



PhD Programme
Foreign Languages, Literatures and Cultures

XXXVIII CYCLE

Hesitations in Original and Dubbed Film Speech
A Diachronic Study

Candidate: Ludovica Calogero

Supervisor: Prof. Serenella Zanotti

Co-supervisor: Prof. Sofia Sánchez Mompeán

Acknowledgements

There are many people I wish to thank and acknowledge. First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Serenella Zanotti, for her kind academic guidance and constant availability throughout the course of this work. Her support and insight have been fundamental in shaping this research, helping me develop a rigorous methodological approach from what were initially sometimes unstructured ideas. I am equally grateful to my co-supervisor, Professor Sofía Sánchez-Mompeán, for enriching my study with valuable tools and thoughtful feedback. I would also like to thank her for making me feel truly welcome during my research period at the University of Murcia, offering not only academic guidance but also personal support throughout my stay.

I am deeply thankful to Professor Charlotte Bosseaux for welcoming me at the University of Edinburgh for my research visit. Her insightful questions helped me organise my thoughts and make decisions during moments of uncertainty. My sincere thanks are also due to the external reviewers, Rocío Baños Piñero and Giselle Spiteri Miggiani, whose comments and suggestions helped improve this work. I would also like to thank Veronica Bonsignori, Sergio Conti and Viviana Masia, who generously offered their expertise, taking the time to respond to my emails or to meet with me when I sought their advice.

The process of writing this dissertation would have been far more challenging and demanding without the backing of my PhD colleagues, or I should say friends. It is not something to be taken for granted to find such a strong support system. In particular, thanks to Andrea, Kristina and Valeria: sharing this journey with you has been a gift.

I wish to thank my family and friends for patiently listening to me talk — and often complain — about this project, sometimes even pretending to understand what I was actually talking about. Thank you for supporting me not only in moments of joy and accomplishment but also during times of stress and exhaustion. I will not name all of you individually (for reasons of space), but I am thinking of each of you at this moment. To Camilla, thank you for helping me understand computer language. To Valerio, thank you for encouraging me to rest without feeling guilty, for helping me recognise the value of my work on days when I feel like I have done nothing, and for pushing me to send the emails I am most afraid of. Thank you for making me feel, even in the simplest things, that you are always proud of me and of my work. Last but not least, I wish to acknowledge those who are no longer here to celebrate this moment. I know you would have been proud and perhaps even more excited than I am about this academic achievement.

Table of Contents

List of Figures	8
List of Tables	9
Introduction	10
Chapter One - Hesitations: A Theoretical Background	16
1.1. Investigating Hesitations: An Introduction	16
1.2. Laying the Foundations: Pioneering Studies on Pauses and Hesitations	17
1.2.1. The Correlation between Location, Functions and Internal Processes	19
1.2.2. From a Word-Based Model to a Syntactic and Phonemic Clause-Based Model ..	22
1.3. Toward a Definition of Hesitations	24
1.4. Hesitations in Speech and their Basic Meaning(s)	27
1.4.1. Functional Categories of Hesitations: Cognitive, Repair and Interpersonal Functions	27
1.4.2. The Basic Meaning of Hesitations: From Delays to Pragmatic Markers	29
1.4.3. Hesitations in Discourse Interactions as Turn-Management Tools	32
1.5. The Multifunctional Framework of Hesitations	34
1.5.1. The Prosodic and Paralinguistic Nature of Hesitations	36
1.6. Hesitations in Conversation: Sociodemographic Factors	37
1.7. The Perception of Hesitations and their Impact on Listeners and Speakers	40
1.7.1. Hesitations and the Impact on the Perception of the Speaker	40
1.7.2. Hesitations and the Impact on Language Comprehension	42
1.8. Hesitations in Writing and in Fiction	44
1.8.1. Hesitations in Writing	44
1.8.1.1. Categorizing Hesitations: Are they Words?	47
1.8.2. Hesitations in Fiction	49
1.9. Hesitations in the Italian Language	53
1.9.1. The Phonetic Pattern of Hesitations in Italian	56
1.9.2. The Effect of Hesitations on Naturalness in the Italian Language	58
1.10. Further Remarks on the Analysis of Hesitations in Film Speech	59
Chapter Two - Audiovisual Dialogue and Dubbing	63
2.1. The Hybrid Nature of Film Dialogue	63
2.2. Film Speech and Orality: A Structural and Functional Analysis	65
2.2.1. Orality and Spoken Language in (Non-)Fictional Dialogue	65
2.2.2. The Communicative Structure of Film Dialogue	67
2.2.3. A Functional Analysis of Telecinematic Dialogue	71

2.3. Scripting and Performing Orality: A Matter of Compromise.....	73
2.4. The Linguistic Features of Audiovisual Dialogue.....	75
2.4.1. Corpus Linguistics Applied to Audiovisual Dialogue	77
2.4.2. The Oral Register of Telecinematic Dialogue	79
2.5. Dubbing Prefabricated Orality in Audiovisual Texts	85
2.5.1. The Dubbing Process: Technical and Non-Technical Constraints	86
2.5.2. The Language of Dubbing: Dubbese and Dubbitis.....	90
2.6. Corpus Linguistics Approaches to AVT: Systematic Linguistic and Translational Behaviours.....	95
2.7. The Translation of Orality Features in Dubbing	100
2.7.1. The Omission of Conversational, Informal and Pragmatic Markers	100
2.7.2. The Influence of the Source Language and the Standardization of Orality Features	103
2.7.3. The Use of Translation Routines	106
2.8. Dubbing Hesitations	110
2.9. Diachronic Studies on Film Speech, Prefabricated Orality and Dubbing	114
2.9.1. The Evolution of Film Speech over Time: Previous Research.....	114
2.9.2. Diachronic Studies on Dubbing and Orality Features	117
Chapter Three - Corpus and Methodology	124
3.1. The Analysis of Hesitations in Cinematic Dialogues.....	124
3.1.1. Objectives and Research Questions	126
3.2. Building the Corpus: Steps and Criteria	127
3.2.1. Language.....	128
3.2.2. Country	129
3.2.3. Year of Release	130
3.2.4. Film Genre	132
3.2.5. Movie Length.....	135
3.2.6. Narrative Aspects.....	135
3.2.7. The Diachronic Corpus of Film Dialogue	136
3.3. Criteria for Soundtrack Transcription.....	140
3.3.1. Transcribing Hesitations	149
3.4. Investigating Hesitations in a Corpus of Film Dialogue	150
3.5. The Location of Hesitations in Film Speech	153
3.6. A Functional Analysis of Hesitations in Film Speech.....	156
3.6.1. The Structural Domain.....	158
3.6.2. The Planning Domain	159
3.6.3. The Interpersonal Domain	161
3.6.4. The Cognitive Domain.....	163
3.6.5. The Hesitation Proper Domain	163
3.6.6. Further Remarks on the Functional Categorization of Hesitations.....	164
3.7. Investigating Hesitations in a Corpus of Dubbed Film Dialogue: Translation Strategies	166
3.8. Methodological Tools: Additional Software and the Reference Corpus.....	168
3.8.1. Praat 6.2.23	168
3.8.2. AntConc 4.0	170

3.8.3. RStudio 25.09.01+401	171
3.8.4. Using the <i>Movie Corpus</i> as a Reference Corpus	175
Chapter Four - Hesitations in Anglo-American Film Dialogue: A Diachronic Analysis	178
4.1. Hesitations in Film Speech over Time	178
4.1.1. Hesitation in Film Speech over Time: A Quantitative Analysis	178
4.1.2. Hesitations and the Genre Variable: A Diachronic Analysis	183
4.1.3. The Evolution of Hesitations in Film Speech: A Discussion of the Quantitative Data	189
4.2. Hesitations in Film Speech over Time: A Functional Analysis	192
4.2.1. The Location of Hesitations in Film Speech: A Turn-Based Analysis	192
4.2.2. The Location of Hesitations in the Turn: A Functional Analysis.....	194
4.2.2.1. Hesitations in Film Speech: The Occurrence of Hesitations at the Beginning of the Turn.....	195
4.2.2.2. Hesitations in Film Speech: The Occurrence of Hesitations in the Middle of the Turn.....	196
4.2.2.3. Hesitations in Film Speech: The Occurrence of Hesitations at the End of the Turn.....	198
4.2.2.4. Hesitations in Film Speech: The Occurrence of Hesitations as Free-Standing Items.....	199
4.3. The Multifunctional Use of Hesitations in Film Speech: Examples from the Corpus.....	200
4.3.1. Hesitations in Film Speech: The Structural Domain.....	201
4.3.2. Hesitations in Film Speech: The Planning Domain.....	208
4.3.3. Hesitations in Film Speech: The Interpersonal Domain.....	220
4.3.4. Hesitations in Film Speech: The Hesitation Proper Domain.....	227
4.3.5. Hesitations in Film Speech: The Cognitive Domain.....	233
4.4. A Corpus-Linguistic Approach to Hesitations in Film Speech: Collocation and Concordancing	236
4.4.1. Collocates of Hesitations in the Corpus	236
4.4.2. KWIC (Key-Word-In-Context) Tool.....	239
Chapter Five - Hesitations in Italian Dubbing: A Diachronic Analysis	243
5.1. Hesitations in Italian Dubbing over Time	243
5.1.1. Hesitations in Italian Dubbing over Time: A Quantitative Analysis.....	243
5.1.2. Dubbing Hesitations and the Genre Variable: A Diachronic Analysis	248
5.2. Dubbing Hesitations into Italian: A Diachronic Analysis of Translation Strategies .	253
5.3. Dubbing Hesitations into Italian: Examples from the Corpus	255
5.3.1. Dubbing Hesitations: Close Rendering	255
5.3.2. Dubbing Hesitations: Omission.....	263
5.3.3. Dubbing Hesitations: Addition.....	269
5.3.4. Dubbing Hesitations: Substitution.....	277
Conclusions	289
References	301

Sitography	330
Filmography	331
APPENDIX A - List of Conventions in the Prosodic Transcription	333
APPENDIX B - Normalized Frequencies of Hesitations per Decade	334
APPENDIX C - Hesitations in English Originals	335
APPENDIX D - Normalized Frequencies of Hesitations in Italian Dubbing per Decade	336
APPENDIX E - Hesitations in Italian Dubbing	337
APPENDIX F - Conditional Inference Tree (CTree)	338
APPENDIX G - Analysis of Variance (ANOVA)	339

List of Figures

Figure 3.1. Spoken Words by Genre	134
Figure 3.2. Example of an Excel Spreadsheet from <i>Sideways</i> (2004)	151
Figure 3.3. Example of a Praat Screenshot from <i>Mississippi Burning</i> (1988).....	169
Figure 3.4. Example of Excel Spreadsheet for Analyses on RStudio (<i>dataITA</i> file).....	173
Figure 3.5. Example of ANOVA Table Retrieved from Gries (2009: 278).....	175
Figure 4.1. Normalized Frequencies of Hesitations in Film Speech over the Years (1,000 words).....	182
Figure 4.2. Hesitations in Comedies and Non-Comedies over the Years	185
Figure 4.3. Variation of Hesitations in Comedies and in Non-Comedies (1950s-2010s). 187	
Figure 4.4. Functional Domains of Hesitation in the Start-Turn Category (Raw Frequencies and Percentages).....	196
Figure 4.5. Functional Domains of Hesitations in the Intra-Turn Category (Raw Frequencies and Percentages).....	197
Figure 4.6. Functional Domains of Hesitations in the End-Turn Category (Raw Frequencies and Percentages).....	199
Figure 4.7. Functional Domains of Hesitations in the Single-Turn Category (Raw Frequencies and Percentages)	200
Figure 4.8. Collocates of Hesitation Markers in the Corpus (4L-4R).....	237
Figure 4.9. Collocates of Hesitation Markers in the Corpus (2L-2R).....	237
Figure 4.10. Twenty-Five Random Instances for <i>hes</i> Sorted to Right.....	240
Figure 4.11. Twenty-Five Random Instances for <i>hes</i> Sorted to Left	242
Figure 5.1. Normalized Frequencies of Hesitations in Italian Dubbing over the Years (1,000 words).....	246
Figure 5.2. Hesitations in Dubbed Comedies and Non-Comedies over the Years	249
Figure 5.3. Variation of Hesitations in Dubbed Comedies and Non-Comedies (1950s-2010s)	251
Figure 5.4. Translation Strategies in Italian Dubbing	254
Figure 5.5. The Use of the Close Rendering Strategy over Time	256
Figure 5.6. The Use of the Omission Strategy over Time.....	263
Figure 5.7. The Use of the Addition Strategy over Time.....	269
Figure 5.8. The Use of the Substitution Strategy over Time.....	277
Figure 5.9. Waveform of <i>di</i> in Extract 52	286
Figure 5.10. Waveform of <i>perché</i> in Extract 53	287
Figure 5.11. Waveform of <i>è</i> in Extract 54.....	288

List of Tables

Table 2.1. Comparable Data from Veirano Pinto (2014) and Zago (2016)	116
Table 3.1. The Diachronic Corpus of Film Dialogue	139
Table 3.2. Sub-Sections of the Diachronic Corpus of Film Dialogue	149
Table 3.3. The Functional Framework	158
Table 4.1. Tokens of Hesitations in the Anglophone Section of the Corpus	178
Table 4.2. Raw Frequencies of Hesitations per Decade	180
Table 4.3. Normalized Frequencies of Hesitations per Decade (1,000 words)	180
Table 4.4. Normalized Frequencies of Hesitations in the Corpus (1,000 words)	183
Table 4.5. Percentage Variation of Hesitations in Comedies and Non-comedies	186
Table 4.6. Average Length of Films per Decades	190
Table 4.7. Hesitations in the <i>Movie Corpus</i> over Time	191
Table 4.8. Number of Hesitations in Different Locations	193
Table 5.1. Tokens of Hesitations in the Italian Section of the Corpus Compared to the English Section	243
Table 5.2. Raw Frequencies and Percentages of Hesitations per Decade in Italian Dubbing	245
Table 5.3. Normalized Frequencies of Hesitations per Decade in Italian Dubbing (1,000 words)	245
Table 5.4. Normalized Frequencies of Hesitations in Italian Dubbing (1,000 words)	248
Table 5.5. Percentage Variation of Hesitation in Dubbed Comedies and Non-Comedies	252

Introduction

Audiovisual Translation (AVT) has grown remarkably over the last twenty years (Chaume 2018a; Zanotti 2022), to the extent that it is now regarded as the most widespread mode of translation (Bosseaux 2015:1) as a result of the proliferation of foreign audiovisual products — i.e. movies, TV series and programs, video games and social media content — distributed internationally. In this respect, new technologies and streaming platforms have changed the distribution and watching patterns of these products (Sánchez-Mompeán & Zanotti 2025), altering the traditional audiovisual map divided into “subtitling”, “dubbing” and “voice-over” countries. In today’s rapidly evolving landscape, AVT has come to encompass a wide “range of transfer methods used in dealing with both long-standing and emergent forms of multimodal and multimedia meaning-making practices” (Zanotti 2022: 441). In spite of the new or more recent forms of AVT such as “dubtitles” (Sokoli 2024), “cyberdubbing” and “cybersubtitling” (Baños Piñero & Díaz-Cintas 2024), this study focuses on a well-established translation mode, i.e. dubbing.

Dubbing is a post-production process that involves the substitution of the original dialogue soundtrack with one translated and re-recorded into a given target language (Chaume 2012; Ranzato & Zanotti 2019: 1); the soundtrack is thus linguistically recreated and synchronized to the images on the screen. This type of translation still represents one of the preferred modalities of fruition of audiovisual content, increasingly adopted not only in countries like Italy that have always heavily depended on this translation mode, but also in new contexts – for example, for film distribution in Russia, a traditionally voice-over country (Ranzato & Zanotti 2019: 4). Dubbing is also becoming more common in Anglophone countries as a direct outcome of the increase of foreign products distributed via streaming platforms, i.e. viewers coming from a non-dubbing background are now exposed to non-English language productions being dubbed into English (Sánchez-Mompeán 2023a: 4; Sánchez-Mompeán & Zanotti 2025: 4). In this respect, scholars have been talking about “a dubbing revolution” (Ranzato & Zanotti 2019) as the result of a marketing strategy carried out by streaming platforms, especially Netflix, to expand the number of subscribers around the world and to increase profits (Roxborough 2019: online). As a matter of fact, dubbing represents the largest portion of revenue of internationally distributed audiovisual products (see Ranzato & Zanotti 2019: 2). This had already been remarked by Romero-Fresco (2013) more than 10 years ago: the author highlighted how, despite the minimal budget allocated

for film translation (between 0.1% and 1%), over half the revenue obtained by top-grossing and award-winning Hollywood films distributed between 2001 and 2011 came from foreign markets. In other words, it is the dubbed or subtitled version of a movie distributed internationally that brings the highest profits (*ibidem*: 202-203).

As pointed out by Sisto (2017: 393), with dubbing, “[t]he resulting vocal exchange has profound significance for the film, as it affects the body and performance of the actor, the rhythm and flow of the language, the communicative texture and timbre of the voice, and the coherence of voice and gesture”. These factors shape, and potentially alter, the viewing experience of international audiences, who incorporate these translated products into their own culture. As claimed by Cornu (2019b: 295):

The film culture of a nation has often been nurtured not only by its home production, but also — and sometimes to a very large extent — by films produced in other countries, the international influence of American cinema being one of the most striking examples of this. [...] Viewers benefit from this through the use of translated versions of the ‘foreign’ films, and often consider such films as part of their own culture and, even, heritage.

In this respect, Toury (1995: 18) claimed that translated texts should be regarded “as facts of the culture that would host them, with the concomitant assumption that whatever their function and systemic status, these are constituted within the target culture and reflect its own systemic constellation”. In the context of AVT, it is important to consider the implications this statement entails since, while the verbal message is transformed and adapted (most commonly, through dubbing or subtitling), images remain the same; hence, they represent a link to the source culture (Ranzato 2010: 13-14). Audiovisual texts are multimodal in the sense that they involve multiple channels of communication such as the acoustic and the visual ones (Díaz Cintas 2009: 9; Chaume 2012: 100). It is the interaction of these channels and their signifying codes that creates meaning; the linguistic code is only one of the multiple semiotic resources involved in these types of texts (Chaume 2004a). Nevertheless, language certainly is one of the central elements that allow to explain the unfolding of events (Kozloff 2000; Pavesi 2005; Bednarek 2018). The centrality of the verbal component in motion pictures has also been highlighted by Cutting (2023), who claims that “conversations have [...] become the core of popular cinema” (*ibidem*: 320). This has also been remarked by Cutting & Candan’s (2015) findings, according to which, in most movies, over half of the shots include some form of conversation. Accordingly, this research

will focus on the verbal component of films — that is, dialogue and its dubbed translation — while also considering how it interacts with non-verbal elements.

As pointed out by Romero-Fresco (2009b: 41), studies conducted on AVT seem to include two basic notions: the independence of audiovisual translation as an autonomous discipline, on the one hand, and its dependence on other related disciplines like Film Studies, on the other. In this regard, “[a]udiovisual texts are usually built according to the conventions of film language, a complex language that overcomes linguistic communication and has its own rules and conventions” (Chaume 2004a: 12). These norms make the language of films a “cross-linguistically stable register” (Zago 2019: 239), characterized by highly predictable features (Taylor 2008). This study aims to test this, by examining some of the conventional conversational features of the language of movies.

Most notably, particular attention will be paid to the fact that film dialogue is written with the intention of mirroring spontaneous spoken dialogue, as if it was not originally created in the written form of a script. In other words, this research aims to investigate “prefabricated orality” (Chaume 2004b; Baños Piñero & Chaume 2009; Baños Piñero 2024), i.e. key features of impromptu speech inserted in scripted dialogues to reproduce day-to-day conversational interactions and convey a sense of naturalness and spontaneity to viewers. As noted by Pavesi & Zanotti (2023: 167-168), “as audiovisual dialogue represents face-to-face communication on screen, it calls for the pervasive expression of interpersonal meanings”. These meanings then must be translated and adapted for international audiences, based on their cultural norms and expectations regarding orality and spokenness, and on the constraints of the selected translation mode (i.e. dubbing, subtitling, or voice-over). As mentioned above, this research revolves around cinematic products adapted for an Italian audience through dubbing; in particular, “[t]hanks to its verbal and prosodic richness to mimic the voices coming from the source texts, [dubbing] is indeed a privileged site of investigation of fictive orality, in all its linguistic and pragmatic conventions” (Ranzato & Zanotti 2019: 7). In this respect, the present study specifically focuses on pragmatic markers by examining a phenomenon hitherto insufficiently researched in AVT, i.e. hesitations.

Hesitations are multifunctional features that frequently occur in spontaneous conversation as the result of the complex processes carried out by speakers engaged in conversations, who must both plan and perform their utterances in real time (Biber *et al.* 1999: 1048). In the context of AVT, hesitations are among the stylistic devices inserted in dialogues to convey a sense of orality and spontaneity to the text. Over the years, a growing tendency to include linguistic items typical of oral speech has been observed in film scripts

(Valdeón 2009: 199; Pavesi 2005, 2009b), making the language of audiovisual products increasingly more emotional and informal (Werner 2021: 64). However, when translating, orality features like hesitations tend to be overlooked, reduced and simplified; this happens because their rendering, or lack of rendering, is considered not to alter the overall structure of the story (Bruti 2019: 197). This study aims to test this through an examination specifically focused on hesitations.

In AVT, hesitations have traditionally been studied in co-occurrence with or among a wide variety of features of spontaneous language — such as interjections, attention signals and disfluency phenomena (Romero-Fresco 2007; Valdeón 2008; Bruti 2019) — or through case studies based on a small corpus of films (Calogero 2024) that do not allow to make generalizations on the way these elements are treated in both original and dubbed film speech. Based on these premises, this study aims to investigate whether an evolution with respect to the way orality features are codified in film speech can be noticed, examining the degree of similarity between film speech and impromptu conversation, while also raising awareness on the role played by these often-overlooked features in pre-planned conversations. Moreover, the study also endeavours to examine the phenomenon from a translational point of view, i.e. offering an overview of the procedures chosen to render hesitations in Italian dubbing over the years. In order to do so, the phenomenon will be studied from a diachronic perspective (from 1950 to 2019), examining fictional dialogues originally created in the English language and then dubbed into Italian.

The study is organized in five chapters: Chapter One provides the theoretical framework for the investigation of hesitations. More specifically, this chapter aims to draw a picture of the phenomenon in spontaneous conversation, i.e. their distribution within the sentence, as well as the meanings and the functions they commonly fulfil in spontaneous speech. This is achieved by reviewing studies conducted on the phenomenon in various fields, namely linguistics, psycholinguistics, cognitive studies and pragmatics, as well as discourse and speech analysis. This chapter will thus clarify the type of elements included among the ambiguous label of “hesitation markers”, illustrating their characteristic features, as well as some of the difficulties and constraints involved in the investigation of these pragmatic items. Having reviewed the way hesitations are employed in spontaneous speech, it will be possible to compare their occurrences in pre-planned dialogues; namely, original English-language movies adapted for an Italian audience through dubbing.

Chapter Two specifically focuses on audiovisual dialogue. Although this dissertation mainly concerns film speech, this chapter draws upon studies conducted on both cinematic

and television dialogue with the purpose of describing the complex nature of audiovisual texts, the participation framework and the multiple functions dialogue fulfils. Most notably, audiovisual dialogue is examined in relation to its adherence to the codes of naturalness and realism. The chapter will address the way orality is codified in fictional products and how it is conveyed in both original and dubbed versions. In particular, this chapter draws attention to the typicalities of dubbed products by focusing on dubbing orality features and, most notably, hesitations. The discussions presented here also allow to pinpoint some of the insights obtained with regard to the evolution of audiovisual language in relation to its adherence to spontaneous speech, alongside the diachronic evolution of Italian dubbing over the years.

Having presented the phenomenon under investigation (i.e. hesitations) and the specific context in which it will be examined (i.e. audiovisual dialogues), Chapter Three aims to describe the criteria and methodology used to build and examine the corpus of film speech developed for the analysis. Firstly, the corpus will be presented by focusing on the rationale followed for the selection of films. Then, drawing upon the field of Corpus Pragmatics, the methodology used to survey hesitations within the corpus will be described, explaining the way these features have been allocated to different locations within the turn, interpreted based on the Functional Framework developed for the investigation, and examined from a translational point of view. Moreover, insights are provided concerning software and computer tools used to carry out the empirical analysis of hesitations such as Praat 6.2.23, AntConc 4.0 and RStudio 25.09.01+401.

Results are reported in Chapter Four and Chapter Five: the former describes the data obtained from the analysis of hesitations in the English section of the corpus. First, the diachronic quantitative data are presented and described; then, the chapter will focus on the results obtained from the qualitative analysis, that is, a review of the way hesitations are employed in film speech with respect to their location, their functions and recurrent patterns. In particular, the results obtained from the investigation of English film speech are compared with the considerations made with regard to hesitations in spontaneous conversations reported in Chapter One.

Finally, Chapter Five focuses on the results derived from the Italian section of the corpus. In this chapter, the comparison between the English and the Italian dubbed soundtracks aims to stimulate a contrastive analysis with a particular focus on the rendering of the pragmatic elements under investigation. The discussion will thus describe the data obtained from the quantitative analysis by comparing them with those extracted from the

source English texts. Subsequently, a qualitative analysis of the translation strategies used to treat hesitations in Italian dubbing is presented.

Importantly, the comparative processes carried out in the analysis (between film speech and spontaneous conversation, on the one hand, and between the source and the target text, on the other) have been executed taking into account various factors at the macrostructural level, that is, the text as a whole and the historical context in which both the source text and the target text are produced, i.e. how social and political changes impacted the cinematic and the dubbing industry and the linguistic preferences specific to each period. In so doing, the purpose of the study is to examine the mechanisms of codification of orality in English-language film speech with respect to hesitation markers and the evolution of film speech as regards orality features. At the same time, the aim is to contribute to AVT research by investigating dubbing's past in order to better comprehend today's and tomorrow's future developments. The ultimate objective is to add to existing AVT research by highlighting further the complex nature of audiovisual texts and their adaptation process, with a specific focus on the adaptation of pragmatic markers. As a matter of fact, in a context where technological innovation and AI are said to jeopardize the professional role of both translators and dubbing actors, the study of these pragmatic markers brings the human element to the forefront, drawing attention to the role of these elements in the actors' performances as conveying a variety of meanings and affecting the overall pragmatic interpretation of both original and dubbed speech.

Chapter One

Hesitations: A Theoretical Background

The fundamental reason for hesitating
is that speech production is an act of creation.
(Chafe 1980: 170)

The aim of this chapter is to provide a theoretical framework for the phenomenon under investigation, i.e. hesitations. By reviewing studies conducted in various fields — namely linguistics, psycholinguistics, cognitive studies, pragmatics, as well as discourse and speech analysis — it aims to draw a picture of hesitations in spontaneous spoken language. The objective is to observe their distribution, functions and meanings in speech to then compare their occurrences in pre-planned film dialogues, specifically in original English movies adapted for an Italian audience through dubbing. The discussions presented in the chapter also highlight some of the difficulties and constraints involved in the investigation of the hesitation phenomena examined here, such as vocalizations, lengthenings, and pauses of silence.

1.1. Investigating Hesitations: An Introduction

Hesitations are natural occurrences in spontaneous speech, resulting from the complex processes carried out by speakers engaged in conversations who must both “plan and [...] execute their utterances in real time” (Biber *et al.* 1999: 1048). Spontaneous speech involves dynamic exchanges in which interlocutors have to coordinate their planning and formulation of words to construct verbal meanings (Clark 2004). At the same time, speakers also need to adapt and adjust their utterances to the needs of the interlocutor(s) (Biber *et al.* 1999: 1045).

As stated by Goldman-Eisler (1958: 61), “the nature of speech production is such that a continuous flow of verbal output is only rarely achieved. In most cases, utterances are series of verbal productions differing in length and broken up into discrete elements by pauses of varying duration”. In other words, some of the linguistic behaviour and

conversational traits typical of spontaneous speech emerge from the speakers' constant need to encode and decode meanings (Voghera 2017). Since speakers only have a limited amount of time to think about and plan their speech, phenomena like repetitions and rephrases, but also pauses and hesitations, commonly occur (Biber *et al.* 1999: 1048; Carter & McCarthy 2006: 168). As pointed out by Rochester (1973: 51), “[w]hen words are uttered spontaneously (i.e. when they are being organized at the time of utterance), they emerge haltingly, interspersed with *ums* and *ahs* and *ers*”.

Since the second half of the twentieth century, the linguistic behaviours typical of spontaneous speech have attracted the attention of scholars from different fields. The study of speech and nonsegmental systems had long been overlooked by researchers. This was due, on the one hand, to the elite status of written language and the consequent rejection of spontaneous conversation as a proper subject for linguistic theories (Clark 2004: 382); on the other, it was difficult to study oral utterances until the technologies to record actual conversations were made available (Rochester 1973: 52). The pioneers of pragmatics started studying language from invented interactions of pre-planned and non-interactive forms of language coming from novels and plays; many then turned to spontaneous language for evidence, but data were primarily based on clean and invented examples that did not reflect the spontaneous and interactive nature of real language use (Clark 2004: 365). It was not until the 1950s that researchers started to investigate the decision-making processes involved in speech production, giving attention to pausal phenomena (Rochester 1973: 52).

1.2. Laying the Foundations: Pioneering Studies on Pauses and Hesitations

The first descriptive and experimental studies on the duration and distribution of pauses and hesitations in speech were carried out in the field of psychology and psycholinguistics. These studies provided early hypotheses and insights regarding the process of speech encoding and the psychological significance of hesitations in speakers' speech.

In particular, Goldman-Eisler (1954, 1957, 1958) set a precedent for the investigation of pauses in speech and their functions. Starting from the analysis of the rate of speech production (1954)¹ — defined as the output of speech measured in syllables per unit time (minutes) (Goldman-Eisler 1958: 61) — the author investigated the gaps occurring in the stream of speech. In fact, speech rate appears to be highly influenced by the frequency and

¹ The study was conducted on twelve elements of speech. Some of these elements were linguistic (parts of speech), others were temporal and included the durations and frequencies of periods of talk and periods of silence.

the duration of pauses (*ibidem*: 63). Of course, gaps occurring in speech can vary in nature, and a distinction must be established between hesitations and pauses due to the biological need of the speaker to stop at regular intervals to breathe (*ibidem*: 64). While breathing and speech production are functionally related, the occurrence of hesitation pauses in conversation is not as functionally obvious (*ibidem* 65-66), since hesitations seem to interfere with “the structural cohesion of syntactic grouping and thus with the intelligibility of the message” (*ibidem*: 66). Interestingly, the dichotomy is not between speech and non-speech, but between a certain activity — which could be speech or breathing — and its delay through hesitations (*ibidem*). As already articulated by Lounsbury (1954), the author supports the idea that hesitation pauses are manifestations of the blocking of an activity when speakers are in situations of uncertainty such as the organization of speech or the selection of the next step (Goldman-Eisler 1958: 66-67).

These hypotheses found some validation in experimental evidence showing the correlation between hesitations and an increase of information (Goldman-Eisler 1957). In fluent speech, speakers often use routine combinations of words shared by the members of a language community. When speakers diverge from these common and almost automatic sequences in favour of a highly individual and unexpected choice, the speech is typically more hesitant (Goldman-Eisler 1958: 67). The words following the pause are expected to be less probable or predictable; in other words, they carry an increase of information (Goldman-Eisler 1957: 1497). The hypothesis was tested through two sets of experiments investigating the relationship between pauses and transitional probabilities. Results showed that redundancy² is closely associated with fluency, while uncertainty (less predictable words) is closely associated with hesitancy. This seems to confirm the hypothesis that pauses in speech anticipate an increase of information (*ibidem*).

All things considered, pausing seems to be a condition for the thinking process taking place when organizing speech. Of course, the speech rate and the rate of hesitation pauses are influenced by a series of factors such as the speaking situation. Speakers, in fact, mix some pre-organized automatic combinations of words with some newly organized sequences, which are tailored to convey specific meanings to the interlocutor and adjusted to the social demands of the communicative situation (*ibidem*). As also underlined by

² The term redundancy indicates “a measure of the predictability of a message or of the limitation of choice in selecting the symbols of a message: a situation characterized by high redundancy exists when, in a given position in a sequence, the range of elements which can occur is limited by the preceding elements and when this range includes items with unequal probabilities of occurrence” (Kahane & Pietrangeli 1959: 167).

Rochester (1973: 75), “pauses and other phenomena of spontaneous speech should be functionally related to changes in the interpersonal situation and/or changes in the responsiveness of the speaker”. Of course, individual factors also play a role, as for individual preferences and the speaking style of subjects³.

Despite the individual differences, is it possible to find regular patterns in the occurrence of hesitations in speech? Psycholinguists have tried to answer this question by observing whether there is some correlation between the location of hesitations in spoken utterances and their functions.

1.2.1. The Correlation between Location, Functions and Internal Processes

Maclay & Osgood (1959) followed the path of previous investigations with the objective of determining whether it is possible to find some regular patterns in the occurrence of hesitations through a descriptive study of utterances produced during the speech of twelve participants in a conference. The aim was to observe the frequency of hesitations and their location in the sentence in relation to the grammatical function of the words they co-occurred with. Four hesitation types were examined: repeats, false starts, filled and unfilled pauses. Filled pauses included the “occurrences of the English hesitation devices [ɛ, æ, ɪ, ə, m]”, among which schwa (ə) was the most frequent, while unfilled pauses included silences of unusual length and the non-phonemic lengthening of phonemes (*ibidem*: 24). Considering the interests of the present research, the results discussed here will only refer to filled and unfilled pauses.

The data show that the majority of both filled and unfilled pauses occur at word boundaries, falling between words of any class (*ibidem*: 27). Both filled and unfilled pauses tend to occur more often before lexical words than before function words (*ibidem*: 32-33); however, there are some differences regarding their distribution within the phrase. Filled pauses are more likely to occur before function words, whereas unfilled pauses are more likely to occur before lexical words (*ibidem*: 33). Consequently, for the constructions statistically analysed in the study, filled pauses occurred more frequently at phrase boundaries, while unfilled pauses tended to occur at word boundaries within phrases (*ibidem*: 34). As a matter of fact, phrases usually start with function words, while lexical words tend

³ The differences between subjects’ style of speech and the way in which individual preferences impact on the selection of hesitation phenomena have also been discussed in Maclay & Osgood (1959: 38-39) and in Blankenship & Kay (1964), as well as in multiple studies developed since then, which will be examined later in the chapter.

to occur within the phrase. Since speakers often start talking before having made all the decisions regarding what they are going to say, filled and unfilled pauses may occur at points where the speaker has to make decisions on how to continue. Filled pauses are expected to be found at points of uncertainty when the author presumably has to make constructional decisions regarding the content of the utterance (*ibidem*: 41), while brief pauses of silence may occur before lexical words as an indication of the undergoing selection process among a wide set of lexical alternatives (*ibidem*: 42).

The conclusions reported in Maclay & Osgood's (1959) descriptive study have then been taken as a starting point for several investigations about the psychological significance of pauses (Goldman-Eisler 1961) and their distribution within the sentence (Blankenship & Kay 1964). Goldman-Eisler (1961) investigated hesitations through a comparative experimental study consisting of two conditions: firstly, speakers had to talk about concrete situations, such as telling the sequence of events of a story; then, speakers were asked to summarize the meaning of the story and make generalizations (i.e. abstract situations) (*ibidem*: 19). The experiment not only allowed the researcher to measure and compare the duration of both types of hesitations (i.e. filled and unfilled pauses), but also to examine individuals' consistency throughout the phases of the experiment and the psychological significance of pauses in relation to the two different tasks. Both filled and unfilled pauses seemed to contain a habitual factor; however, deviations from the habitual rate were detected. Unlike unfilled pauses, deviations from the habitual rate of filled pauses seemed to be related to the content of the story and not to cognitive factors such as the abstraction of the speech or difficulty of choice. This may indicate an emotional factor. Consequently, the two hesitation phenomena could reflect different internal processes: silent pauses seem to reflect responses to more challenging cognitive tasks, while filled pauses seem to reflect the emotional attitude of the speaker (*ibidem*: 25).

Based on the assumption made by Maclay & Osgood (1959: 40) that hesitation phenomena are related to grammatical and lexical selection, Blankenship & Kay (1964) suggest that the occurrence of hesitations falls more frequently before structural units than before lexical choices (*ibidem*: 369). In particular, the authors observed seven hesitation phenomena in American English speech, i.e. non-lexical intrusive sounds (which coincide with what has so far been defined as filled pauses), sentence correction, word change, repeat, stutter, omission of a part of a word, and sentence incompleteness. Interestingly, the authors did not include two other possible types of hesitation phenomena because they are closely related to the individual style of the speaker, i.e. non-phonemic lengthening of phonemes

and unfilled pauses. For the purposes of the present study, particular attention will be given to the category of non-lexical intrusive sounds, which will be referred to as “filled pauses”. Their results show that the category of filled pauses generally locates at word boundaries and they seldom occur at morpheme boundaries; moreover, this category was the largest single type of hesitation measured in the study (*ibidem*: 362). Such data is particularly relevant because it proves how frequent and relevant filled pauses are to speakers in spontaneous speech. In particular, Blankenship & Kay raise the following hypothesis: the higher the number of hesitation phenomena, the more redundant (or predictable) the message (*ibidem*: 370). If, according to Goldman-Eisler (1958: 67), the close relation existing between pauses and information, on the one hand, and fluency of speech and redundancy, on the other, indicates that the insertion of hesitation pauses in speech is a necessary condition for the increase of information, it can also indicate that hesitation phenomena can operate as elements of redundancy in the language system. If hesitation pauses operate as a redundant element and as some kind of ‘signal’ by limiting the range of co-occurrent items in speech, they may also help hearers in the process of understanding the message (Blankenship & Kay: 371). Their view opposes to Lounsbury’s (1954) suggestion that, unlike juncture pauses — i.e. pauses falling at the boundaries of syntactic units (*ibidem*: 98) — which help the hearer’s understanding, hesitation pauses often constitute “an annoyance to the hearer and interfere with, rather than aid in grasping the sentence as a whole” (*ibidem*: 99).

The difference between juncture and hesitation pauses will later be reconsidered by Barik (1968) in the light of Boomer’s (1965) study on the location of hesitations in the phonemic clause (further discussed in §1.2.2). Boomer’s results also provide interesting data regarding the duration of filled and unfilled pauses and, on the basis of findings regarding the minimum length for a pause to be identified as a hesitation, Barik (1968) suggests that hesitations occurring between two phonemic clauses are not necessarily either juncture pauses or hesitation pauses; they may be a combination of the two⁴. In other words, pauses

⁴ Boomer (1965) claims that juncture pauses — i.e. those pauses occurring at the end of phonemic clauses made for the listener’s understanding — are facultative accompaniments of terminal junctures and do not have the same extra-linguistic function of hesitation pauses (Barik 1968: 156). According to Barik, if juncture pauses are free of extra-linguistic functions and only serve to delimit an encoded speech unit, their duration should be shorter than the duration of a hesitation pause, denoting that the speaker is conducting a searching and encoding operation (*ibidem*: 157). This hypothesis seems to go against Boomer’s findings. Boomer (1965), in fact, defined hesitations as filled and unfilled pauses of at least 200 msec. However, nearly half of the junctures in his corpus were not followed by pauses longer than the minimum duration of 200 msec. For those pauses that did meet the pause criterion, the mean duration of the juncture pause was significantly longer than the mean duration of intra-clause hesitation pauses (*ibidem*: 155). Therefore, a big proportion of pauses classified by

between two phonemic clauses may be interpreted as a combination of a juncture pause (referring to the first clause) helping the listener's comprehension and a hesitation pause (referring to the second clause) reflecting the decision-making operation carried out by the speaker.

As can be inferred from the discussion presented above, scholars felt the need to investigate hesitation phenomena in units higher than the word, i.e. at a clause level. The major studies conducted along these lines are further examined hereinafter.

1.2.2. From a Word-Based Model to a Syntactic and Phonemic Clause-Based Model

The considerations discussed thus far have supported the claim that hesitations typically occur at the beginning of an encoding unit, meaning at points where choices and decisions must be made by the speaker. However, it has not been established whether the nature of these units is grammatical or phonological. To determine this, scholars have started moving toward the idea that the analysis of hesitations may require units larger than the word, such as the clause. Boomer (1965), for instance, proposes the adoption of the phonemic clause as the encoding unit, that is, a unit phonologically defined and independent from the presence or absence of hesitations (*ibidem*: 149); on the other hand, Hawkins (1971) relates pauses to the syntactic structure of the utterance, considering the clause as the highest unit of analysis.

Boomer (1965) rejects the idea that hesitations randomly occur in the utterance before any difficult word and supports the hypothesis that hesitation phenomena mainly occur at the beginning of the units composing phonemic clauses. Boomer's data shows that this is true for both filled and unfilled pauses; however, both types of hesitations seem to typically occur not at the very beginning of the clause (i.e. in first position) but rather in second position, after the first word of the clause⁵. These findings require the author to reflect on the theoretical issue regarding the nature of grammatical encoding in speech since the location of filled and unfilled pauses in second position has some implications for what concerns the psycholinguistic process of speech encoding. More specifically, the author's data seem to support the view previously presented by Goldman-Eisler (1958), according to which the selection of the initial word in a phonemic clause circumscribes the sets of possible

Boomer as juncture pauses may represent, at least in part, hesitation pauses associated with the first word in the second clause (Barik 1968: 157).

⁵ In Boomer's (1965) view, these results are strictly related to the decision of the author to consider phonemic units as the encoding units. In fact, a word-transitional model would predict the opposite outcome, since lexical words, which are the more informative elements, usually occur at the end of the phonemic clause (*ibidem*: 155).

constructions the speaker can select afterwards. In line with this view, hesitations occur in second position as they follow an initial decision made by the speaker regarding the structure of the phonemic clause. After that first decision, the speaker can proceed to make lexical choices (Boomer 1965: 156).

Although Boomer's interpretation of results needs to be carefully considered, his findings regarding the location of hesitations in the phonemic clause share some similarities with Hawkins's data concerning the position of hesitations in the syntactic clause. Hawkins (1971) investigated the typical grammatical positions in which pauses tend to occur and described the location of a pause by considering the grammatical function — or functions — of the first word coming immediately after the pause. Results show that the majority of hesitations occur at boundaries between clauses, which is where much of the speech planning occurs (*ibidem*: 285). Among the different grammatical positions studied by Hawkins, the most frequent ones coming after a pause are all associated with the beginning of the clause (*ibidem*: 282). These results are in line with Boomer's (1965) findings regarding the location of hesitations in the phonemic clause. On the basis of the similarities between his and Boomer's findings, Hawkins (1971: 286) suggests that "the two types of clauses are very frequently co-terminous".

To sum up, the preliminary studies conducted on hesitations have provided some general insights regarding their location in speech, as well as the significance of pauses in speakers' production and cognitive processes:

- The occurrence of pauses is subject to cognitive, social and affective-state variables (Goldman-Eisler 1958, 1961; Rochester 1973). Silent pauses, for instance, seem to reflect responses to challenging cognitive tasks, while filled pauses seem to reflect the emotional attitude of the speaker (Goldman-Eisler 1961).
- Pauses constitute a blockage occurring when speakers are in a situation of uncertainty regarding the organization of speech and the selection of the next step (Goldman-Eisler 1958).
- Pauses indicate the speaker's thinking process regarding lexical and/or structural choices (Macley & Osgood 1959; Boomer 1965).
- Hesitations co-occur with an increase of information (Goldman-Eisler 1957, 1958).

- Based on a word-transitional model, both filled and unfilled pauses tend to occur more often before lexical words than before function words; however, within the phrase, filled pauses are more likely to occur before function words and unfilled pauses before lexical words (Maclay & Osgood 1959).
- Hesitations mainly occur at clause boundaries, whether it is a phonetical clause (Boomer 1965) or a syntactical clause (Hawkins 1971).
- In speech, hesitations not only help the process of encoding (Boomer 1965) but also the process of decoding (Blankenship & Kay 1964; Barik 1968).

As pointed out by Rochester (1973), many of the works cited here have encountered methodological issues and criticism⁶. Nevertheless, these studies laid the foundation for linguistic, pragmatic and psycholinguistic lines of research on hesitation phenomena, opening up numerous questions and debates that paved the way for future inquiries, such as: what types of linguistic phenomena can be categorized under the “hesitation” label? How can we categorize hesitations and what is their linguistic status? How do speakers use hesitations in spontaneous conversation? Are there any patterns in the way hesitations are used in speech? Do hesitations impact language comprehension? The various topics these questions touch upon will be discussed in the following paragraphs, beginning with the terminological debate.

1.3. Toward a Definition of Hesitations

The terminology used to refer to the elements studied in this investigation and the word “hesitation” itself represent a problematic aspect since there is no universally agreed-on definition of what hesitations are and which linguistic, paralinguistic and prosodic features can be included under this label. In fact, the label used to refer to the elements under study changes according to the research field, the phenomena considered, and the perspective from which they are observed.

As can be inferred from the studies previously examined (§1.2.1), an issue that needs to be analysed concerns what can be included under the label “hesitations”. Scholars have traditionally used a broad definition of the term, including a wide range of phenomena in

⁶ Boomer (1965: 149), for instance, criticised Goldman-Eisler’s (1968) investigation for being based on a small and highly selective corpus of data; that is, data were based on a biased sample of spontaneous speech including only well-constructed sentences and, consequently, unrepresentative of ordinary speech. Similarly, Maclay and Osgood’s (1959) results have been challenged by Boomer (1965), while Boomer’s findings have been further reviewed by Hawkins (1971).

their research. Maclay & Osgood (1959), for instance, considered not only filled and unfilled pauses, but also repeats and false starts. Similarly, Blankenship & Kay (1964) examined seven hesitation phenomena in American English, among which were non-lexical intrusive sounds, sentence correction, word change, repeat, stutter, omission of a part of a word and sentence incompleteness⁷. Likewise, much variety can also be noticed in more recent studies where the same phenomena are investigated using different naming or transcriptions: some researchers mainly focus on filled pauses like *uh* and *um* (Clark & Fox Tree 2002) or *er* and *erm* (Kjellmer 2003); some focus on what they called “vocalized fillers” (*uh/um*), pauses of silence and lengthenings (Betz & Kosmala 2019), while others also include features like repeats and false starts (O’Connell & Kowal 2005). Gilquin (2008), for instance, identified three categories of hesitation markers: firstly, pauses of silence and pauses filled by vocalizations; secondly, syllable lengthening, truncated words and repetitions⁸; and, thirdly, small words like *kind of* and *well* (see also Boonsuk *et al.* 2019). Similarly, Fox Tree (2007) also examined *um* and *uh* among various discourse markers such as *like* and *you know*.

Although there is general agreement that in English the label “hesitation” should at least include vocalizations consisting of schwa sounds [ə], which can be accompanied by nasalization ([ə̃m]) or lengthening ([ə:], [ə:m]) (Tottie 2011: 173), there is no universal agreement on how these vocalizations ought to be transcribed. As pointed out by Kjellmer (2003: 171), these elements are variously represented in the literature by *ah*, *er*, *erm*, *mm*, *uh*, and *um*; in written American English, they are usually transcribed as *uh* or *u(h)m*, while in British English, which is non-rhotic, as *er* or *erm* (Tottie 2011: 173). Clark & Fox Tree (2002: 92) also point out that each language has its own set of fillers or hesitations — usually one or two syllables built around the central vowel in the language — and provide a list of fillers in several languages other than English such as German, Dutch, Spanish, French and Hebrew, just to name a few.

Since the current investigation focuses on texts originally produced in English and subsequently translated and adapted into Italian, the discussion presented here will mainly focus on hesitation phenomena within the English and Italian context. Moreover, it must be pointed out that the aim of this study is not to differentiate the meanings and functions among the alternative transcriptions of hesitations (distinguishing, for instance, between the

⁷ For a detailed description of the seven phenomena, see Blankenship & Kay (1964: 360-361).

⁸ From the examples reported in Gilquin (2008), it can be assumed that under the label “truncated words” are included those words left unfinished, pronounced omitting the final part, but also what is commonly also referred to as false starts. For a definition of false starts, see Biber *et al.* (1999: 1062).

nasalized and the non-nasalized variant). In this investigation, the different transcriptions used are considered variants of the same phenomenon and are employed interchangeably. In case there is any significant difference among the transcriptions, it will be made explicit. Moreover, when citing the investigation carried out by previous scholars, the transcription reported in the discussion will align with the one used by the authors in their study to present their results.

Another aspect that needs to be taken into account concerns the terminological debate. Based on the features included under the label “hesitation markers”, these elements have been referred to as “(un)filled pauses” (Maclay & Osgood 1959; Goldman-Eisler 1968), “editing expressions” (Levelt 1989), “fillers” (Clark & Fox Tree 2002), “filler words” (Lynch 2002), “discourse markers” (Fox Tree 2007), and also “autonomous fillers” (Candea *et al.* 2005). Based on how frequently speakers resort to hesitation markers, Christenfeld (1995: 171) distinguishes between “ummers” and “um-avoiders”. Erard (2007) simply refers to these phenomena as *ums*, but categorizes them as “verbal blunders”. Corley & Stewart (2008) opt for “hesitation disfluencies”, implying that these elements constitute a blockage to fluent speech; while Voghera (2017) uses the term “phonetic disfluencies” (as different from “textual disfluencies”) to refer to phenomena like pauses filled by vocalizations, nasalizations, and lengthenings that do alter or interrupt the utterance sound chain but do not change its verbal sequence, i.e. a moment of pause needed by the speaker as a control mechanism for speech programming (*ibidem*: 73). Culpeper & Kytö (2010: 379), instead, use the term “pragmatic noise” to refer to those items, not fully lexical, having a cognitive expressive function.

Scholars have gradually rejected the negative connotation associated with hesitations as a mere form of disfluency in favour of a more positive view. Schegloff (2010), for instance, simply refers to them as *uh(m)s*; similarly, Kosmala & Morgenstern (2018) resort to *uh* and *um*, as well as to the term “fillers”. This change of attitude has led scholars to look for new terms that could highlight the role of hesitations as communicative solutions used to fulfil various pragmatic-conversational functions. Tottie (2011: 193), for instance, calls attention to the fact that the label “filler” also constitutes “a rather negative and uninformative default term that is dependent on the ideal of fluency and that says nothing about the discourse functions of these items”. Thus, Tottie (2011) suggests using the term “planners” or “planning signals” (also shared by Jucker 2015a, 2015b, 2021).

In line with the abovementioned debates and discussions, this study aims to use the terms “hesitation(s)”, “filled/unfilled pause(s)” or “fillers” in a manner that is entirely neutral

and unbiased, free of any possible negative connotation. In the present discussion, these labels are used interchangeably to discuss and review the features under analysis. Nevertheless, when citing the investigation carried out by other scholars, the label used to discuss their results may conform to the label employed in the original work.

To sum up this section, the different terminologies used by scholars throughout the years are a direct representation of the way in which hesitation phenomena are reputed. Scholars have slowly departed from the perception of hesitations as mere disfluencies, moving toward a more positive attitude regarding their role not only in speech production but also in listeners' comprehension. The next paragraphs will provide a more detailed overview of this change of attitude toward these phenomena, discussing the pragmatic functions and meanings associated to them, i.e. how they are used in conversation and for what purposes.

1.4. Hesitations in Speech and their Basic Meaning(s)

1.4.1. Functional Categories of Hesitations: Cognitive, Repair and Interpersonal Functions

Traditionally, hesitations have been classified as disfluencies, meaning a linguistic device used to interrupt the fluency of speech without adding any propositional content to the utterance (Maclay & Osgood 1959; Fox Tree 1995), expressing some sort of difficulty, uncertainty (Corley & Stewart 2008) and poor communication skills (Lynch 2002). In other words, hesitations were mainly considered performance errors or unsystematic noise going against the 'ideal delivery', that is, the correct way to perform an utterance (Clark & Clark 1977: 261). Nevertheless, in spontaneous conversation, minor disturbances in the flow of speech are completely normal and do not hinder the comprehension of the discourse. It has been estimated that about 6% of words uttered in spontaneous speech include disfluencies such as repeats, restarts and hesitations, like the fillers *um* and *er* (non-including silent pauses) (Fox Tree 1995: 709-710). Biber *et al.* (1999) also report that the combined frequency of *uh* and *um* in a mixed corpus of both British and American spoken English is over 13,000 per million words (*ibidem*: 1054). These elements are not necessarily averse to comprehension; on the contrary, if these minor interferences are not disruptive enough, listeners usually miss them or simply ignore them (Lickley 1995; Bard & Lickley 1998). Lickley (2015) suggests that a distinction should be made between "surface fluency" and "perceived fluency", that is, what the listener perceives as opposed to what the speaker

actually utters⁹. Moreover, based on how frequent the occurrence of hesitation phenomena is in speech, fluency should be considered as the exception rather than the rule (*ibidem*: 451).

Chafe's (1980) study was one of the first works supporting the idea that hesitation phenomena do not hinder communication but, on the contrary, they are an inherent part of the communicative process taking place in real time. According to the author, "[t]he speaker is interested in the adequate verbalization of his thoughts. Pauses, false starts, afterthoughts, and repetitions do not hinder the goal, but are steps on the way to achieving it" (*ibidem*: 170). Hieke (1981: 150) also sustains this conception by stating:

Hesitations [...] form an integral part of speech production in the positive sense, a view quite in opposition to the attitude that there is fluency on one hand and hesitancy on the other. Not only are hesitations a normal component of fluency if they occur in moderation, but now pauses and the other hesitations can actually be considered well-formedness phenomena rather than disfluencies, at least as far as they serve as devices by the speaker to produce more error-free, high-quality speech.

In this respect, hesitations constitute a reflection of the cognitive effort required by the online planning necessary in communicative interactions, rather than a blockage contaminating the fluidity of the spoken engagement. More specifically, in Chafe's (1980) view, hesitations reflect a way for the speaker to show where in the sentence it is easier to move forward and where it is more challenging. From the data analysed, which comes from twenty female speakers of American English talking about a short movie they had been previously shown, Chafe notices that hesitations are more present between sentences than within sentences, meaning that "the breaks at coherence which are found at sentence boundaries [...] require more time-consuming mental processing than does the moving from focus to focus within a cluster" (*ibidem*: 175). In other words, moments of transition in speech require extra time to process. In particular, among some of the factors that seem to foster the use of hesitations in speech, Chafe includes finding and clarifying the initial sentence, moving to the next focus within a sentence, and moving across a sentence boundary (*ibidem*: 180); at the same time, the content of the speech also plays a role. More specifically, people tend to hesitate when deciding what to talk about next, as well as when deciding how to talk about the chosen topic (*ibidem*: 171).

⁹ Lickley (2015) also investigates what it means to be 'fluent' in different contexts of speech production, i.e. in various speech pathologies and in language learning. The author underlines how the word "fluency" referred to speech production can be used in many different ways. It is therefore difficult to achieve a universally shared definition of "fluency" and, consequently, "disfluency".

Over the past few decades, much research has been carried out on the underlying reasons for speakers to hesitate in speech, examining the fundamental meanings of hesitations with reference to their linguistic and paralinguistic nature, their functions and their status within the language system. In this regard, scholars have slowly outclassed the basic view of hesitations as linguistic elements signalling a delay (Clark & Fox Tree 2002) in favour of a broader spectrum of pragmatic functions.

1.4.2. The Basic Meaning of Hesitations: From Delays to Pragmatic Markers

A major study that influenced much of the research conducted on hesitations over the last two decades is the investigation carried out by Clark & Fox Tree (2002)¹⁰. The authors refused the simple categorization of filled pauses as something opposed to silent pauses and attributed to hesitations the status of English words, i.e. interjections whose basic meaning is to signal the speaker's intention to introduce a delay. The delay may be either brief (*uh*) or long (*um*) and it is quantified based on the duration of the adjacent silence co-occurring with *uh* and *um*. As a result, *uh* indicates a minor delay, while *um* indicates a major delay (*ibidem*: 79). According to the authors, *uh* and *um* can have other functions too; however, those functions are mere implicatures derived from the basic meaning of delay (*ibidem*). The meanings speakers can implicate through these delays may indicate, for instance, that they want to hold or cede the floor (*ibidem*: 89), but also planning problems co-occurring with repairs in which the delay may reflect the time to identify the problem and/or formulate the repair, as well as a variety of other meanings such as deciding what to say next, doubt, uncertainty or searching memory for a word (*ibidem*: 90). The hypothesis proposed by the authors is that, as it happens for any other word, speakers consciously plan, formulate and articulate fillers such as *uh* and *um* (*ibidem*: 80). As a result, since they are not automatic, speakers can have selective control over them and intentionally reduce their occurrences according to the circumstances, for instance, in formal contexts (*ibidem*: 99). Clark & Fox Tree's (2002) data also seem to be in line with the traditional considerations made in relation to the location of hesitations (see §1.2.1; Chafe 1980). Since planning speech is more difficult at unit boundaries, i.e. locations where "speakers have to select the message, formulate the syntax [...], and create its overall prosody" (Clark & Fox Tree 2002: 94), this is where speakers are expected to signal a delay. This is confirmed by corpus data, according

¹⁰ Their data comes from the London-Lund corpus. It consists of face-to-face conversations among British adults recorded between 1961 and 1976 (Svartvik J., Quirk R., (eds.). 1980. *A Corpus of English Conversation. Lund Studies in English*, 56. Lund: Liber/Gleerups).

to which the majority of fillers occur at unit boundaries (where the delay is presumably longer) as opposed to other positions such as at the beginning or in the middle of the unit (where the delay is presumably shorter) (*ibidem*).

The claim that *uh* and *um* are interjections whose basic meaning is the signalling of minor and major delay has later been questioned and rejected by multiple scholars (O'Connell & Kowal 2005; Schegloff 2010; Tottie 2015). O'Connell & Kowal (2005) test this hypothesis by observing the occurrences of *uh* and *um* in six media interviews by Hillary Clinton and measuring the silent pauses occurring either before or after *uh* and *um*¹¹. According to the authors, a media-quality corpus of interview conversations between expert speakers represents the ideal data to test Clark & Fox Tree's hypothesis, since it can be assumed that if *uh* and *um* did signal some delay, professional speakers would expertly use these rhetorical devices (*ibidem*: 560). Results show that the number of silent pauses does not significantly correlate with fillers; as a result, since in most cases fillers do not co-occur with silence, *uh* and *um* do not seem to signal any sort of delay. Moreover, when pauses did occur after *uh* and *um*, no significant difference was found in their duration. Similarly, Tottie (2015) found that in about one-half of the instances observed in a corpus of American English, hesitations "did not serve to introduce delays but ended them" (*ibidem*: 49), while other scholars, like Schegloff (2010), consider Clark & Fox Tree's (2002) view limited in its scope and opt for a different interpretation. More specifically, much of the data analysed by Schegloff (2010) does not show *uh* and *um* being followed by delays, and in cases where they are followed by silences, *uh* and *um* are considered to embody a delay rather than merely announcing it (*ibidem*: 171).

On the other hand, Clark & Fox Tree's (2002) hypothesis found some new support in Betz and Kosmala's recent study (2019). The scholars not only tested the hypothesis that there is a correlation between silence and hesitation, on the one hand, and between hesitation type and silence duration, on the other; but they also took into account the notion of distance to test whether fillers can influence the duration of silences. Their results, based on a small corpus of data¹², showed that the length of silence occurring in the proximity of *um* and *uh* systematically changes according to the type of filler, with longer silences following the nasalized filler *um*. This also seems to confirm the assumption that the longer the filler, the

¹¹ This is in line with Clark & Fox Tree's (2002) idea that the distinction between a minor (*uh*) and a major (*um*) delay can be defined based on the pauses adjacent to *uh* and *um*.

¹² Their corpus includes a collection of recorded interview-style face-to-face interactions between 16 native English speakers (aged 18-23). The participants knew each other quite well (Betz & Kosmala 2019: 1-2).

longer the silence that comes after. Finally, distance also seems to influence the duration of silences. In fact, silences are longer when directly adjacent to fillers; however, silences before fillers tend to be shorter than silences following them (*ibidem*: 3)¹³.

Regardless of the new support, it must be pointed out that, in testing the ‘delay hypothesis’, scholars have moved beyond the simple idea of hesitations co-occurring with speech production troubles. Schegloff (2010), for instance, used a conversation analytic approach to investigate telephone conversations and demonstrated how hesitations can be used to orient participants in spoken interactions. The author recognizes that hesitations can be used to announce something to some extent problematic, which consequently tends to be delayed — as is the case when introducing bad-news announcements (*ibidem*: 168), dispreferred responses or sequences¹⁴ — but they are also used to indicate “that a sequence’s ending has resisted consummation and is being tried again” (*ibidem*: 166) or to mark “the reason for the call” (*ibidem*: 146). In other words, hesitations represent tools to organize the verbal exchange, for example, to indicate the transition from the initial “identification sequence” or the typical small talk and formalities of telephone conversations to the actual reason that prompted the call (*ibidem*: 166).

In summary, within the functional debate, scholars have gradually rejected the idea of hesitations as mere disfluencies used to announce delay, uncertainty or reluctance, and have instead focused on hesitations as pragmatic markers in conversational interactions. In spoken dialogic exchanges involving multiple speakers, these features can serve a variety of functions. To begin, hesitation can function as an important tool for turn organization (Stenström 1990; Norrick 2014; Tottie 2014a).

¹³As pointed out by Betz & Kosmala (2019), the dissimilar results obtained from their study in comparison with O’Connell & Kowall’s (2005) investigation may be influenced by the different speaking styles required by the different communication settings from which the data has been extracted. In fact, Clark & Fox Tree (2002) used spontaneous face-to-face conversations, O’Connell & Kowall (2005) observed trained professional speakers, while Betz & Kosmala (2019) focused their analysis on interview-style data of spontaneous speech. The spontaneity aspect may be the reason for Betz & Kosmala’s (2019) closeness to Clark & Fox Tree’s (2002) results.

¹⁴ In adjacent pairs we can recognize preferred and dispreferred sequences. Preferred sequences are those that advance the course of action launched, while dispreferred sequences do not advance that course of action; on the contrary, they block it or interfere with its realization (Schegloff 2010: 141).

1.4.3. Hesitations in Discourse Interactions as Turn-Management Tools

Different scholars have studied hesitations as turn-management tools in conversation¹⁵. Fischer (2000: 115-116) defines them as discourse particles having different functions depending on their occurrence in the turn: “a turn-initial occurrence of *uhm* would thus be a turn-taking device, a turn-final occurrence would constitute a turn-yielding device, and a turn-medial occurrence would be a turn-holding device”. Norrick (2014) claims that elements like *uh* and *um* can constitute free-standing turns in conversation; he classifies them as “phatic interjections”, i.e. a sub-category of primary interjections that can constitute an utterance by themselves (*ibidem*: 256).

Stenström (1990) had already discussed the use of pauses and fillers in discourse as interactional and social devices fulfilling the function of either turn-taking, turn-holding and turn-yielding (*ibidem*: 247-248)¹⁶. According to the author, hesitations constitute a device used to avoid too long silences that may hamper speakers participation in the communicative exchange. In fact, “[i]f a person remains silent too long at a potential speaker shift he will never get the turn. If he wants to take the turn he has to say something, even if he has not yet made up his mind exactly what to say” (*ibidem*: 248). The same happens at the moment of turn-holding. If a speaker wants to keep the conversational turn, it is preferable to avoid staying silent for too long; otherwise, other participants may interpret the silence as a form of turn-yielding.

Tottie (2014a) proposed a more detailed analysis of the elements under investigation by also considering their location within the turn. The investigation of the *Santa Barbara Corpus of Spoken American English*¹⁷ showed that *uh* and *um* more frequently occur in turn-medial position (see also Kjellmer 2003), followed by turn-initial and turn-final positions (Tottie 2014a: 393). More specifically, each location in which *uhm*¹⁸ occurs in the turn can be related to different communicative functions: turn-initial *uhm* typically has a planning nature, signalling that the speaker is thinking about what to say. It can also indicate the speaker’s intention to add something to the conversational turn or that the speaker wants to

¹⁵ In conversation analysis, “[a]ctions accomplished by talking get done in turns-at-talk” (Scheff 2007: 3). The organization of taking turns to talk in communicative interactions is fundamental to conversation (Sacks *et al.* 1974).

¹⁶ Corpus study carried out on the *London-Lund Corpus of Spoken English*, compiled by Jan Svartvik.

¹⁷ Du Bois J.W., Chafe W.L., Meyer C., Thompson S.A., Englebreton R., Martey N. 2000-2005. *Santa Barbara Corpus of Spoken American English*, Parts 1-4. Philadelphia: Linguistic Data Consortium. Available online at <https://www.linguistics.ucsb.edu/research/santa-barbara-corpus>.

¹⁸ In line with Tottie’s (2014a) study, and for the sake of brevity, *uhm* is used here as a label including both the nasalized (*um*) and the non-nasalized (*uh*) variants of the same variable (*ibidem*: 381).

take over the turn and contribute to the verbal exchange. This last case, however, occurs more rarely in the corpus. In general, the most frequent use of *uhm* in this turn-initial position is to answer questions (*ibidem*: 394-396). Turn-medial *uhm*, instead, typically serves the function of turn-holders, meaning that the turn has not yet come to an end and that the speaker still has something to add. Tottie rejects the idea that speakers utter hesitations in this location to prevent someone else from taking the turn, as suggested by Stenström (1990), and puts forward the hypothesis of vocalizations having a planning function, i.e. speakers take some time to think about how to phrase the utterance in the best possible way (Tottie 2014a: 397). Turn-final *uhm* consists of vocalizations that do not add any content to the conversational turn, and it is usually used by speakers to yield the turn to their interlocutors. However, it is not always easy to distinguish between turn-holding and turn-yielding in conversation, since “[a] speaker’s turn-holding can easily be interpreted by the interlocutor as turn-yielding” (*ibidem*: 401). In the case of turn-yielding, Stenström (1990) considers filled pauses “a speaker’s last effort to keep the turn when he does not know how to continue” (*ibidem*: 251), while Tottie (2014a) considers them unintentional symptoms of planning (*ibidem*: 402).

More specifically, Tottie underlines how all three functions of turn-taking, turn-holding and turn-yielding can be related to a more general planning function, respectively: thinking about what to say, thinking about how to phrase what to say, and unintentional symptoms of the thinking process. Interestingly, Boonsuk *et al.* (2019) have more recently observed the position of hesitation markers in everyday informal conversation so as to identify the more frequent patterns in the English language. They noticed that speakers mainly resort to hesitations in the middle of their utterances rather than in initial position, the latter being where researchers have presumed hesitations to occur more frequently as a result of the higher cognitive load required by the planning process at an initial stage. These findings rather show speakers’ tendency to use hesitations in middle position to hold their conversational turn and to help organize what to say next and how to say it, offering support to Tottie’s (2011) use of “planning signals” or “planners” to refer to these features.

The present scrutiny has briefly touched upon the fact that the functions hesitations fulfil in spoken interactions tend to intersect and overlap. In this regard, Kjellmer’s (2003) multifunctional scheme appears to be particularly relevant to this investigation.

1.5. The Multifunctional Framework of Hesitations

The functional framework of filled pauses adopted by Kjellmer (2003) in his investigation of the *Cobuild Corpus* (Collins Birmingham University International Language Database)¹⁹ appears particularly valuable for the present study. Kjellmer (2003) states that “like other speech components, [filled pauses] should be looked upon in most cases as task-performing elements, employed to bring about certain effects” (*ibidem*: 181) and continues: “the speaker may well make use of them unconsciously in aiming for a certain effect that s/he achieves without understanding how it was brought about” (*ibidem*). In other words, speakers may use these elements to fulfil certain functions both in a conscious and in an unconscious manner and, since they operate to some extent under the level of consciousness, they do not hinder communication or negatively alter our perception of speakers’ fluency and eloquence (*ibidem*: 191).

Kjellmer (2003) identified five main functions of filled pauses (transcribed as *er* and *erm* in the corpus), i.e. hesitation proper, signposting speaker turns, attracting attention, highlighting, and correction. These functions are not always easy to distinguish and some overlap is highly possible. Firstly, the “hesitation proper” function indicates that filled pauses can be used to express uncertainty. This is supported by the fact that hesitations tend to co-occur with each other, as well as with other phenomena like false starts. As pointed out by the author, “[s]uch repetitions indicate uncertainty and are in themselves hesitation markers” (*ibidem*: 182). This display of filled pauses may be due to various reasons: filled pauses may be used to provide the speaker some thinking time to decide what to say next and how to organize the rest of the sentence (*ibidem*), but they may also indicate that the speaker has nothing significant to add to the conversation or that s/he is looking for the right word (*ibidem*: 183). The “hesitation proper” function corresponds to what Tottie (2014a) simply defines as “planning” and, in Kjellmer’s view, it is strictly related to the “correction function”, in the sense that filled pauses mark corrections in repairs and reformulations, signalling to the listener that the speaker wants to change and redraft a word, a part of a word or a whole sentence, and that what comes before the correction should be ignored (Kjellmer 2003: 188-189).

The “signposting speaker turns” function includes the turn-management tools of turn-taking, turn-holding, and turn-yielding described in the previous section. Kjellmer (2003:

¹⁹ Sinclair J. 1987. *Looking up: an account of the COBUILD Project in lexical computing* (Collins Cobuild dictionaries). London/Glasgow: Collins ELT.

184) confirms Stenström's (1990) observation that filled pauses often serve as turn-initiators, that is, to indicate that the speaker wants to take over the turn. Moreover, in the corpus, *er* and *erm* tended to occur next to answer particles (like *yeah*, *yes*, *right*) (*ibidem*), corroborating Tottie's (2014a) observation that hesitations at the beginning of a turn often collocate with sequences of answering questions. As pointed out by Tottie (*ibidem*), Kjellmer (2003) also draws attention to the fact that the three turn-management functions often co-occur with each other, making it difficult to distinguish between turn-holding and turn-yielding (*ibidem*: 185-186). Closely linked to the turn-taking function of filled pauses is the "attracting attention" function: speakers may, in fact, utter hesitations to attract attention to themselves and take the floor, as well as to establish contact with other interlocutors (*ibidem*: 186). Finally, in line with the function of "attracting attention", speakers may also use filled pauses to direct the listeners' attention toward an important topic, a specific word or combination of words. This is the case, for instance, of hesitations preceding semantically heavy items or words that are not frequently used in the language. The speaker can insert hesitations before these elements in order to signal that they may require more attention and more mental effort to process. This function corresponds to what Kjellmer calls "highlighting" (*ibidem*: 187), and it can also be used to alert to implications and innuendos that the listener may otherwise not grasp. In this sense, filled pauses may, for instance, precede irony, euphemisms or disapproving understatement (*ibidem*: 188).

This last aspect has been more recently emphasized by Staley & Jucker (2021), who define it as a "face-saving" or "mitigating" function. Drawing attention to the multifunctional nature of hesitations, the authors claim that in addition to the two main functional groups in which hesitation phenomena can be classified — functions related to trouble, repair and planning, on the one hand, and functions related to discourse structuring, turn-management, new or semantically heavy content, on the other — a third category can be established, i.e. hesitations as a tool to hedge and mitigate suggestions or recommendations that could potentially cause face threats (*ibidem*: 29) (see also Tonetti Tübben & Landert 2022). Moreover, in line with the multifunctional nature of elements like *uh* and *um*, this face-saving function might not be completely free of a planning and a hesitating component (Staley & Jucker 2021: 32).

In view of this multifunctionality and the possible overlaps, how does the speaker manage to interpret hesitations and their functions in the correct way? To answer this question, prosody and intonation patterns play a key role in inferring the proper interpretation (Kjellmer 2003; Lickley 2015). As stated by Poyatos (2002b), "interpersonal

communication would be most limited if we did not possess the many qualifiers or voice types which truly attest to its unique complexity". He argues that "different qualifiers succeed each other in our spoken discourse, especially in certain situations, in such a way the same verbal message can acquire, through a concrete type of voice, a very specific and efficacious value" (*ibidem*: 23).

1.5.1. The Prosodic and Paralinguistic Nature of Hesitations

According to Poyatos (2002b), a series of nonverbal activities play a role in human interactions; such activities include prosodic, paralinguistic and kinetic elements. Among these features, hesitations are defined as "paralinguistic alternants", i.e. non-verbal segmental utterances that can be modified by kinesics and paralinguistic suprasegmental features and that "can be encoded also out of awareness just like gestures and other nonverbal behaviours" (*ibidem*:142-143). In line with the functional description of hesitations outlined above, such features fulfil a wide range of functions: hesitation pauses, for instance, may be due to "a break in one's train of thought, not daring say anything, or out of insecurity, or caused by the listener or listeners' disturbing feedback" (Poyatos 2002a: 257); while unfilled pauses (also referred to as "silent alternants") can express, to name a few: hesitation, word search, word withholding, but also memory searching, thinking pause and remembering (*ibidem*: 165-166).

It is central to remark that hesitations are not completely independent of the prosodic context in which they occur. Shriberg & Lickley's (1993) study offers support to this statement, claiming that there is a relationship between the intonation of filled pauses and their prosodic context. By examining both American and British English, their findings show that the intonation and pitch of filled pauses in the middle of clauses are related to their surrounding sounds. In this sense, filled pauses appear to lend continuity to speech rather than disrupting it, supporting the view that filled pauses do help spoken delivery, guiding and smoothening communicative interactions (Kjellmer 2003: 191). In Brennan & Williams' (1995) view, both filled and unfilled pauses can be considered paralinguistic features processed by the participants of communicative interactions together with other linguistic information and can affect, along with other paralinguistic features, the pragmatic interpretation of the utterance (*ibidem*: 383). More specifically, they examined whether listeners are sensitive to the information speakers display in their speech through paralinguistic features, conducting three experiments in a question-answering context. When

answering questions, speakers commonly convey metacognitive information about their knowledge state, that is, information on whether they know how to answer the question or not. Such metacognitive information may be conveyed through the insertion of filled pauses that delay the response, but also through intonation. Incorrect answers, for instance, typically present a rising intonation (*ibidem*: 388). This is also true for uncertain answers; that is, in those cases where speakers are not certain about the correctness of their answer, they tend to resort to a rising intonation (see Smith & Clark 1993: 33; Brennan & Williams 1995: 388). Brennan & Williams' (1995) investigation showed that listeners are indeed attentive to filled pauses and prosodic features such as intonation to interpret the information speakers exhibit in their speech regarding their knowledge state (*ibidem*: 396). Listeners, in fact, would give a higher rating to speakers' knowledge when the answer would include a falling intonation (i.e. when the speaker sounded more confident), while they would give a lower rating when the answer would follow a filled pause (*ibidem*: 394)²⁰.

To conclude, even if intentionality is in doubt (i.e. whether speakers use hesitation features consciously or not), no question arises regarding the fact that these elements do not just randomly happen; they are uttered by speakers with a purpose, that is, to convey something to the listener who may be more or less observant of these features.

1.6. Hesitations in Conversation: Sociodemographic Factors

Thus far, the considerations made on hesitations mainly focused on studies conducted on cognitive and interpersonal factors influencing the use of hesitations in speech; social factors, however, also play a role. In order to obtain a clear framework of hesitation phenomena in ordinary spoken language, it is worth remarking some of the sociodemographic factors that might differentiate the use of hesitations in spontaneous conversation and influence speakers' choice of fillers. In particular, studies have tried to investigate the correlation between the occurrence of filled pauses and socio-demographic factors such as age, gender and socio-economic class (Rayson *et al.* 1997; Tottie 2011; Stenström 2012), also taking into account personality traits (Laserna *et al.* 2014) and situational factors, like context and register (Tottie 2014b), topic domains, or the relationship between speakers (Bortfeld *et al.* 2001).

²⁰ In Brennan & Williams's (1995) view, the fact that listeners are attentive to the presence of filled or unfilled pauses supports the idea that paralinguistic features can be used to coordinate individual mental states during communicative interactions (*ibidem*: 397).

As for the relation of filled pauses with age, gender and socio-economic factors, scholars have noted that men tend to use fillers more than women (Tottie 2011: 182-183; Stenström 2012: 551), with the filler *er* appearing among one of the most characteristic words of male speech (Rayson *et al.* 1997: 4)²¹. This is also confirmed by Staley & Jucker's (2021) recent study, although the small-size nature of their corpus does not allow for quantitative-based conclusions. Laserna *et al.*'s (2014) analysis of "filler words" in daily conversations, instead, associated the use of filled pauses with age rather than with gender (*ibidem*: 335). This data is in line with Rayson *et al.*'s (1997) findings, according to which the filler *er* falls within the words more frequently used by speakers over 35 (*ibidem*: 8). These results are also supported by Tottie (2011), who reports that the highest frequencies of fillers usually occur among older speakers, possibly due to the growing difficulty in carrying out certain cognitive tasks with increasing age (Bortfeld *et al.* 2001: 143)²². Finally, speakers' social background also brings about some differences in the use of pauses and hesitation phenomena (Stenström 2012: 553). As reported by Tottie (2011: 190), speakers of the highest socio-economic class present a higher frequency of fillers. This data aligns with Rayson *et al.*'s (1997) findings; namely, in the *British National Corpus*, the filler *erm* seems to occur among the words more frequently used by upper social classes (*ibidem*: 9).

Interestingly enough, Tottie's (2011) investigation also provides some data regarding the speakers' selection of the nasalized/non-nasalized filler in speech. Tottie studied the occurrence of the fillers *uh* and *um* (as variants of a single variable) as sociolinguistic markers in British English, observing two sub-corpora of the spoken component of the *British National Corpus*²³: the context-governed part (BNC-CG) and the demographic part (BNC-DEM). The former consists of recordings from different formal domains, while the latter includes spontaneous speech in more informal situations (*ibidem*: 178). The analysis

²¹ Rayson *et al.*'s (1997) survey was conducted on the spoken English component of the *British National Corpus*.

²² Bortfeld *et al.* (2001) studied a corpus of task-oriented conversations between two people and observed how certain factors can affect the rate of fillers (*uh*, *ah*, *um*, *er*), repeats and restarts in conversation. The corpus contained data obtained from conversations uttered by 48 pairs of people. These couples included 24 pairs of male and female strangers and 24 pairs of married couples. The task involved three between-subjects variables (age, gender and relationship between the speakers) and two within-subjects variables (topic domain and task role). Participants had to match a set of twelve picture cards containing pictures of either a familiar or an unfamiliar domain, alternating in the role of matcher and director. The objective was for the matcher to line up the picture cards in the same order as the cards of the director (*ibidem*: 129-131). In line with the claim that fillers typically occur when the cognitive load required for planning is higher, hesitations are expected to be more frequent in the directors' speech (i.e. the ones taking most of the initiative while carrying out the task) and when discussing unfamiliar topics. While directors did indeed use fillers more than matchers, elements like *uh* and *er* presented a higher rate when discussing familiar topics rather than unfamiliar ones (*ibidem*: 142).

²³ *British National Corpus (BNC), XML Edition*. 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium.

shows that even though *uh* seems to be the overall favourite filler in both sub-corpora, there are some significant differences concerning the use of the nasalized variant *um* in relation to the three sociolinguistic factors of gender, age, and socio-economic class. Firstly, while men seem to use more fillers than women in spontaneous conversation, women use a higher proportion of the nasalized variant *um* (*ibidem*: 182-183). The same happens with age: data shows that older speakers have the highest frequencies of fillers; however, a higher proportion of the nasalized variant can be noticed among younger speakers (*ibidem*: 186-187) (as also shown by Stenström 2012: 550). Finally, as for the third sociolinguistic factor, speakers of the managerial classes tend to prefer *um* compared to the manual-worker classes (Tottie 2011: 190).

In addition to the socio-demographic factors already observed, there is another aspect worth mentioning that could influence the use or the type of filler selected by speakers, i.e. language variation. As already pointed out by Clark & Fox Tree (2002), hesitations (or fillers) may change not only from language to language but also within the same language. In other words, hesitations may vary across dialects and language varieties. As it will be shown in §1.9, such variation constitutes an extremely critical factor in the Italian language as well. As for the English language, Tottie examined the occurrence of hesitations in both British (2011) and American English (2014b). In particular, Tottie (2014b) investigated the occurrences of *uh* and *um* in the first quarter of the *Santa Barbara Corpus of Spoken American English* (SBC) and compared the results to the data obtained from a previous study (Tottie 2011) carried out on the *British National Corpus* (BNC). Although the results are to be considered only preliminary due to the different sizes of the corpora, it appears that *uh* and *um* are more frequent among British Speakers (in BNC) compared to American speakers (in SBC) (Tottie 2014b: 9-10).

Moreover, the investigation on SBC shows no significant differences linked to gender in American English (Tottie 2014b: 10), while no statistical significance could be established regarding age and socio-economic class (*ibidem*: 11). However, similarly to the *British National Corpus* (Tottie 2011), the *Santa Barbara Corpus* data shows that the use of *uh* and *um* is dependent on extra-linguistic context and register. As a matter of fact, in the corpus of American English, it appears that in more private, intimate and familiar situations, speakers are less likely to insert the fillers *uh* and *um*. On the other hand, the frequency of *uh* and *um* is higher in non-private contexts (Tottie 2014b: 13). This, of course, does not mean that one context calls for more fluent or disfluent speech (*ibidem*: 25), just like the distinction between private and non-private contexts does not automatically translate into a

degree of formality or informality. From the above, what could be hypothesized is that, in non-private contexts, speakers feel the need to plan their speech more carefully than they do in private ones.

To conclude this section, the studies examined here have illustrated that it is possible to pinpoint some trends regarding the occurrence of hesitations in spontaneous interactions; most notably, there are socio-demographic and situational factors that can influence speakers' use and selection of hesitations. These factors can also impact on both speech delivery and listeners' perception. The responsiveness and sensitivity of people to hesitation phenomena and the effects hesitations may have on the way both speakers and their speech are perceived by listeners will be further investigated in §1.7.

1.7. The Perception of Hesitations and their Impact on Listeners and Speakers

As pointed out by Brennan & Williams (1995), “research on paralinguistic features of speech has examined what speakers produce but not what listeners perceive” (*ibidem*: 384). The effects hesitations may have on listeners' impressions and comprehension had already been noted by Blankenship & Kay (1964), who claimed that hesitations can be used to favour the listener's comprehension, making the message more redundant. Hesitations can, in fact, affect how listeners process spoken language (Corley *et al.* 2007) and also shape their impression and perception of the speaker (Christenfeld 1995).

1.7.1. Hesitations and the Impact on the Perception of the Speaker

Several studies have tried to survey through questionnaires people's perception of hesitation markers (Fox Tree 2007), whether they influence the impression listeners have of those who resort to hesitations, and whether people are aware of the actual way in which these fillers affect our perception of the other (Christenfeld 1995; Brennan & Williams 1995).

Christenfeld (1995) conducted two experiments with the aim of examining whether the presence of hesitations in speech is actually noticed by listeners and whether *ums* influence the audience's perception of the speaker. The first experiment involved a questionnaire in which subjects had to rate ummers, i.e. those speakers who utter *um*, *er*, and *uh* a lot in their speech (*ibidem*: 173). Results showed that such speakers were rated not only uncomfortable, monotonous and ill-prepared but also nervous, disfluent, unsophisticated, and lacking in confidence (*ibidem*). This is in line with the results obtained from the study

of Brennan & Williams (1995), in which speakers had to rate their ability “to accurately assess and monitor their own knowledge” when answering questions (*ibidem*: 384)²⁴. The rating was lower when the answer was articulated after a filled pause (*ibidem*: 394). The same estimation has been noted by Fox Tree (2007), who investigated the folk notion of *uh* and *um* as discourse markers next to *you know* and *like*. Fox Tree noted that people commonly describe their use of *uh* and *um* as the result of speech production trouble and nervousness (*ibidem*: 306).

Nevertheless, people’s reports and impressions declared through questionnaires do not always correspond to the way hesitation phenomena are actually perceived and processed in spoken interactions. Christenfeld’s (1995) second experiment was developed for this exact reason, i.e. to evaluate the actual responsiveness of listeners to *ums*, while Fox Tree (2007) compared people’s evaluation of hesitations with their actual uses in speech.

Christenfeld (1995) established a set of different conditions to examine the audience’s actual level of awareness of hesitations and how they are perceived. Subjects were exposed to three different versions of the same radio talk show recording: the original one containing 25 pauses filled with vocalizations; an edited version in which filled pauses had been replaced with silent pauses; and finally, another edited version with no pauses at all. In addition, three more conditions were established: a group of subjects received no instructions to complete the test, another group was asked to focus on content, while the last group was asked to focus on style. The results showed no sensitivity in the group with the focus on content, some sensitivity in the group with no instructions, and the greatest sensitivity among the group with the focus on style (*ibidem*: 180). Moreover, as for the perceived quality of the speech, filled pauses created a better impression among listeners than pauses of silence; in fact, the audio containing filled pauses was rated as more relaxed than the one containing silent pauses. The latter gave the listeners a sense of anxiety (*ibidem*: 182). More specifically, Christenfeld’s (1995) findings suggest that hesitations can be processed in an automatic way, but not always are since there are certain conditions that can raise listeners’ awareness of these phenomena (in the abovementioned experiment, focusing on style translated into a higher sensitivity to fillers). Moreover, even though hesitations can impact listeners’ perception of the speaker, the elimination of these features does not automatically alter or improve the impression of the audience. Thus, the audience does not

²⁴ Hart (1965) defines this ability as the “feeling of knowing” (FOK).

seem to judge the quality of the speech (how good or bad it is) solely based on the rate of *ums* (*ibidem*: 181).

After having established a clearer picture of listeners' level of awareness of hesitations and how hesitations are perceived by speakers, Fox Tree (2007) focused on the actual way these elements are used in people's speech. Fox Tree (2007) compared the folk notion of *uh* and *um* with the data obtained from a previous study (Fox Tree 2006) in which subjects were asked to first tell and then retell a personal story. When telling their stories, participants first tried to limit their use of hesitation markers; however, when retelling the story, *uh* and *um* were among the most reused items. More specifically, hesitations were mainly displayed at the beginning of the narrative segment, before a list of elements, and in remembering detailed information like names and times. All these locations correspond to points where speakers might be more likely to encounter production trouble and, consequently, need more time to plan their speech (Fox Tree 2007: 309). Results show that people have a sense of the prototypical meaning of *uh* and *um*, or the function they fulfil. Such awareness reflects to some extent how people use these elements in speech (*ibidem*: 310). However, in their actual use, speakers seem to use hesitations not so much to express nervousness (as stated by speakers in the first part of the experiment), but rather to signal that they need time to think and process their ideas. In other words, they seem to resort to hesitation as planning signals.

1.7.2. Hesitations and the Impact on Language Comprehension

Since speech interactions usually involve both coding and decoding activities, hesitations should not be investigated solely in relation to the speaker's perspective but also with respect to those receiving the message, i.e. how filled pauses can impact language comprehension.

Initially, researchers have focused on short-term effects of hesitations in language comprehension tasks (Fox Tree 2001). Through a word monitoring task during which listeners had to detect words from recordings where filled pauses were either present or had been edited out, Fox Tree (2001) showed that subjects were faster at carrying out the task when words were preceded by hesitations²⁵. As a result, hesitations can be beneficial to

²⁵ The task involved speech stimulus uttered by one speaker and containing at least one filled pause (either *uh* or *um*). The filled pauses present in the speech were immediately followed by a new word that had not been previously uttered in the speech segment (target word). For each of the critical stimuli and for half the filler stimuli, other versions of the recordings were created. In such versions the critical filled pauses had been digitally removed. The aim of the test was to verify whether listeners were able to detect the editing in the stimuli or not (Fox Tree 2001: 322).

listeners' ability to recognize words in upcoming speech (*ibidem*: 323), increasing their level of attention.

Similarly, Corley *et al.* (2007) tried to observe long-term effects of hesitations in natural circumstances. The authors carried out an experiment using Event-Related Potentials (ERPs) to measure subjects' speech processing time, followed by a surprise recognition memory test to evaluate the long-term impact on the listener. In particular, "ERPs provide an ideal means to investigate how listeners process disfluent speech in a situation which is a close analogue to everyday language comprehension" (*ibidem*: 660). To carry out the experiment, an N400 effect was established. The N400 effect consists of an ERP component that provides information regarding language processing. More specifically, in the investigation, the N400 effect was related to the predictability of words within their linguistic context, meaning any positive or negative change in voltage recorded at the scalp could indicate whether a word was more predictable or unpredictable in a certain context. According to the authors, since hesitations tend to precede unpredictable words in speech (see Goldman-Eisler 1958), changes in N400 could demonstrate whether hesitations impact the processing of the input depending on its predictability (Corley *et al.* 2007: 660). Firstly, the experiment showed that the N400 was reduced when target words were preceded by the hesitation marker *er*. This reduction reflected a greater ease of integrating words into their contexts. Then, a recognition memory test was conducted. Results showed that words following the hesitation marker *er* were more likely to be recognized and remembered, meaning that the presence of hesitations does indeed influence the way these words had been processed by subjects (*ibidem*: 666). In short, hesitations affect not only the short-term processing of language, but they also have long-term effects (*ibidem*: 667)²⁶.

Interestingly, Corley *et al.* (2007) have also tried answering the question of what it is about hesitations that justifies this processing change. The authors note that "in the disfluent utterances, the *er* necessarily introduces more time between the context and the (predictable or unpredictable) target word" (*ibidem*: 666). This consideration corroborates with the findings that hesitations provide more thinking time not only to speakers but also to listeners: "the cognitive function of pausing to give the speaker time to think may be mirrored in giving the hearer more time to comprehend" (Tannen 1985: 99). Although Tannen is referring to silence, this statement can also be cited with regard to pauses filled

²⁶ Following Corley *et al.*'s (2007) investigation, Collard *et al.* (2008) also used ERPs to demonstrate that hesitations do affect listeners' attention signals, also impacting the way in which either a concomitant or a subsequent utterance is processed.

with nasalization, i.e. giving listeners time to fully absorb the message that the speaker wants to communicate (Tottie 2014b: 26). In Tottie's view, this could be seen as a sign of fluency. In fact, although *uh* and *um* "do not add propositional content to the flow of speech, they contribute to fluency in a way that (silent) pauses do not, by adding to the flow of language rather than interrupting it" (*ibidem*).

1.8. Hesitations in Writing and in Fiction

The review of studies previously examined has discussed hesitations in spontaneous or semi-spontaneous speech, mainly focusing on impromptu conversations, interviews, (re)tellings, etc. Nevertheless, since the present investigation aims to study hesitations in film dialogue, which represents a hybrid between speech and writing (Chaume 2012: 81), it is necessary to provide an overview of studies conducted on hesitations in different modes, that is, the presence of hesitations in written language as well as in different types of fictional texts.

1.8.1. Hesitations in Writing

Although the phenomenon studied here is mainly related to oral language, researchers have also observed and analysed the presence of hesitations in writing (Jucker 2015a; Rühlemann & Hilpert 2017; Tottie 2017, 2020). This is particularly interesting for the present investigation, which focuses on hesitations in film speech. Film speech can be considered a hybrid between speech and writing since dialogues are first developed in the written form of scripts and, subsequently, performed orally by actors. In writing, the elements under investigation are intentionally displayed with a stylistic purpose in mind; most notably, writers can use these elements to express hesitation and planning (Jucker 2015a: 173), but also as stance-markers, often indicating irony (Tottie 2014b: 20).

Much of the research on the written mode has been carried out through a diachronic perspective. Jucker (2015a), for instance, reviewed two centuries of written texts contained in the *Corpus of Historical American English* (COHA)²⁷. The corpus includes texts from the 1810s to the 2000s drawn from four genres: fiction, popular magazines, newspapers, and non-fiction books. The analysis shows that, despite being used quite moderately in COHA, a growing frequency of *uh* and *um* as planners can be attested between the 19th and the 21st century (*ibidem*: 172).

²⁷ Davies M. 2010. *The Corpus of Historical American English: 400 million words, 1810-2009*. Available online at <http://corpus.byu.edu/coha/>.

In this respect, much of the diachronic research on this topic has aimed at observing the progressive colloquialization of language in writing, i.e. written norms becoming closer to the norms of spoken language (Rühlemann & Hilpert 2017: 105). In particular, researchers tried to attest the increasing frequency of certain features typical of spoken language in written texts, also providing interesting insights about the pragmatic functions they fulfil in such contexts. Rühlemann & Hilpert (2017), for instance, investigated the presence of three hesitators (*uh*, *um* and *er*) among a list of elements typical of spoken language — defined “inserts” by Biber *et al.* (1999: 56) — in the *TIME Magazine Corpus* (Davies 2007)²⁸. The analysis highlighted an increase in the use of inserts in writing from 1923 to 2006. With regard to hesitators, the scholars claimed that “all three forms are modestly strongly correlated with the decades. As regards the rate of change on the earlier decade, only *um* sees its highest relative growth in the 1990s; *uh* and *er* change most dramatically in earlier decades” (Rühlemann & Hilpert 2017: 112). Similarly, a growing frequency of occurrence of hesitation in writing has been attested by Tottie (2017) in her investigation of *uh*, *um* and *er* in American English journalistic prose (COCA, Davies 2008)²⁹. Through an analysis of written language from the 1960s until the early 2000s, the study shows how, after a modest start in the early 1990s, an increase of fillers occurred in journalistic prose between 1995 and 2004. By 2010/2011, the frequency was 3.76 times as high as it was in the first period (Tottie 2017: 115)³⁰.

The results obtained from these studies provide a first glimpse at any attested development in the conventions of certain types of written texts (see also Jucker 2015a: 171). The diachronic investigations conducted on written language showcase the progressive insertion of orality features in the written mode; in addition, it is worth commenting on the functional way in which hesitation features are displayed in writing. In this regard, Tottie

²⁸ Davies M. 2007. *TIME Magazine Corpus: 100 million words, 1920s–2000s*. Available online at <http://corpus.byu.edu/time/>.

²⁹ Davies M. 2008. *The Corpus of Contemporary American English: 450 million words, 1990–present*. Available online at <http://corpus.byu.edu/coca/>.

³⁰ Two sub-corpora of COCA were observed for the analysis: The Magazine sub-corpus and The Newspaper sub-corpus. Tottie (2017) shows that there are some differences between the frequencies recorded for the two sub-corpora. In fact, considering the total sample, 60% of tokens of the elements investigated occurred in the Magazine sub-corpus and 40% in the Newspaper sub-corpus. With the passing of time, these proportions changed “from almost a fifty-fifty distribution in the first period to about a quarter from News and three-quarters in Magazines in the last period” (*ibidem*: 115). However, it is important to note that the selection of magazines included in the corpus changed much over time, just like the magazines domains or the types of publications. On the contrary, the sources constituting the sample of newspapers remained constant. Thereby, the addition of lightweight publications and the domain of Women/Men/Fashion and magazines focusing on looks and lifestyle has brought about significant consequences for the study of *uh* and *um*. In other words, there is a difference in register and colloquialization that should be taken into consideration (*ibidem*:116–117).

(2017) observes the correlation between the functional use of hesitations in writing and their location in the sentence.

Findings show that *uh*, *um* and *er* mainly occur in either medial or initial positions. More specifically, the majority of occurrences happen in sentence-medial position (*ibidem*: 117), where they are used to express the writers' attitude toward the content of their upcoming text (*ibidem*: 120). Hesitations occurring in sentence-initial position, instead, are mainly used to express the writer's attitude — generally a negative or contradictory attitude (*ibidem*: 122) — toward a previous message usually said or written by someone else (*ibidem*: 120). In particular, medial hesitations are used with a “tongue-in-cheek” function, implying that what is being said should not be taken seriously, as well as to signal irony or euphemism (*ibidem*). Moreover, writers also tend to rely on these elements to express things tactfully (*ibidem*: 121). As for the occurrence of hesitations in initial position, those are usually used in negative answers to rhetorical questions or in questions making fun of previous assumptions (*ibidem*: 122).

To sum up, in writing, features like *uh*, *um* and *er* are used to achieve a stylistic effect and to express the writer's attitude as markers of authorial stance and increasing subjectivity. In initial position, they fulfil the function of attitude adverbs, while in medial position they operate as style adverbs. As pointed out by Tottie (2017: 123), their function is metalinguistic: they are used to highlight preceding or following elements and to produce a humorous or ironic effect; moreover, they can be used for politeness purposes since “in the context of writing, hesitate does not mean ‘pause because of uncertainty’. It has a strong connotation of reluctance or unwillingness, reluctance to be tactless, to hurt or insult someone” (*ibidem*: 124). There are some similarities between some of the uses found in spontaneous speech and some of the uses found in writing. Turn-initial hesitations, for instance, are often used in constructions of answers to questions, as is also the case in spoken language (Kjellmer 2003; Tottie 2014a); however, while in speech people hesitate because they are unsure about how to answer, in writing subjects “merely pretend to hesitate, out of reluctance to say something tactless or hurtful” (Tottie 2017: 124-125). In these texts, hesitations are purposely and deliberately displayed in order to convey specific meanings to those who are reading. As it will be described in §1.8.2, the same happens in fiction.

Before discussing hesitations in fictional texts, there is one last aspect that needs to be further examined since studies conducted on the written mode provide the opportunity to reflect more carefully upon the long debate regarding the status of hesitations as ‘words’ in the conventional sense.

1.8.1.1. *Categorizing Hesitations: Are they Words?*

Traditionally, psycholinguists have had different views on the status of elements like *uh* and *um*. According to Erard (2007), most speakers do not consider these elements as words with meanings and codified uses, but as undesirable noise unintentionally produced by speakers. Clark & Fox Tree (2002) recognized these elements as conventional English words, namely, interjections whose basic meaning is to signal either a major or a minor delay. According to them, just like any other word in the language system, speakers purposefully and intentionally plan and include these features in their speech. Corley & Stewart (2008), like O'Connell & Kowal (2004, 2005), are opposed to this view. The latter are unwilling to categorize *uh* and *um* as interjections (O'Connell & Kowal 2004), suggesting that more research needs to be done with regard to their structure and functions before categorizing them among this word class (O'Connell & Kowal 2005: 568). Corley & Stewart (2008), instead, investigated whether speakers intentionally produced hesitations in their speech or whether they emerged as an involuntary consequence of difficulties encountered during the planning process. Through a review of studies on the topic, the authors concluded that, even though listeners are attentive to pauses filled with vocalizations, there is no conclusive evidence that such elements are intentionally uttered by speakers. Therefore, they reject their classification as words in the conventional sense. According to the authors, hesitations provide information to the listener similarly to facial gestures or changes in the tone of voice. This information corresponds to: "pay attention, the speaker's in trouble and the next part of the message may not be what you predicted" (*ibidem*: 600).

More recently, Tottie (2017) investigated the word status of *uh* and *um*, making a distinction between the written and the spoken mode. According to the author, in writing, the elements under investigation do fulfil the criteria to be categorized as 'words' in the conventional sense since they are used as intentionally produced items, separated from other linguistic items either by blank spaces or punctuation marks. Moreover, they are used with specific meanings or, at least, communicative functions (*ibidem*: 109). As previously mentioned in §1.8.1, in writing, Tottie classifies them as stance adverbs, which, depending on their location in the sentence, can be distinguished between attitude adverbs (in sentence-initial position) and style adverbs (in sentence-medial position). In speech, on the other hand, Tottie describes the status of these elements as "a continuum with low-to-high degrees of wordhood" (*ibidem*: 127). In other words, Tottie considers them as nascent words on the

cline of lexicalization, where one end of the cline is represented by their cliticized use (as in *anduh* or *butuh*, for instance), and the other end of the cline is represented by their occurrence between silent pauses; hence, fulfilling the phonological criterion of wordhood (*ibidem*: 125), as illustrated by Biber *et al.* (1999: 51).

In addition, Tottie (2017, 2020) considers other factors in support of the lexicalization and pragmaticalization of *uh* and *um* in spoken language, as in the case of hesitations being used for word-search and in answers to questions. In such instances, *uh* and *um* are perceived as words in the mental lexicon chosen by speakers for specific purposes (Tottie 2017: 125-126). Moreover, the author investigates the transition of these elements from fillers (or filled pauses), used as involuntary planners in speech, to fully-fledged words in writing (Tottie 2020)³¹; this is achieved by focusing on the word-search use of *uh* and *um* in speech through a sub-corpus of the *Santa Barbara Corpus of Spoken American English* (SBC). The author hypothesizes that the word-search use of these items in writing actually corresponds to a “make-believe word-search function”, which aims to emulate in the written mode the way these elements are used in spoken language (*ibidem*: 31). In other words, the word-search use of *uh* and *um* in spoken language might have provoked their lexicalization in written language as “a rare case of coinage or root creation, i.e. lexicalization occurring without any kind of prior traditional word-formation process taking place” (*ibidem*: 29). In particular, the results obtained from the investigation of the *Santa Barbara Corpus* show that, despite being rather uncommon, the word-search use of *uh* and *um* in speech mainly involves content words such as nouns and adjectives, which are linguistic items extremely relevant for the interaction. In this respect, the word-search use of *uh* and *um* is so remarkable in speakers’ minds that it results in their adoption in the written mode (*ibidem*: 37). The conclusion drawn by Tottie is that the word-search function of *uh* and *um* in speech can be considered a stage in the word-formation process these elements (nascent words) are undergoing (*ibidem*: 38).

In line with O’Connell & Kowall (2005) and Tottie (2015), this study rejects the idea that *uh* and *um* should be classified among interjections or discourse markers (Clark & Fox Tree 2002; Fox Tree 2007) for two main reasons: on the one hand, the classification of hesitations in one of these two classes would inevitably imply the detailed inclusion of other

³¹ Within the word-search process, a distinction is made between “lexical retrieval” and “lexical selection”. Lexical retrieval is used to indicate that there is only one target word, which is considered to be the ‘correct’ one. Lexical selection is used to indicate that there can be multiple possible choices, among which one is considered to be better than the others (Tottie 2020: 32).

linguistic elements categorized as interjections or discourse markers that, for one thing, have been largely investigated; on the other hand, the inclusion of filled pauses into one of these categories would feed into a research approach that participated in the neglect of hesitations as a subject of enquiry for a very long time. The view proposed in the present study is that the phenomenon needs to be investigated as a category in its own right. In this regard, Tottie (2011) points out that “one reason for being neglected [...] is certainly their low status as vocalizations and their lack of recognition as words” (*ibidem*: 193). The present survey does not aim to resolve the debate regarding the status of these features as ‘words’ in the conventional sense, since “the category word is itself a fuzzy one, and uhm is a good example of a borderline case” (Tottie 2015: 51)³². Considering the pre-planned written origin of film dialogues, Tottie’s (2017) distinction between the written and the spoken mode could constitute an interesting view for the categorization of hesitations in the current investigation. However, even though originally created in the written form of scripts, viewers ultimately consume film dialogues through the spoken mode (either in the original or dubbed version). The present investigation supports the view that, considering the multifunctional role hesitations play in discourse, they should be studied as the pragmatic markers that they are; hence, as paralinguistic cues (Brennan & Williams 1995) — also defined “paralinguistic alternants” by Poyatos (2002a, 2002b) and “paralinguistic signals” by Culpeper & Kytö (2010) — capable of conveying a variety of pragmatic meanings.

1.8.2. Hesitations in Fiction

Although fictional texts are clearly different and less spontaneous than spoken interactions, they have slowly come to be accredited as a rightful source for pragmatic investigations. As a matter of fact, since fictional texts portray characters engaged in communicative behaviours, they can be subjected to pragmatic inquiry (Jucker 2015b: 63).

In studying the language of scripted dialogues created for different types of fictional texts (either for the screen, for the stage or for books), Bublitz (2017) made a paramount distinction between orality and spoken language. Spoken language refers to “the vocally produced segmental and non-segmental (prosodic) sounds” produced by speakers, while orality concerns “speaking as a mode” (*ibidem*: 241). In other words, some of the features vocally produced while speaking “are a characteristic, though not a constitutive feature of

³² Tottie (2015) uses *uhm* as one variable, including four variables: the non-nasal variant *uh*, the nasal variant *um*, and the lengthened variants *uh=* and *u=m* (*ibidem*: 41).

spoken language” (*ibidem*). These features also include hesitations and the investigation of their incorporation in fictional texts can constitute a rich source of pragmatic data, as long as fictional language is analysed “on its own terms and not as a less than perfect substitute for spontaneous spoken communication” (Jucker 2015b: 63). Indeed, fictional language is only a representation of impromptu spoken language. Authors of fictional texts insert these elements for specific purposes and with the aim to recall the general stereotypes associated with them rather than faithfully reproduce the way they are actually used in spontaneous conversation (*ibidem*: 65).

The extensive use of these pragmatic elements in fiction is also supported by Jucker’s (2015a) study, where he investigates the diachronic development of *uh* and *um* in the *Corpus of Historical American English* (COHA). The corpus comprises two centuries of written texts (from the 1810s to the 2000s) from fiction, popular magazines, newspapers and non-fiction books. Among the four genres included in the corpus, the fiction category witnesses the highest frequency of hesitations: this category includes between 48% and 55% of data for each decade (*ibidem*: 170), where they mainly fulfil hesitation and planning functions (*ibidem*: 173), but they can also be used by authors to indicate that the speaker is lying (*ibidem*: 176). To put it in other words, *uh* and *um* are often used in fiction as stylistic devices employed by authors to characterize characters and their communicative behaviour (Jucker 2015b: 67).

Nevertheless, fiction does not only include narrative texts, but it also involves performed fiction. More recent studies (Jucker 2021; Tonetti Tübben & Landert 2022) have focused on the simultaneous examination of different types of corpora confronting performed fiction, spontaneous conversation and writing in order to underline differences with regard to the occurrence of orality features; this is achieved by comparing fictional and non-fictional texts, as well as scripted and non-scripted texts.

Jucker (2021) studied a small range of orality features such as contractions and inserts — the latter also include hesitations — across seven corpora: four corpora of performed fiction from movies and TV series³³, the *Santa Barbara Corpus of Spoken American English* (SBC) reporting spontaneous face-to-face interactions, and two sections of the *Corpus of*

³³ The four corpora of performed fiction are: the *Corpus of American Soap Operas*, the *TV Corpus*, the *Movie Corpus*, and the *Sydney Corpus of Television Dialogue*. Differences between the frequencies of the four categories of inserts studied in performed fiction (response forms, discourse markers, hesitators and interjections) may result from the differences existing between the four corpora investigated. In fact, different corpora may adopt different transcription principles (Jucker 2021: 356). Moreover, fictional genre also plays a role since different genres of movies and television series call for different interactional features, linked to the specificities of the product (*ibidem*: 357).

Contemporary American English (COCA) including written fiction, on the one hand, and academic writing, on the other. Results confirm the expected trend: orality features are most frequent in the corpus of spontaneous face-to-face conversations (SBC), while they are almost completely non-existent in academic writing and rather rare, but present, in written fiction (COCA). The data obtained from the four corpora of performed fiction lie in the middle of these two extremes (*ibidem*: 355). Fictional language shows frequencies closer to spontaneous interactions than to written language; however, this does not mean that fictional language is a perfect and absolute substitute for spontaneous spoken language (*ibidem*: 356). Orality features function as indexical cues in fiction, meaning that they do not have to perfectly replicate the frequencies of spontaneous face-to-face interactions to convey a sense of orality (*ibidem*). In fact, the nature of fictional language itself makes it impossible for it to fully replicate spontaneous face-to-face conversations since this type of language is generally scripted in advance, meaning that actors do not have to actually plan or elaborate what to say next. They utter lines previously memorised and they already know in advance what the interlocutor is going to answer (*ibidem*: 342). The only exception to fictional language is when improvisation comes into play, as is the case in improvisation theatre, in which actors have to come up with their lines as the story progresses (*ibidem*). This last context of improvised fictional language has been taken into consideration by Tonetti Tübben & Landert (2022) in their investigation of planners.

Tonetti Tübben & Landert (2022) compared the use of *uh* and *um* across three different genres: spontaneous conversation (*Santa Barbara Corpus*), scripted fictional dialogue (TV dialogue) and improvised theatre (i.e. spontaneously produced fiction). In so doing, they compared two dimensions of text production: the dimension of spontaneous vs scripted and the dimension of fiction vs non-fiction. In the production contexts here observed, *uh* and *um* are categorized as pragmatic markers fulfilling a range of interpersonal and textual functions, i.e. a cognitive, a repair, a discourse management and an interpersonal function. The analysis aims to survey how the occurrence of *uh* and *um* changes across texts depending on the spontaneity factor, in addition to how the functions they fulfil in the text are influenced by the (non-)fictional context of occurrence (*ibidem*: 351). Results show that planners are more frequent in spontaneous dialogues (both fictional and non-fictional) than they are in scripted dialogues. Thus, spontaneity seems to influence the general frequency of occurrence of planners. At the same time, the spontaneity factor also impacts on the distribution of their functions: in spontaneous dialogues, planners are mostly used with a cognitive, repair and discourse management function; in fictional dialogues, they mainly

fulfil interpersonal functions (*ibidem*: 372). In the first case, the cognitive, repair and discourse management functions can be related to the real-time language production occurring in impromptu conversations; in the second case, the significant presence of planners fulfilling interpersonal functions may be due to the “higher presence of interpersonal conflict and tension in fiction” (*ibidem*: 371). However, this does not mean that the other functions do not occur in fictional texts; on the contrary, they can occur both intentionally, for characterization purposes and uses that contribute to making dialogues sound natural, and unintentionally due to the effort required to recall lines from memory (*ibidem*: 370-371). The analysis, in fact, shows that the planners chosen for the study feature in all three genres of spontaneous conversation, scripted fictional dialogue and improvised theatre. This can indicate that there are some similarities between the use of *uh* and *um* in spontaneous conversation and in scripted fiction. What differs between the two contexts (spontaneous vs scripted) is the content, on the one hand, and the conditions in which these texts are produced, on the other; at the same time, the content and the production conditions impact on the frequency of planners’ functions (*ibidem*).

The studies analysed here confirm that some of the functions and peculiarities of hesitations established in spontaneous spoken language are also found and reproduced in fictional texts. In particular, it is interesting to notice that the differences between the two contexts of spontaneous conversation and fictional dialogue are not always as intense as they could be thought. Moreover, as advocated by Jucker (2015b: 64), fictional texts can constitute a fertile resource for the investigation of pragmatic markers:

[...] we are dealing with a representation of communication constructed by the author according to artistic considerations. Such considerations vary from one period to the next; they vary according to the literary genre and from one author to the next. But these differences do not disqualify fictional language from serving as data in pragmatics. On the contrary they call for a detailed analysis of fictional language in its own right.

In this investigation, pragmatic markers are observed in film speech through the investigation of dialogues in original English movies and their dubbed Italian versions. Considering the languages involved in the study, it is of paramount importance to review the research conducted on hesitations in the Italian context. This will be discussed in §1.9.

1.9. Hesitations in the Italian Language

This dissertation revolves around hesitations in film speech through the investigation of original English movies and the approach used when adapting them in Italian through dubbing. Much of the research cited thus far focused on the English language in order to draw a picture of the nature of hesitations, their meanings and functions in spontaneous spoken language with the purpose of then comparing it with the orality reproduced in English movie dialogues. Nevertheless, considering the translational approach adopted in this study, it is essential to examine how hesitation phenomena are specifically produced in the Italian language.

Some insightful experimental studies on hesitation phenomena in Italian date from the early 2000s. Tissi (2000), for instance, conducted an experiment on ten Italian students studying to become interpreters and asked them to simultaneously interpret two extracts of a tape-recorded spontaneous speech from German into Italian. These texts contained pauses (grammatical and/or communicative pauses vs non-grammatical pauses), hesitations (vocalized hesitations and vowel/consonant prolongations) and interruptions (repeats, restructuring and false starts). The aim was to investigate whether the presence of these phenomena would affect the interpreters' comprehension and delivery of the translation. Interestingly, various similarities can be established between Simultaneous Interpretation (SI) and dubbing. As claimed by Stenzl (1983: 40):

While in translation the message is conveyed entirely by graphic means, interpretation involves not only linguistic elements and what they convey, but also intonation, voice quality, changes in pitch and loudness, pauses and non-linguistic elements such as facial expressions, gestures etc, which can all contribute to the message and may have to be verbalized by the interpreter.

The prosodic and non-linguistic elements involved in SI are also relevant in dubbing, since the original voice is substituted with a new one in a different language.

As for the effect of hesitations on the delivery of the translation, Tissi's results show some great individual differences in the productions uttered by the Italian interpreters (*ibidem*: 115). Nevertheless, some trends could be established. Notably, the Italian texts contained fewer but longer pauses of silence than the source text and the majority of them were grammatical pauses, while prolongations were more present in the interpreted texts than in the original ones (*ibidem*: 120-121). The fact that Italian speakers use prolongations more frequently than vocalizations corroborates with the results more recently obtained by

Di Napoli (2020: 25) and Schettino & Eklund (2023: 82). Moreover, this is in line with Savino & Refice's (2000) claim that in Italian, "speakers do not need to insert a "spurious" vowel-like segment in their speech: they can achieve the same by simply lengthening a word ending vowel" (*ibidem*: 422).

As for filled pauses, Tissi's (2000) findings show that these are the elements exhibiting greater variability among the subjects, ranging from 0 to 34 in the first text and from 1 to 53 in the second (*ibidem*: 118). Despite the great individual variability and the consequent difficulty in identifying trends, the occurrences found in the Italian version can be used to show how these features can serve as strategic tools speakers resort to in order to achieve communicative goals. Filled pauses, for instance, were used by students when encountering difficulties in carrying out the interpreting task, but they also resorted to these features while waiting for new items to be translated (*ibidem*: 121). Moreover, Tissi stressed that in Italian, if supported by intonations, syllable lengthenings can have a function of "discourse scansion" (*ibidem*: 118)³⁴, that is, they have a communicative value in conveying meanings to the listener. In particular, after a lengthening at the end of a word — that is, a lengthening expressing hesitation — one of the following words may present the prolongation of the tonic vowel. These types of tonic vowel prolongation serve to draw listeners' attention back to an important informative element, usually a lexical item (*ibidem*).

The fact that vocalizations and hesitation phenomena are used as specific communication tools and vary extensively according to the speaker's individual preferences is also supported by Giannini (2003b). Giannini investigated hesitation phenomena in Italian in order to verify whether there is some sort of relationship between the tonal pattern and the Italian stressed/unstressed vocalic system (*ibidem*: 2653). Her analysis found that speakers from different geographical areas resort to different realizations of hesitation phenomena³⁵, depending on their personal phonological inventory (*ibidem*: 2655)³⁶.

³⁴ The word "scansion" is related to prosody in the literary field and it is defined as "[t]he analysis of poetic metre in verse lines, by displaying stresses, pauses, and rhyme patterns with conventional visual symbols" (Baldick 2015). Tissi (2000) seems to apply this literary concept to spoken discourse, alluding to the communicative value of features like intonation, pauses and rhythm in spoken language.

³⁵ The regional variation within the Italian context is reflected by the different varieties investigated in research revolving around spoken language in Italian. Di Napoli (2020), for instance, focused on Roman Italian task-oriented dialogue, while Schettino & Eklund (2023) studied data concerning the Neapolitan variety of the Italian language.

³⁶ More specifically, speakers from southern Italy produce vocalizations inside the unstressed vocalic area, resulting in a sound similar to a central vowel. The central vowel uttered by southern speakers is not part of the Italian phonological inventory, but it is part of the diatopic variation spoken in Naples and Bari (Giannini 2003b: 2654-2655).

As for the tonal pattern, results show that while vocalizations show a descending pattern, lengthenings are not subjected to any significant variation. As also noted in a previous investigation (Giannini 2003a), such tonal difference may reflect the different roles these phenomena play in on-line programming (Giannini 2003b: 2655). This is also in line with Schettino & Eklund's (2023) findings on the duration of prolongations compared to the duration of filled pauses. As also noted by Giannini (2003a: 165) and Di Napoli (2020: 26), lengthenings are generally shorter than filled pauses. According to Schettino & Eklund (2023), the different durations measured for prolongations and filled pauses offer some support to the idea that these two hesitation phenomena may be used for different goals and may be linked to different speech management processes (*ibidem*: 84). Both Giannini (2003a) and Di Napoli (2020) show the interactional function of filled pauses and prolongations in relation to their positions in conversational turns. Filled pauses predominantly occur in turn-initial position, while prolongations occur in turn-medial position (Giannini 2003a: 166; Di Napoli 2020: 26), as well as in turn-final position (Di Napoli 2020: 26). In line with the results obtained for the English language, also in Italian turn-initial vocalizations signal a turn-taking function, as well as the need to buy some time to plan the utterance (Giannini 2003a: 167). Moreover, Di Napoli (2020) notices that turn-initial filled pauses occasionally occur after a question, providing additional time to answer, while turn-final prolongations are typically followed by holds (holding the floor) rather than speaker changes (*ibidem*: 27). The same turn-holding function is noted by Giannini (2003a) regarding prolongations in turn-medial positions (*ibidem*: 167).

One last aspect that needs to be mentioned concerns the fact that, next to durational traits, Schettino & Eklund (2023) have also investigated the distributional picture of lengthenings and their segmental traits in Italian. The analysis was carried out by focusing on both dialogical and monological data concerning the Neapolitan variety of the Italian language. The distributional data seem to demonstrate that, in Italian, prolongations occur in more or less the same measure both in content words and in functional words (*ibidem*: 84; see also Di Napoli 2020: 26). As for the segmental features, lengthenings typically involve the final vowel of a word. When prolongations involve consonants, these are usually followed by a schwa sound, reproducing the canonical Italian syllabic structure CV (consonant-vowel) (Schettino & Eklund 2023: 83).

1.9.1. The Phonetic Pattern of Hesitations in Italian

Studies conducted on hesitations in the Italian language have also focused on the phonetic patterns of hesitation phenomena related to their function in Italian. Schettino *et al.* (2021a), for instance, investigated three hesitation markers (lengthenings, filled pauses and lexicalized filled pauses) with the objective to ascertain whether specific phonetic features differentiate between various communicative functions in Italian discourse.

Three possible functions were established for the phenomena under study: a hesitative, a word-searching and a structuring function³⁷; in addition, three parameters were considered for the phonetic analysis: pitch contour shape, pitch range, and distance between the pitch means of the hesitation and its corresponding intonational unit. Notably, this research was conducted on the C.H.R.O.M.E corpus³⁸, which is a database containing nearly one and a half hours of semi-spontaneous speech produced by three female expert tour guides. Results show that flat and peak contours are the most frequent outputs: flat contour mainly involved prolongations and filled pauses, while peak contour typically occurred with lexical fillers. As for function, peak contour appeared to be related to a structuring or a focusing function, while a flat profile seems to characterize hesitative fillers. In other words, lengthenings and filled pauses are mainly used in relation to a word-searching function and, more generally, to address planning problems; lexicalized fillers and silent pauses are mainly associated with structuring and focusing functions (Schettino *et al.* 2021b: 256). As regards distance between hesitations and their surroundings, hesitations fulfilling a word-search function appear more phonetically distant, thus more evident, than hesitations with a focusing function. These results seem to support the claim that phonetic patterns can discriminate between different functions hesitations may fulfil in Italian (Schettino *et al.* 2021a). As for duration, filled pauses are typically longer than the other hesitation phenomena observed, while all hesitation types studied are longer when fulfilling a word-searching function. The opposite happens to prolongations when they fulfil a focusing function; in this case, they tend to be significantly shorter (Schettino *et al.* 2021b: 254).

³⁷ The word-searching function involves lexical retrieval and the consequent selection of a target word; the structuring function involves items occurring at clause (syntax) or topic-comment (information structure) boundaries; the focusing function serves to pinpoint key concepts in a certain context or to indicate that semantically heavy elements are about to follow; finally, the hesitative function regards the basic function of speech planning (Schettino *et al.* 2021b: 249-250).

³⁸ The corpus was developed for the CHROME project, i.e. *Cultural Heritage Resources Orienting Multimodal Experience*, which aims at shaping multimodal data for the development of Virtual Agents to be used in museums (Origlia *et al.* 2018).

What is particularly interesting about the C.H.R.O.M.E corpus (Origlia *et al.* 2018) is the fact that the speech included in the corpus is a pre-planned speech uttered by tour guides and characterized by a descriptive-informative style that aims to provide listeners with information about touristic tours in a clear and efficient manner (*ibidem*: 255). Therefore, some similarities can be drawn between the language involved in the C.H.R.O.M.E corpus and film speech. As a matter of fact, film dialogues are also pre-planned texts subsequently performed orally by actors with the aim to provide viewers with the required information regarding characters and the narrative events. As noted by Schettino *et al.* (2021b), the occurrence of hesitations in the pre-scripted texts included in the C.H.R.O.M.E corpus validates the hypothesis that these elements help communication (*ibidem*: 255). More specifically, confronting the utterances produced by the three guides, the authors attested how speakers can use different hesitation strategies to tackle their sociolinguistic and psycholinguistic needs; namely, to handle their speech and, at the same time, perform their communicative goals (*ibidem*).

Finally, though hesitations typically occur as standalone elements, the C.H.R.O.M.E corpus also allowed to pinpoint some of their combination patterns and distribution within tonal units in Italian discourse (Schettino *et al.* 2021c). As for hesitation combinations, hesitations tend to occur in clusters when they fulfil a word-searching function, often occurring at boundaries between tonal units (*ibidem*: 31-32). Clusters more frequently involve two elements in which lengthenings are typically followed by a filled pause; filled pauses are mostly followed by a lexical filler or a silent pause; silent pauses are usually followed by a filled pause. As for the position of hesitations, filled and unfilled pauses usually occur at the beginning of the prosodic unit and they typically fulfil a structuring function at the clause or information structural level (*ibidem*).

To conclude, research on hesitation phenomena in Italian has concentrated on various types of speech: pre-planned speech, spontaneous speech uttered in a specific geographical area and the simultaneous interpretation of spontaneous speech into Italian. Despite the different linguistic contexts analysed and the individual preferences of speakers, also due to their personal phonological inventory, some patterns for the Italian language could be established:

- In Italian, prolongations occur more frequently than other hesitation phenomena (Tissi 2000; Di Napoli 2020; Schettino & Eklund 2023) in both function and content words. They usually involve word-final vowels and, when involving

consonants, a schwa sound is inserted, reproducing the canonical Italian syllabic structure (consonant-vowel) (Schettino & Eklund 2023).

- The prosodic and durational differences existing between prolongations and vocalizations may reflect different roles in planning and speech management, as well as different communicative functions (Giannini 2003a, b; Schettino *et al.* 2021a; Schettino & Eklund 2023)
- Vocalizations show more individual differences among speakers than prolongations, which display more consistent values among speakers (Giannini 2003a). This may be due to the fact that lengthenings are part of either lexical or functional words whose time balance within the text has to be somehow preserved (*ibidem*: 166).

1.9.2. The Effect of Hesitations on Naturalness in the Italian Language

More recently, a new line of research was established in order to survey how hesitation phenomena are perceived by speakers in Italian and their role in guaranteeing the naturalness of prefabricated speech. These factors were investigated using neural speech synthesisers, i.e. systems that can be trained to replicate the voice of a target speaker as well as a specific speaking style (Schettino *et al.* 2022: 195). More specifically, the speech synthesiser analysed here was trained on the C.H.R.O.M.E corpus presented in §1.9.1.

The experiment was developed in the following way: participants were asked to listen to a pair of synthetic utterances and select which one of the two sounded more natural and which one of the two sounded more appropriate for the context of use, namely to give voice to a virtual avatar serving in a museum. In this way, both naturalness and appropriateness were tested in relation to the communicative context or the situational application of the computational model (*ibidem*: 198). Results show that the presence of hesitation phenomena such as vowel lengthening, filled pauses and silent pauses, does affect listeners' perception of the synthesised utterances. In fact, those utterances containing hesitations were judged as more natural than those that did not contain hesitations. On the other hand, non-hesitant utterances were judged more appropriate to give voice to a virtual agent. In other words, utterances containing hesitation phenomena are perceived as more natural sounding, but less likely to be associated with a non-human voice. This may be due to the fact that people do not expect machines to produce elements that sound natural and spontaneous; they rather expect a virtual avatar to sound artificial (*ibidem*: 200-201). Nevertheless, despite being

perceived as less appropriate for the intended context of application, hesitation phenomena make utterances sound more natural in Italian.

At a later stage, these results were further investigated through a memory test (Schettino *et al.* 2024). This follow-up study confirms the previous findings according to which hesitation phenomena increase the naturalness effect of synthesised utterances; moreover, it also shows that phenomena like lengthenings, filled pauses and silent pauses can benefit listeners' comprehension (as also proved by Fox Tree 2001; Corley *et al.* 2007). The memory test was developed in the following way: subjects were asked to listen to a presentation of approximately three minutes uttered by the avatar introducing visitors to a museum and talking about the tour. Two versions of this presentation were created: with and without hesitation phenomena. Participants were asked to rate the quality of what they heard, its naturalness and its pleasantness; then, they were asked to complete a memory task regarding 10 pieces of information about the tourist site mentioned in the recordings (Schettino *et al.* 2024: 9). Results of the task-oriented listening test show that participants listening to the recording containing hesitations were more able to recall information than those who listened to the recording without hesitations. Therefore, it seems that hesitations are beneficial for listeners' comprehension and recollection of information (*ibidem*: 12). Moreover, both versions of the recording were judged in a similar way, i.e. they were perceived to be of good quality. This means that hesitations do not inevitably have a negative impact on the way a text is liked or perceived (*ibidem*).

As already mentioned, the studies cited here have investigated hesitation phenomena in modelling human communicative behaviours to develop speech synthesis systems that would support human-machine interaction systems. Although the present investigation does not revolve around such systems, the results obtained from the abovementioned tests provide interesting information regarding how hesitation phenomena are perceived by speakers and their role in guaranteeing the naturalness of prefabricated speech (Schettino *et al.* 2022), yielding useful insights about another form of prefabricated speech, namely film speech.

1.10. Further Remarks on the Analysis of Hesitations in Film Speech

The objective of this chapter is to provide a global overview of the trends and patterns regarding the occurrence of hesitations in ordinary spoken language so as to then compare it to the occurrences in film speech. The investigation of hesitations in spoken language has been carried out by paying particular attention to impromptu conversation, considering face-

to-face and telephone interactions, but also storytelling and interviews; however, multiple studies on the topic also provide data obtained from semi-spontaneous interactions and task-oriented conversations.

In particular, the overall picture drawn in this first chapter has also allowed to pinpoint some of the main difficulties encountered when investigating phenomena like hesitations, such as the inherent ambiguity of the label “hesitation” and its overlaps, the question of inter-speaker variability, as well as the difficulty in comparing data from different corpora. As it was recently underlined by De Iacovo *et al.* (2020), the evaluation and the interpretation of pauses in speech, either filled or unfilled pauses, require special attention and carefulness since many factors and conditions need to be considered. Firstly, individual variation is an aspect that needs to be taken into account; speakers differ enormously in the way they use hesitations (Tissi 2000; Clark & Fox Tree 2002; Lickley 2015; Betz & Kosmala 2019; Kosmala & Morgenstern 2019; Staley & Jucker 2021) and, in various contexts, speakers’ individual preferences can be considered as characteristic of a particular geographical area (Giannini 2003a: 168; Tottie 2014b). Secondly, multiple researchers have underlined the difficulties in comparing works on hesitations due to the diversity of the materials and the different methodologies adopted for the analysis (Borfeld *et al.* 2001: 129; Giannini 2003a: 163). Finally, it is also important to consider that different corpora may adopt different transcription principles (Jucker 2021), not to mention the fact that there are various ways in which phenomena like hesitations can be considered, counted and categorized (Borfeld *et al.* 2001).

As previously shown in §1.3, there is no universally agreed-on definition of hesitations, just like a wide variety of linguistic and paralinguistic features can be labelled under this term. In this investigation, the focus will be on filled pauses; nevertheless, prolongations and pauses of silence will also be considered. Filled pauses are intended as vocalizations consisting of schwa [ə] or vowel sounds, which can be accompanied by nasalization ([ə̃m]) and/or lengthening ([ə:] [ə:m]) (Tottie 2011: 173). Prolongations, instead, involve the lengthening of either a vowel or a consonant sound, typically occurring at the end of a word. This form of hesitation is particularly relevant in Italian, since it is more frequently used compared to other hesitation phenomena such as vocalizations (Di Napoli 2020; Schettino & Eklund 2023). Finally, unfilled pauses are also taken into account. In line with Maclay & Osgood (1959), unfilled pauses are considered as silences of unusual length, which makes it possible to distinguish them from the necessary biological need to pause when speaking. As pointed out by Poyatos (2002b: 301):

There are no ‘pauses’ or ‘interactive silences’ of any type we can truly call ‘unfilled’; or, put differently, a pause or silence would not be such, because, by definition, it is really an interactive segment with specific structure determined by the co-occurrent behaviors [...], which in turn affect the preceding behaviors (in their decoding) and the succeeding ones (in both encoding and decoding).

In this investigation, however, pauses of silence are mainly observed in combination with the other hesitation phenomena examined. This is due to the need to ensure the possibility to confront the data obtained from the present analysis of film speech with a reference corpus in which silences, or even prolongations, may not be reported and signposted. On the other hand, pauses filled with vocalizations present prototypical transcription forms and language-specific conventions that make it possible and easier to carry out queries in the investigation of corpora.

In this regard, it is important to remark that this investigation does not include in the analysis of hesitations those linguistic phenomena often studied in concomitance with silences, vocalizations, nasalizations and prolongations such as repeats and false starts, generally referred to as disfluencies. Based on the reference literature mentioned in the current chapter, the more general inclusion of disfluencies in the analysis would fuel a negative view of hesitations as mere errors and, consequently, flatten the possible discussions on the topic. In the light of Tottie’s study (2011), the current analysis supports the idea that hesitations constitute a communicative tool that speakers can resort to for different purposes. In this respect, my use of the words “hesitation”, “filled and unfilled pauses” or “fillers” wants to be free of any connotation; they constitute neutral labels used to refer to pragmatic markers that should be investigated in their own terms, trying to understand their coding processes in film speech and their functions within the mechanisms of on-screen interaction in a fictional context (see Poyatos 2002b: 301).

The present survey aims to observe hesitations with regard to their uses and meanings, based on the cognitive, repair, discourse management and interpersonal functions established in spontaneous language, but also considering the prototypical functions identified in constructed fictional dialogue. Most notably, in the specific context of film speech, it is important to pay attention to the fact that scriptwriters use these elements in a stereotypical way to foster specific meanings to the audience, such as a sense of orality and naturalness, as well as a shortcut to characterization and characters’ communicative behaviour (Jucker 2015b). These uses may not faithfully coincide with how speakers actually

resort to hesitation phenomena in spontaneous conversation. Jucker (2015b), for instance, noted that hesitations like *uh* and *um* can be used by fictional authors to characterize a lie (*ibidem*: 176), while Villar & Castillo (2017) showed that, generally speaking, *um* is more often used in true statements than in false ones³⁹. While it is important to draw attention to stereotypical meanings associated to these elements — often exaggerated and taken to the maximum level to get easy access into the viewer's minds — this investigation does not want to flatten the possible discussions by considering hesitations merely as a stereotypical way to express trouble and errors. Rather, the analysis will explore how these pragmatic elements are displayed in constructed fictional dialogue through a diachronic perspective, drawing upon previous studies conducted on written language and fictional texts over time (Jucker 2015a; Rühlemann & Hilpert 2017; Tottie 2017).

The results obtained from previous diachronic research provide a first glimpse at any attested development, which may reflect changes in the fictional, literary and journalistic conventions rather than any changes in the use of language in the real world. This is also relevant for the current study, since the objective is to offer a diachronic investigation not only of the procedures chosen by translators to render hesitations in the Italian adaptations, but also of the evolution of film scripts and the role of these elements in film dialogue, meaning any changes in the conventions regarding orality in audiovisual translation (AVT).

An extensive and comprehensive analysis of film speech and orality in audiovisual translation will be provided in the second chapter of this dissertation.

³⁹ Villar & Castillo's (2017) data was obtained by studying the episodes of the British comedy television programme *Would I Lie to You* (BBC One) in which two teams of three celebrities have to convince the opposing team of the veracity of personal facts previously verified as true or false. The goal of the opposing team is to understand whether the fact narrated is true or not (*ibidem*: 552).

Chapter Two

Audiovisual Dialogue and Dubbing

Dubbing “requires a complex juggling of semantic content, cadence of language and technical prosody ... while bowing to the prosaic constraints of the medium itself”
(Whitman-Linsen 1992: 103-104).

The aim of the second chapter is to provide an extensive and comprehensive analysis of audiovisual dialogue. By focusing on film speech as well as on TV dialogue, it aims to describe the complex nature of audiovisual texts, the participation framework and the multiple functions dialogue fulfils. The objective is to observe dialogue in relation to its adherence to the codes of naturalness and realism, drawing attention to the way orality is conveyed in both original and dubbed products. By reviewing studies conducted on AVT and prefabricated orality, the chapter draws attention to the typicalities of dubbing, also focusing on the phenomenon under investigation, i.e. hesitations. The discussions presented in the chapter also allow to pinpoint some of the insights obtained regarding the diachronic evolution of audiovisual language and its adherence to spontaneous speech, alongside the evolution of Italian dubbing over the years.

2.1. The Hybrid Nature of Film Dialogue

The language of audiovisual texts such as movies and TV series is a particular type of language, “written to be spoken as if not written” (Gregory & Carroll 1978: 42); that is, a speech originally created in the written form of a script, then performed orally by actors. As a result, the mode of speech typically found in films and television — but also in plays and radio talk (*ibidem*) — presents some linguistic features that make it “distinguishable from other written texts through a greater comparative incidence in them of the favourite grammatical patterns of spontaneous speech” (Gregory 1967: 192), as well as from impromptu conversation itself. These texts exhibit a lower frequency of grammatical and lexical features typical of spontaneous speech, due the fact that texts associated with the

audiovisual medium “probably use language to create their situations to a greater extent than we do when speaking spontaneously” (*ibidem*).

Since the language of cinematic products comprises a selection of linguistic features typical of both the written and the spoken mode, Gregory’s (1967) considerations provide the groundwork to discuss the peculiarities of film speech and its hybrid nature. These dialogues are prepared in advance; nevertheless, they are always written with the final aim of being delivered orally. In fact, when moving from the script to the vocal delivery, the actor must try to hide the pre-planned nature of the text by performing it in such a way that it should sound unscripted or improvised. This particular sense of orality sought by both writers and performers of audiovisual dialogues has been referred to as “recited speech” (Nencioni 1983), “staged orality” (Heyd 2010), and “prefabricated orality” (Chaume 2004b; Baños Piñero & Chaume 2009), while Marzà & Chaume (2009) refer to an oral register that could be defined as “false spontaneous”. These labels also underscore the fact that, despite being uttered orally, fictional texts are carefully constructed by a team of professional writers (Heyd 2010: 60), who stage items typically associated with oral and spontaneous interactions in their scripts in order to convey specific information to the audience. The aim is not to perfectly replicate spontaneous conversation, but rather to select specific features typical of oral language that should recreate a spontaneous-sounding conversation, i.e. mimic impromptu interactions. In this regard, various scholars have investigated the language of fictional dialogue as consisting of a combination of features typically used in both written and oral texts (Whitman-Linsen 1992; Baños Piñero & Chaume 2009; Chaume 2004, 2012; Pavesi 2008, 2018; Romero-Fresco 2009a, 2012; Zabalbeascoa 2012); i.e., “an intermediate genre between the spoken and the written domains” (Heyd 2010: 60). Hence, the language of audiovisual texts is said to have a hybrid nature since dialogues are “characterized by a strange kind of oral discourse, an *orality* which may seem spontaneous and natural, but which is actually planned or, as we might say, false, *prefabricated*” (Chaume 2012: 82, emphasis in the original).

Much research has been conducted on the extent to which the discourse of audiovisual products such as films and TV series — also defined as “telecinematic language” (Piazza *et al.* 2011) — comes close to spontaneous spoken language, outlining the unique and specific peculiarities of audiovisual dialogue. The present contribution specifically focuses on cinematic products and film speech; nevertheless, some of the discussions examined in this chapter also apply to television dialogue. In fact, even though “[p]roducts for cinema and television follow specific conventions and fulfil viewers’ expectations in

different ways” (*ibidem*: 1), film and television discourse share some similarities such as the participation framework (Bubel 2006, 2008; Dynel 2011a) — namely, “the creators and the recipients of a fictional artefact” (Locher & Jucker 2021: 251)⁴⁰ — and some analogous functions (Kozloff 2000; Bednarek 2018)⁴¹. The studies conducted on telecinematic language and, more specifically, on film speech could be distinguished in two lines of research relevant for the present discussion: on the one hand, the investigation of the linguistic features of the original, non-translated speech, focusing on its degree of similarity to spontaneous conversation, as well as on specific linguistic features or issues (see Taylor 2004; Quaglio 2009; Valdeón 2009; Freddi 2011; Forchini 2012; Bublitz 2017; Bednarek 2018; Pavesi 2020); on the other hand, the examination of film speech from a translational point of view (see Pavesi 2005, 2019; Valdeón 2008, 2011; Romero-Fresco 2009a, 2009b; Baños Piñero *et al.* 2013; Sánchez-Mompeán 2020b, 2023b; Baños Piñero 2024).

In line with the abovementioned lines of research, the present discussion will provide an overview of the characteristic features of film speech and, more generally, of telecinematic discourse; i.e. its functions, its inherent verbal realism, and the way it is produced and transmitted to the audience. Subsequently, studies conducted on the translation of film dialogue will be reviewed, specifically focusing on the dubbing process.

2.2. Film Speech and Orality: A Structural and Functional Analysis

2.2.1. Orality and Spoken Language in (Non-)Fictional Dialogue

Before diving into the discussion of orality features and verbal realism in film speech, a distinction between orality and spoken language must be established. As reported by Bublitz

⁴⁰ This concept was introduced in the context of media by Goffman (1981). The author suggests that “[w]hen a word is spoken, all those who happen to be in perceptual range of the event will have some sort of participation status relative to it” (*ibidem*: 3) and that “[e]very utterance and its hearing bear the marks of the framework of participation in which the uttering and hearing occur” (*ibidem*: 4). The author then describes the participation framework as “the circle, ratified and unratified, in which the utterance is variously received, and in which individuals have various participation statuses” (*ibidem*: 226).

⁴¹ On the one hand, films constitute a single isolated narrative experience, while television series present a serial nature that requires a consistent exposure to the story (Piazza *et al.* 2011: 1). In addition, films and TV series differ in various aspects such as their business models, their target audiences, rules and regulations, as well as their production processes (Bednarek 2018: 6). Nevertheless, the discourse of films and television presents a similar participation framework (see Bubel 2006, 2008; Dynel 2011) and similar functions (Kozloff 2000; Bednarek 2018). Some of the similarities between the discourse of films and TV series have also been more recently pointed out by Werner (2021) in his investigation on the developments of telecinematic discourse over time. The analysis shows that film and TV series do present some analogous patterns and trends, reinforcing the correctness of labels such as “telecinematic language” or “telecinematic discourse” (*ibidem*: 65).

(2017), spoken language can be defined as “the product of the act of speaking”, that is, “the vocally produced segmental and non-segmental (prosodic) sounds which prototypically are accompanied by supporting kinesic signs (gestures, mimics, gazes and posture)” (*ibidem*: 241). Orality can be understood as having to do with “speaking as a mode” (*ibidem*) and as a concept. In the former case, “orality refers to a concrete (phonic) realization of language” (*ibidem*); in the latter, orality is an abstract principle that can be applied to both written and spoken language (*ibidem*). Based on this distinction, “[o]ral features are thus spoken features in the sense that they are vocally produced in the act of speaking. As such they are a characteristic, though not a constitutive feature of spoken language” (*ibidem*).

Based on these premises, Bublitz (2017) connects the concepts of orality and spontaneous speech with the language of fictional and non-fictional dialogues. In the author’s view, non-fictional dialogue is characterized by orality both as a mode and as a concept. In fact, non-fictional dialogue can be defined as spoken dialogue that displays features of orality ascribable both to the unplanned nature of the conversation and to the socio-cultural speech situation in which the verbal exchange occurs, as well as to the dynamic context in which interlocutors negotiate meaning (*ibidem*: 242). Conversely, in fictional dialogue, “orality is artificially generated in the act of writing by the author, who copies the concept from natural face-to-face conversation” (*ibidem*). Hence, in this case, orality “does not dynamically emerge as the function of a situation and context but is previously designed by a non-participating other” (*ibidem*). In the context of telecinematic products, this “non-participating other” has also been referred to as “collective sender” (Dynel 2011a), who designs the plot of the story and the dialogues to fulfil specific functions and transmit specific information to the audience, who then constructs and infers meanings from them. As pointed out by Messerli (2017), any form of fiction is the outcome of a creative process resulting in a fictional artefact that will have to be received by some recipients: “these processes will lead to a form of mediated communication between creators and recipients, in which meaning-making and understanding processes are anchored on the cultural artefact” (*ibidem*: 26).

Further discussion on the participation framework, that is, “who interacts with whom on what levels” (Locher & Jucker 2021: 251) in this mediated form of communication, is provided hereunder.

2.2.2. The Communicative Structure of Film Dialogue

The wide range of information conveyed to viewers through dialogue is built on a complex system, which is far from the traditional dyadic model of communication where “the addresser sends a message to the addressee” (Jakobson 1960: 353). More specifically, the communicative system of telecinematic products involves a “screen-to-face discourse” (Bubel 2006: 46) in which there is more than one level of communication, namely the communication occurring between characters on-screen and the communication occurring from characters to viewers (see also Brock 2015 and Messerli 2017)⁴². Thus, the lines uttered by the on-screen actors are multi-addressed, that is, equally directed at the other actors involved in the scene (at the diegetic level) and at the audience who is watching (at the extra-diegetic level) (Bublitz 2017: 255; Locher *et al.* 2023: 1). In particular, scholars have investigated the role of viewers, traditionally conceived as unratiſed participants, eavesdroppers or overhearers (Kozloff 2000; Bubel 2006, 2008; Richardson 2010) — to use Goffman’s (1981)⁴³ terms — meaning that “they have no conversational rights or responsibilities” (Bubel 2008: 64) in the interactions depicted on the screen. Various scholars, however, consider this perspective inaccurate and support the view that audiences do “have a place within the *participation framework*” (O’Keeffe 2006: 18, emphasis in the original). In this regard, Dynel (2011a, 2014) puts forward an interesting hypothesis according to which once the speaker purposefully directs, more or less overtly, his or her utterance to an individual, the latter assumes the position of a ratiſed hearer. In fact, as also

⁴² For a review of the participation framework of telecinematic discourse, as well as of literary fiction and dramatic discourse, see Messerli (2017). For a detailed investigation of participation structures in TV programmes, see Bednarek (2010: 14-17) and Brock (2015).

⁴³ Goffman’s work on participation framework constitutes an important landmark in the investigation of the interaction between speaker and listener. Goffman (1963, 1967) distinguishes between “focused” and “unfocused” interactions; one type of focused interaction is verbal interaction, also defined as “talk” (Goffman 1981). In Goffman’s initial formulation, the production format of an utterance — that is, the speaker — can be distinguished between “animator”, “author” and “principal” (*ibidem*: 144); for the present discussion, however, the distinction established between “listeners”, “hearers” and “recipients” is particularly relevant. More specifically, hearers can be differentiated between “ratiſed” and “unratiſed” (*ibidem*: 9-10). Unratiſed hearers are then distinguished between “bystanders” and “overhearers/eavesdroppers”, depending on their intentional engagement in the act of listening; on the other hand, the distinction between “overhearers” and “eavesdroppers” depends on the speakers’ awareness — or lack of awareness — of the unratiſed activity. As for the ratiſed hearers, a distinction is made between the “addressed recipient” (explicitly addressed by the speaker) and the “unaddressed recipient” (who is not addressed) (*ibidem*: 133). As for audiences, when discussing theatrical performances, Goffman (1981) suggests that they should be considered as some “out-of-frame eavesdroppers”, listening to actors delivering their performance according to a script (*ibidem*: 83). As for radio talk audiences, Goffman claims that it “is treated as if it were a ratiſed participant, albeit one that cannot assume the speaking role” (*ibidem*: 234), in line with the notion of “imagined recipients” (*ibidem*: 138), also applicable to viewers of both television and films (Dynel 2017: 74).

pointed out by Bubel (2008: 66), in telecinematic products “[u]tterances are designed with overhearers in mind” insomuch that dialogue is “purposely designed for the viewer to overhear” (Kozloff 2000: 15). If the discourse is designed and produced to be transmitted to the audience, then viewers should not be considered as overhearers. On the contrary, they ought to be regarded as a separate type of ratified hearers, i.e. the recipient (Dynel 2011a: 1629; 2017: 74). To use Dynel’s (2014: 29) words:

Conversing characters bear resemblance to conversationalists in real-life interactions, the most significant difference being that the former’s interactions are, by nature, available to an outside hearer, i.e. the film viewer. This is the underlying goal of the film production crew, who design film discourse for the audience’s (recipients’) benefit.

In Dynel’s (2011a) work, the production crew is also referred to as “the collective sender”. In particular, according to Dynel, in film discourse the communication system works on two layers: the “collective sender’s layer” and the “fictional layer” (*ibidem*: 1634)⁴⁴. The collective sender constructs characters’ conversations and the meanings conveyed through them specifically for the viewer’s benefit (Dynel 2011a: 1642, 2011b: 463), while the fictional layer includes the fictitious world and the characters depicted on screen, where protagonists, places and events “are understood as representations of real people, real places and real events” (Messerli 2017: 31).

Therefore, film discourse works on two levels: the inter-character(s)’ communicative level and the recipients’ level (Dynel 2011a: 1632). The interactions on the characters’ level can be described by the same participation framework of real-life spontaneous interactions; however, the meaning conveyed through these engagements is contingent upon the collective sender (*ibidem*: 1634), i.e. the scriptwriters, actors, directors, cameramen, editors, and so on (*ibidem*: 1632). The collective sender designs the film and the on-screen

⁴⁴ More recently, Locher & Jucker (2021) introduce a model of communication of performed fiction that further extends the complexity of the participant structure previously proposed. The authors specifically focus on both the recipient’s and the production’s side of communication. When discussing the recipient’s side of communication, the authors distinguish between “core recipients” (i.e. target audience: known, ratified and addressed), “incidental recipients” (i.e. irregular and chance viewers who are known to be there but are not ratified) and “accidental recipients” (i.e. who should not be there and are unrated, for instance, children watching a movie suitable only for adults) (*ibidem*: 44-45). On the other hand, when discussing the intricate nature of the production side of communication, Locher & Jucker distinguish between the intradiegetic level of communication between characters, the extradiegetic level of communication between creators and the audience, and what the authors define as the “superdiegetic level” (*ibidem*: 53), which typically includes a live audience (*ibidem*: 56). Locher & Jucker’s model, however, seems to be more applicable to television than to cinematic products. Since this study specifically focuses on films and film speech, Dynel’s (2011a) work remains the most relevant framework for the present discussion.

interactions — that is the fictional layer — at all stages of the production process (pre-production, production, and post-production) (*ibidem*: 1634). Viewers should be able to interpret the scene and infer the correct meanings — that is, the meanings intended by the collective sender — based on their background knowledge (*ibidem*: 1636). This process is not only relevant to verbal but also to non-verbal communication, for example, facial expressions — that is to say, all the stimuli audiences receive from the various semiotic modes involved in audiovisual products (Dynel 2011a, 2011b). As claimed by Dynel (2011a: 1642):

This is because film discourse is designed for viewers, and their listening and watching are by no means surreptitious or accidental. On the contrary, recipients' listening to scripted dialogues is fully legitimate, being presupposed and carefully facilitated. [...] Essentially, although character converse with one another, their conversations are actually directed to the recipient. Characters may not talk to us, but they certainly talk for us.

In other words, although viewers are not the direct addressee of characters' dialogue, they still are its recipient (Locher *et al.* 2023: 21)⁴⁵. Writers create dialogues based on the fact that viewers should be able to use them to make correct assumptions about the meanings conveyed through them, while also relying on other factors such as viewers' background cultural and cognitive knowledge, episodic and conceptual memory, emotional frames, and interactions with other fictional artifacts (Bublitz 2017: 245; Locher *et al.* 2023: 8). In this sense, fictional products should be considered as “dynamically created in the sense that their meaning [...] is essentially made up of the communicative efforts of the sender and the processing effort of the recipient” (Locher *et al.* 2023: 8).

The collective sender, aware of the fact that the conversations happening on-screen at the inter-character level undergo viewers' interpretation (Dynel 2011a: 1632), designs exchanges with the aim to transmit specific information to viewers. At the same time, viewers are induced to think that those interactions belong entirely to characters (*ibidem*: 1635), forgetting about the production crew. In other words:

⁴⁵ As also underlined by Dynel (2011a: 1634), there are cases in which “characters seem to acknowledge the presence of the audience”; this occurs, for instance, in films like *High Fidelity* (Frears, 2000) and in TV series like *Fleabag* (2016-2019) and *House of Cards* (2013-2018), where characters frequently break the fourth wall to address viewers directly. In these cases, the viewer could be described as “an explicit, ratified participant who is given the illusion of first-level participation on a par with the other inside participants” (Sorlin 2015: 134). Indeed, as pointed out by Dynel (2011a: 1634), “this type of address is still a matter of the fictional layer designed by the collective sender”; thus, although the two communicative levels (i.e. the characters' level and the recipients' level) seem to merge, “the distinction between the two layers still obtains” (*ibidem*).

Normally, the viewer willingly forgets about the real production end and engages in characters' interactions. Obviously, some chunks of discourse or cinematic techniques may entail conscious inferring on the viewer's part, especially if they are cognitively demanding. Nonetheless, this does not normally entail the viewer's leaving the fictional world and investigating, for example, what the scriptwriter and the actor [...] aim to convey (Dyrel 2011a: 1632).

This willingness of the audience to step into the fictional world and accept its illusion has been defined as "suspension of disbelief". The phrase "willing suspension of disbelief" was first used by Coleridge (1817) "to explain how readers experience poetry and the hypnotic process whereby they willingly allow themselves to be transported from story spectator to story participant" (Kallan 2009: 65). This concept is applicable to what occurs to audiences when watching a film; that is, viewers are still able to enjoy the product, regardless of the potential linguistic and prosodic discordances they may find, especially when watching a movie adapted from a different language (on the concept of suspension of linguistic disbelief, see Romero-Fresco 2009a, 2020; on the suspension of prosodic disbelief, see Sánchez-Mompéán 2020). As claimed by Ferri (2007: 16), "the synergy between viewer and film creates the filmic experience" and dialogue plays a central role in enhancing this synergetic bond.

All things considered, in film dialogue, meanings are intentionally produced and conveyed by the collective sender in order to suspend viewers' disbelief while also conveying information necessary for viewers to understand the story, characterize characters and the relations that exist between them. In other words, fictional dialogue fulfils a number of functions relevant for the narrative development and the elements found in the verbal exchanges among characters are intentionally and consciously included by the collective sender solely to favour viewers and their interpretation of the overall product (Kozloff 2000: 19). As claimed by Zago (2021), "the illocutionary force of an utterance in a film or a TV series is "greater" than that of an utterance in spontaneous unscripted spoken English, as what the characters tell each other regularly acquires additional functions for the audience" (*ibidem*: 168).

An overview of the wide range of functions dialogue fulfils and the information transmitted to viewers through the complex system at issue is provided below.

2.2.3. A Functional Analysis of Telecinematic Dialogue

Before diving into the linguistic characteristics of audiovisual texts, it is important to briefly discuss the main functions dialogue fulfils in fictional telecinematic products. In fact, this aspect inevitably impacts on the linguistic choices made by the professionals involved in writing the original script and, eventually, adapting it into another language. In particular, the present discussion mainly relies on Kozloff's (2000) model of the functions of cinematic dialogue, as well as on Bednarek's (2017, 2018) adaptation of this paradigm to TV series.

One of the basic functions of fictional dialogue has already been introduced in §2.1, that is, the requirement of fictional dialogue to exhibit "adherence to the code of realism" (Kozloff 2000: 33); namely, to provide authenticity to those engaged in the viewing process. At the same time, there are various concomitant functions that telecinematic dialogue fulfils (Kozloff 2000; Pavesi 2005: 31; Bednarek 2017, 2018; Locher & Jucker 2021: 128), which can be distinguished into two macro-categories: functions related to the entertainment nature of audiovisual products, on the one hand, and functions related to commercial, ideological and aesthetic factors, on the other (Kozloff 2000: 33-34)⁴⁶. The former group includes those functions that are considered to be fundamental to the communication of the narrative (*ibidem*: 33); indeed, dialogue constitutes one of the main elements to actually create the fictional world in which the narrative occurs, anchoring characters to the diegetic world and allowing viewers to orient, and constantly reorient, within the narration (*ibidem*: 33-35). This function has been defined as "anchorage of the diegesis and characters" (*ibidem*: 33). According to Bednarek (2017), dialogue can also be associated with the function of "anchorage to modality" in the sense that it allows viewers to interpret the credibility or reliability of a message, i.e. if something is real in the diegetic world or if it is a dream, a fantasy or a desire (*ibidem*: 133)⁴⁷. Moreover, dialogue provides background information to viewers, explaining the narrative development and the causality of events (Kozloff 2000: 37; Bednarek 2018: 42). Indeed, film dialogue represents one of the main resources to unravel

⁴⁶ Kozloff (2000: 33-34) established a total of nine functions: 1) anchorage of the diegesis and characters; 2) communication of narrative causality; 3) enactment of narrative events; 4) character revelation; 5) adherence to the code of realism; 6) control of viewer evaluation and emotions; 7) exploitation of the resources of language; 8) thematic messages/authorial commentary/allegory; 9) opportunities for "star turns". Subsequently, Bednarek (2017) proposes a tenth function, that is, the creation of continuity. This function concerns TV dialogue and serial films such as *Harry Potter*, *Twilight* or *James Bond*, namely, movies that include more than one chapter (*ibidem*: 142).

⁴⁷ Bednarek (2017) puts forward two additional sub-categories to the function dubbed by Kozloff (2000: 34) "exploitation of the resources of language", that is, linguistic innovation (i.e. inventions and play with language) and intertextuality (i.e. allusions to other texts) (Bednarek 2017: 138). These additional sub-categories mainly regard television dialogue.

the events and make the story advance (Pavesi 2005: 30); it structures what happens in a scene, but it can also function to manage scene changes (Bednarek 2018: 45). Dialogue also represents a fundamental tool to display characters, their identity and narrative roles, their personality and psychological traits, as well as the different relationships that exist between them (Kozloff 2000: 36, 43-47; Culpeper 2001; Pavesi 2005: 31; Bednarek 2010, 2018: 46-52, 2023; Locher *et al.* 2023: 46-47)⁴⁸. In addition to these narrative functions, dialogue can also communicate thematic and moral messages (Kozloff 2000: 56-58) through allegory, irony and on-screen verbal storytelling (*ibidem*: 52-55). Dialogue can be used poetically, as well as to convey jokes and humour (*ibidem*: 52-53), to control viewers' evaluation and emotions, guiding the responses of the audience and drawing their attention toward certain specific plot elements (*ibidem*: 49-50).

To sum up, “dialogue must provide adequate characterisation as well as communicate the who, where, what, how, and why of the narrative to the audience” (Bednarek 2018: 53). This allows to put forward an important aspect underlined by Kozloff (2000: 34) in relation to movie dialogue and by Bednarek (2018: 76) in relation to TV dialogue, namely the fact that telecinematic dialogue is polyfunctional. In other words, a piece of dialogue usually fulfils multiple functions at the same time. As pointed out by Kozloff (2000: 62), “dialogue is not the *only* means of accomplishing these ends” (emphasis in the original); considering the multi-semiotic nature of audiovisual texts, sounds and images “work together to create a universe that is an integral part of the film narrative structure” (Bosseaux 2015: 16). Meanings are conveyed through both linguistic and extra-linguistic elements, the latter including gestures, actions, facial expressions, sets, clothes and mise-en-scène (*ibidem*). Bednarek (2018), for example, emphasizes how realism can be created by means of authentic music and sound effects, as well as through camera techniques; as a matter of fact, “dialogue works in conjunction with camera shots to locate the narrative in space, in conjunction with shot sequencing to clarify the modality of happenings, and in conjunction with performance and shot content to elicit viewer's emotional responses” (*ibidem*: 76). Nevertheless, considering the interests of the current investigation (i.e. film speech and its dubbed translation), the discussions presented in this work will focus on the verbal component of motion pictures, that is, dialogue. This is particularly relevant in relation to the translation component of this study since, in dubbing, visuals remain the same; what is inevitably

⁴⁸ As for the phenomenon studied in the present investigation, hesitations are stereotypically used in scripted dialogue to make listeners infer the awkwardness of a character; in other words, to portray a character as embarrassed, shy or insecure (Richardson 2010: 64; Bruti 2019: 197).

subjected to change is the text, i.e. the dialogues uttered by the actors on screen and the way their lines are addressed to both characters (i.e. at the diegetic level) and viewers (i.e. at the extradiegetic level), as well as the way the new dialogue interacts with the visual elements.

As briefly pointed out above, the functions of dialogue discussed here need to be understood in relation to the participation framework through which information is conveyed to the audience. As a matter of fact, this aspect inevitably affects the linguistic choices made by the professionals involved in the pre-production, production and post-production processes of creating the original script and then translating it and adapting it into a target language.

Before discussing the linguistic features typically included in audiovisual dialogues and the distinctive traits of telecinematic language, it is necessary to consider several factors that influence the selection of orality features in fictional dialogue.

2.3. Scripting and Performing Orality: A Matter of Compromise

As previously noted, scripted fictional dialogue is a particular type of communication that, even though writers pursue realism and naturalness, can never fully correspond to the unscripted nature of spontaneous conversation. As a matter of fact, although dialogue gives audiences the impression of being realistic, a selection of features is always needed in order to avoid hindering viewers' comprehension and interpretation of the dialogue (Chaume 2012: 81), allowing them to follow the narrative events and draw the correct inferences regarding the story.

As pointed out by Sheila Lawrence, screenwriter of TV series such as *Ugly Betty*, *Gilmore Girls* and *Desperate Housewives*, it is all about finding the balance “between wanting to sound the way we imagine people to sound, but not be quite as ineloquent as people often are in “real” life” (Bednarek 2019: 19). An excessive number of oral features would indeed make the viewing experience unengaging and hard to follow, while a very small number of elements typical of oral language would make dialogues sound unnatural and artificial (Valdeón 2009: 206). In his investigation of fictional language, Jucker (2021) more recently points out that “a higher incidence [of orality features] would reduce the comprehensibility and intelligibility of what is being said by the characters as well as the overall coherence”; consequently, “it appears that the performance of orality can be achieved with a reduced frequency of features of reality” (*ibidem*: 356), which do not perfectly replicate the complexity of spontaneous language but serve as indexical cues. In fact,

scriptwriters are said to be working on a default assumption of “reduced orality” (Bublitz 2017: 244). To use Valdeon’s (2008) words, the language of audiovisual products can be said to be the result of a process of intralingual translation, since the language of the script always undergoes some kind of transformation, which “implies the deletion of a number of elements of orality, because some of the characteristics of spontaneous speech (such as repetitions, rephrasing, etc.) would render the final text inadequate for the requirements of an audience in search of entertainment” (*ibidem*: 120). A reduced orality, however, still manages to imitate unscripted conversation and evoke a sense of spontaneity among viewers (Bednarek 2018: 123; Jucker 2021: 356).

In this respect, the role of actors’ improvisation in conveying a sense of spontaneity to the dialogue has also been highlighted. Baños Piñero (2024), for instance, compared the original script of a Spanish TV series (*Siete vidas*) with the dialogues actually uttered in the soundtrack, noticing how actors can intervene both in terms of form and content (*ibidem*: 35-36). Likewise, 20 years before, Taylor (2004) observed the same in his investigation of the movie *Notting Hill* (Michell, 1999). Taylor (2004: 80) reports that even though “the scriptwriter made an effort to include genuine spoken language features (discourse markers, ellipsis, hesitation, etc.)”, when the script is acted out, the number of items typical of spontaneous spoken language increases (*ibidem*: 84); in other words, actors add these elements as “they create for themselves an (artificial) context of situation and attempt to ‘live it’” (*ibidem*: 79-80), adapting the language of the script to the context in which they find themselves in order to interact with the scene in a more realistic way (*ibidem*: 80). As also reported by Nencioni (1983), the lines written by the original scriptwriter may be performed differently with regard to spontaneity as the result of the creative contribution of the actor (*ibidem*: 176). To use Jucker’s (2021: 342) words: “If their utterances appear spontaneous and unscripted, it is because they are performed in that way”.

At the same time, albeit improvisation occurs, it should be noted that dialogues go through a complex process of rehearsal, editing and approval. As pointed out by the writer and producer David Mandel⁴⁹, television dialogue, for instance, “has the value of the edit room, which is to say, I’ll go through and I’ll pull out annoying pauses, and *ums* and *uhs* and whatnot, [...] I can take that struggle out” (Bednarek 2019: 22). Similarly, when answering the question “How naturalistic should TV dialogue be?”, he states: “I do think there are

⁴⁹ David Mandel is the Emmy Award-winning showrunner and Executive producer of *Veep*; he also wrote for shows such as *Saturday Night Live* (1992-1995) and directed shows like *Curb Your Enthusiasm* and *The Comedians* (Bednarek 2019: 6).

limits. For example, I could take a 55-minute show and I can go through and tighten up some of the pauses, the *ahs* and the *uhs* and whatnot, and I might get two minutes out” (*ibidem*: 24). On the one hand, these words underline the fact that scriptwriters must abide by the constraints of the audiovisual media, along with the norms and conventions established for the telecinematic product (see also Baños Piñero & Chaume 2009; Valdeón 2011: 230; Bruti 2019: 193). On the other hand, Mandel’s words show that certain features of orality, among which hesitations, are often considered “mishaps of performance” (see Rossi 1999: 181) regardless of their interactive value; as a result, they are easily removed from speech for the sake of time constraints and speech smoothness. As a matter of fact, although professionals recognize that the inclusion of linguistic devices typical of spontaneous speech is necessary and omitting them could limit the dialogue (Bednarek 2019: 26-27), they also point out that “with pure verisimilitude you do run the risk of losing some drama” (*ibidem*: 25)⁵⁰. Considering the editing process dialogues go through, it is possible to state that, if included in the final version, features of orality are purposefully and specifically used to achieve particular effects. This also comprises pragmatic features. In fact, as suggested by Kozloff (2000), the inclusion of these elements constitutes a specific stylistic choice (*ibidem*: 64) since “pauses, repetitions, and hesitation phenomena are edited out unless they fulfil a specific narrative function” (Bednarek 2018: 213).

A more detailed examination of the distinctive traits of telecinematic language is provided hereafter.

2.4. The Linguistic Features of Audiovisual Dialogue

Both professionals (see Bednarek 2018, 2019; Minutella 2021a: 86)⁵¹ and scholars (Taylor 2004; Chaume 2004b, 2012; Pavesi 2005, 2008; Quaglio 2009; Valdeón 2009; Romero-Fresco 2009a, 2012; Zabalbeascoa 2012; Bruti 2019; Baños Piñero 2014, 2024) have recognized the importance of including orality markers in dialogue in order to make the product sound authentic, natural and spontaneous, also highlighting the impact that credible-sounding dialogues have on the commercial success of films and TV series (Bruti 2019: 192-193).

⁵⁰ Bednarek (2019) reporting the answer of the writer/producer Jane Espenson (*Buffy the Vampire Slayer*, *Once Upon a Time*, *Star Trek*, *Dr. Who*) to the question “How naturalistic should TV dialogue be?”.

⁵¹ In her book, Bednarek (2019) presents insights from discussions with five Hollywood screenwriters regarding writing dialogues for television and the way language is used in TV series. More recently, Minutella (2021a) reports some of the discussions she had with Italian dubbing professionals working for cinema and television on the use of Anglicisms in Italian dubbing.

Features of orality have indeed come to be recognized as an inherent part of the staged orality of performed fiction (Jucker 2021: 357). Over the years, a significant body of research has emerged on the language of fictional products, like movies and TV series, with the aim of establishing how close fictional audiovisual language is to real-life conversation and what are the fundamental elements of spontaneous speech that are employed in fictional texts and adapted to the fictional context in order to achieve an acceptable degree of naturalness, making the audience suspend their disbelief. In doing so, Corpus Linguistics (CL) has proven to be a powerful methodology to observe the similarity between fictional dialogue and face-to-face conversations, also “confuting the idea that movie language cannot be used to explore features of spoken language” (Forchini 2013: 505). As also noted by Pavesi (2022: 550), “[c]orpus linguistics has been paramount to the investigation of naturalness, orality and authenticity of film dialogue; at the same time, it has contributed to defining film language’s specificity”. In fact, as recently noted by Zago (2021), “film discourse presents a good degree of stability and conventionalisation” (*ibidem*: 169), i.e. it constitutes a register marked by regular and conventionalised linguistic patterns (*ibidem*: 173). Since film discourse aims to reproduce conversational settings — that is, characters interacting with one another — the linguistic stability of film discourse is related to interactional features typical of any conversation (*ibidem*). In the context of telecinematic language, spontaneous conversation is taken as a reference point to establish the level of orality of a fictional text, while corpora of conversational spoken language are used for comparison purposes; that is, to establish the similarities and/or differences existing between prefabricated orality and spontaneous speech (*ibidem*)⁵².

Further discussions on the specificities of telecinematic language established through a corpus-based approach are provided in the following section. Considering the interests of this research, which focuses on movies originally produced in the English language and dubbed into Italian, the discussions illustrated here will mainly refer to studies conducted on original Anglophone films and/or TV series.

⁵² It must also be noted that some scholars have recently emphasized the need to analyse telecinematic discourse per se, in its own right, rather than comparing it to unscripted data (Bednarek *et al.* 2021).

2.4.1. Corpus Linguistics Applied to Audiovisual Dialogue

Although no corpus can capture the language system as a whole (Forchini 2013: 506), much progress has been made in establishing the defining features that characterize the language of audiovisual dialogues by means of corpus-based investigations⁵³.

As reported by Pavesi (2019b: 330), “[a] corpus is a principled collection of texts stored electronically to be analysed quantitatively and qualitatively by using computer-assisted techniques”. In corpus linguistics, tools such as frequency lists and concordances of individual words and clusters — such as lexical phrases, n-grams or fixed expressions (Baker 2006: 91-92; Pavesi 2019b: 319) — make it possible to inquire into the distinctive features, norms and patterns of audiovisual texts, transcending single case studies. This, in turn, allows for generalizations on the typicalities of audiovisual registers (Pavesi 2019b: 315) through both quantitative and qualitative analyses. In fact, as also pointed out by Halliday (2003: 23), quantitative features are an intrinsic component of the meaning potential of language; this data, however, needs to be functionally interpreted through narrowly focused research so as to provide qualitative explanations for the phenomena identified through quantitative analysis (see Aarts 2001).

For this purpose, researchers have focused on the collection of transcriptions of film and television dialogue in order to study telecinematic language empirically. These collections of texts have been developed using various criteria: focusing on a single product — as in the case of Quaglio (2009) and Bednarek (2010) in their analyses of television dialogue in the TV series *Friends* and *Gilmore Girls*, respectively — or on a group of products selected for the aims of a specific case study and sharing some common features, for example the same genre, as in the *Corpus of American Soap Operas* (Davies 2012)⁵⁴. Studies conducted on these collections of texts describe some of the linguistic features that typify conversation in specific fictional contexts, but do not allow to obtain comprehensive results on telecinematic language as a whole. To this end, it has been of pivotal importance

⁵³ A distinction needs to be established between corpus-based and corpus-driven approaches. Corpus-driven research is developed based on the observation data, that is, regularities and exceptions that bring, first, to a hypothesis, secondly, to generalizations, and, finally, to a theoretical statement (Tognini-Bonelli 2001: 17; Bruti 2020: 382-383). Corpus-based research, instead, is conducted in such a way that corpus data are used to validate or demonstrate already existing theoretical statements and hypotheses (Bruti 2020: 383). As summarized by Bruti (*ibidem*), “corpus-based studies are deductive in nature, as they start from linguistic information that generates a hypothesis, corpus-driven ones are essentially inductive, because the research is sparked off by the observation of data”. Nevertheless, scholars have moved towards the idea that they should not be conceived as non-binary opposites, but as “existing at theoretical poles on a linear scale, with most research falling somewhere in the middle” (Baker 2016: 19).

⁵⁴ Davies M. 2012. *Corpus of American Soap Operas*. Available online at: <http://corpus.byu.edu/soap/>.

to develop more extensive corpora that make it possible to examine telecinematic dialogue as a language variety; among these, important results have been obtained from the *TV Corpus* (Davies 2021)⁵⁵ and the *Sydney Corpus of Television Dialogue* (SydTV)⁵⁶ (Bednarek 2020), developed for the investigation of TV series, along with the *Movie Corpus* (Davies 2021)⁵⁷, the *North American Movie Corpus* (NAMC) (Veirano Pinto 2014) and the *American Movie Corpus* (AMC) (Forchini 2021)⁵⁸, developed for the examination of film dialogue.

With regard to film speech, for instance, Forchini's (2012) investigation allowed to showcase the similarities existing between film speech and spontaneous speech, hence claiming that film dialogue is closer to impromptu conversation than may be expected. The author provides an empirical study comparing spontaneous conversation (the *Longman Spoken American Corpus*)⁵⁹ with movie dialogue through a multi-dimensional (Biber 1988)⁶⁰ and a corpus-driven analysis of 11 American movies belonging to different genres (the *American Movie Corpus*). More specifically, the author distinguishes between comedies, non-comedies (namely, drama, sci-fi, thriller, action, adventure, crime, fantasy movies) and borderline films, i.e. those cases in which genre categorization is not clear-cut (*ibidem*: 54). Observing features typical of spontaneous speech, such as personal pronouns, coordinating conjunctions, imperatives, contractions and discourse particles, Forchini claims that the difference between impromptu conversation and film speech is not as significant as one may think; on the contrary, film dialogue considerably reproduces spontaneity both in

⁵⁵ The *TV Corpus* includes subtitles from 75,000 episodes, 325 million words and six English-speaking countries, from the 1950s to the 2010s (available online at: <https://corpus.byu.edu/tv/>).

⁵⁶ The *Sydney Corpus of Television Dialogue* (SydTV) constitutes a corpus of US American fictional television series. The corpus contains dialogue from one first-season full episode of sixty-six different series for a total of approximately 275,000 words (Bednarek 2020: 110-111). Half of the selected series are categorized as comedy genres, while the other half are categorized as drama genres (*ibidem*: 111). The corpus includes pilot episodes, final episodes, as well as other episodes occurring at the beginning, middle and end of the respective series so as to represent "different moments of textual time within the series" (*ibidem*). Finally, the corpus contains both network television (forty-two episodes) and cable television (twenty-four episodes) (*ibidem*: 114).

⁵⁷ The *Movie Corpus* includes subtitles from 25,000 movies, 200 million words and six English-speaking countries, from the 1930s to the 2010s (available online at: <https://www.english-corpora.org/movies/>).

⁵⁸ Further information and the corpus itself can be retrieved online at <https://www.americanmoviecorpus.net/>.

⁵⁹ The *Longman Spoken American Corpus* (LSAC) is extracted from the *Longman Spoken and Written English Corpus*. Together with the *Longman Written American Corpus*, the LSAC belongs to the *Longman Corpus Network* (Forchini 2012: 51).

⁶⁰ Multi-dimensional analysis is a corpus-based approach developed by Biber (1988) to inquire into register variation "by quantitatively detecting patterns of co-occurrence among lexico-grammatical features in texts" (Pavesi 2022: 550). The multi-dimensional analysis approach resorts to factor analysis to identify co-occurrent patterns across a wide range of linguistic features (Gray 2022: 241). This methodology had been employed to compare the register of spontaneous spoken English with film dialogue both synchronically and diachronically (see Forchini 2012; Veirano Pinto 2014; Zago 2016).

lexical and syntactic terms (*ibidem*: 115-116, 120). The only difference highlighted in her investigation concerns the fact that movie conversation has been labelled as abstract, while face-to-face conversation has been labelled as non-abstract (*ibidem*: 95-96). As for the rest, both conversational domains are described as “informal, non-narrative, situation-dependent and not highly persuasive” (*ibidem*: 119). Since these features have traditionally been linked to impromptu conversation, it is possible to conclude that movie language presents a considerable amount of spontaneity, regardless of film genre (*ibidem*).

Similarly, Veirano Pinto (2014) also carried out a multi-dimensional analysis, investigating 640 North American movies produced between 1930 and 2010, observing how multiple situational variants can impact the language of films. The author specifically took into account genre (comedies, dramas, action/adventure, horror/suspense/mystery), year of release, movie length, directors, rating, etc., to investigate whether these variants affect the degree of spontaneity of the dialogue and, if so, how. Results illustrate that factors like movie length and genre can have an impact on the degree of spontaneity of dialogues; in fact, longer movies (over 2 hours long) tend to present more markers of planned speech, as do action/adventure movies, where dialogue is used to “verbalize the events that serve a narrative purpose” (*ibidem*: 123).

Thus, even though film dialogue shares similarities with spontaneous speech, it does have its own specific features resulting from the dramatic and narrative principles that underpin these fictional products. The specificities that characterize the register of telecinematic language are further discussed below.

2.4.2. The Oral Register of Telecinematic Dialogue

Studies conducted on testing the degree of similarity between audiovisual dialogues and impromptu conversations have established that fictional dialogue is much neater and more well-organized than spontaneous conversation (Rossi 1999; Pavesi 2005; Quaglio 2009; Bednarek 2012). This is observed, for instance, by Rossi (1999) in his investigation of Italian film speech.

Rossi examined six movies (one dubbed and five original Italian movies)⁶¹ distributed between 1948 and 1957 with the aim of evaluating the peculiarities of film speech as against spontaneous conversation (*ibidem*: 95). For the investigation, the author selected

⁶¹ More specifically, Rossi (1999) investigates five Italian movies, namely *Ladri di biciclette* (De Sica, 1948), *Catene* (Matarazzo, 1949), *Totò a colori* (Steno, 1952), *Le amiche* (Antonioni, 1955), *Poveri ma belli* (Risi, 1957), and one original American film dubbed into Italian, i.e. *Nata ieri/Born Yesterday* (Cukor, 1951).

a few linguistic features related to the diaphasic dimension (formal/informal) (*ibidem*: 95-96) with the aim to examine different styles of Italian cinematic language, namely the functional use of certain linguistic choices such as the use of dialect as opposed to standard Italian (*Ladri di biciclette*), as well as the language of films adapted from a novel (*Le amiche*) or from another language (*Nata ieri/Born Yesterday*) (*ibidem*: 450). Despite the inevitable differences found among the styles of the six films observed, the analysis shows that a general simplification of the language can be noticed both at the pragmatic and at the morphosyntactic level with a low incidence of variability index across the six titles. More specifically, this low incidence of variability suggests that the simplification observed does not accidentally occur as a consequence of the author's personal style, but it can be ascribed to prefabricated orality (*ibidem*: 452-453). In other words, the low degree of variation across the six movies seems to be related to a diamesic dimension, i.e. cinema being a means of communication with its own requirements that makes it different from other written and oral means such as the need to find a compromise between written and spoken language, between spontaneity and intelligibility (*ibidem*: 453-454). All things considered, Rossi's study was one of the first investigations conducted on over a hundred linguistic features — at different levels (phonetic, morphological, syntactic, lexical, textual and pragmatic) — that allowed to obtain data regarding the specificities of film speech, also proving that the cinematic medium itself has an impact on the language of movies (*ibidem*: 452), making film speech different from any other type of speech (*ibidem*: 459-460), as well as from real-life language. Indeed, “filmspeak” — “i.e. the way people (are made to) speak in films” (Alvarez-Pereyre 2011: 51) — involves more neat and less homogeneous dialogues compared to spontaneous language (Rossi 1999: 458).

As more recently pointed out by Bublitz (2017), fictional dialogue presents “a lower rate of features of verbosity, vagueness, non-fluency, engagement and narration” (*ibidem*: 244) as a consequence of the fact that fictional dialogue not only has to comply with the conversational dimension, but also with the diegetic dimension; in other words, dialogues need to sound real, credible and authentic in order to favour the immersion of the viewer in the narration, while also making the plot advance (Freddi 2011: 158). This is specifically observed by Freddi (2011) in her investigation of phraseology in film speech. Drawing on the Anglophone component of the *Pavia Corpus of Film Dialogue* (PCFD)⁶², the author

⁶² The *Pavia Corpus of Film Dialogue* is a parallel corpus consisting of British and American films and their Italian dubbed versions; further details on this corpus will be provided in §2.6.

identifies the most frequent 4-word clusters in film speech and compares their occurrences in two corpora of spoken language: the spoken component of the *British National Corpus (BNC)*⁶³ and the *Corpus of Contemporary American English (COCA)*⁶⁴. Although some similarities occur, the comparison process highlights important differences, with very few features being as frequent in the PCFD as they are in BNC and COCA. According to the author's findings, two types of clusters commonly occur in film dialogue, namely: *what*-questions, like *What are you doing?* and fixed expressions such as *Thank you very much*, *Nice to meet you* and *How are you doing?* (*ibidem*: 150). The former items are commonly used in dialogue to indicate some kind of conflict (*ibidem*), as well as to elicit missing information necessary to make the plot advance (*ibidem*: 151). The latter elements, instead, typically fulfil conversational-interpersonal functions (*ibidem*: 150). These findings thus provide empirical data regarding the fact that linguistic elements are employed in fictional dialogue based on the two levels on which dialogue operates: the diegetic level, where linguistic elements “have a plot-advancing function and contribute to the representation of conflict” (*ibidem*: 157), and the conversational-interpersonal level, where linguistic elements mimic “the spontaneity of natural conversation as linked to the interpersonal sphere of dialogic interactions” (*ibidem*).

Hence, compared to spontaneous conversation, characters commonly express themselves through short lines usually uttered in a single conversational turn. Moreover, verbal exchanges typically occur in limited settings and normally revolve around a limited range of topics (Quaglio 2009), which are discussed by means of regular linguistic patterns. In other words, the depiction of a selection of prototypical communicative situations calls for the routine use of a set of words and expressions. For this reason, scholars postulate that formulaicity may be considered a characteristic feature of telecinematic dialogue (Quaglio 2009; Bednarek 2010, 2012; Forchini 2012), contributing to the repetitive and highly predictable nature of on-screen conversation (Taylor 2004: 80). According to Freddi (2009), it is possible to talk about a phraseology of filmic speech, i.e. chunks of words that recurrently occur in dialogue (*ibidem*: 102). According to the author, these elements primarily function as pragmatic tools, which help “achieve a move in dialogue” (*ibidem*: 122). Moreover, unlike spontaneous speech, the themes of the conversations follow a linear

⁶³ *British National Corpus (BNC), XML Edition*. 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium.

⁶⁴ Davies M. 2008. *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at <http://corpus.byu.edu/coca/>.

development, allowing the audience to easily keep up with the story (Pavesi 2005: 32-33). Thus, prefabricated spoken interactions are more concise than those occurring in spontaneous speech, exhibiting a higher degree of cohesion and coherence — which also bears evidence of the written origin of the text (Rossi 2011: 26) — and a reduced and stereotyped use of phatic and interactional elements (Pavesi 2005: 34).

On this matter, Ryan & Granville (2020) recently focused on the conversational pragmatic specificities of film speech. The authors reveal some systematic and recurrent ways in which cinematic dialogue remarkably deviates from spontaneous conversation as a consequence of the dramatic and narrative principles that underpin these fictional products. In particular, the authors investigate invitation sequences together with phone call openings and closing sequences in the scripts of twenty films (excluding futuristic, fantasy and period films, as well as slapstick comedies). The analysis shows a lack of pragmatic authenticity with attenuated sequences (*ibidem*: 10). This occurs in both invitations and call-opening/call-closing sequences; in the former case, the ordinary pre-invitations — used to minimise the risks of rejection (*ibidem*: 3) — are not reproduced. In the latter case, phone call openings are short and concise; in like manner, phone calls usually terminate in an abrupt way, with non-verbalized endings (*ibidem*: 10).

The same also emerges from Chepinchikj & Thompson's (2016) investigation, where conversation analysis (CA) is used to observe the verbal and non-verbal (prosodic and paralinguistic) features of film discourse. In particular, the data — obtained from the analysis of three scenes extracted from three Woody Allen's movies (*Husbands and Wives*, 1992; *Melinda and Melinda*, 2004; *You Will Meet a Tall Dark Stranger*, 2010) — shows that film discourse does indeed present some of the verbal and non-verbal features that typically occur in spontaneous conversations such as features of adjacency pairs, gaps and restarts, whereas other features like overlaps and interruptions are rarely observed (*ibidem*: 48-49). Most notably, in the context of CA, the study shows that when film interactions break the conversational rules, it is done to achieve a specific narrative function and meaning that “even though not directly conceived by the participants themselves (the actors playing the film characters)” (*ibidem*: 48), it is purposefully designed by the screenwriter and the director of the film to convey specific insights regarding the characters' interactions and relationships (*ibidem*).

Similarly, Pavesi (2020) uses a corpus-based approach to demonstrate how screenwriters employ specific linguistic elements in film speech to guide the audience's perception of important visual stimuli and control the narrative structure of the cinematic

discourse. More specifically, the author investigates film dialogue through demonstrative pronouns such as *this* and *that*. Drawing on 12 Anglophone films retrieved from the PCFD, demonstratives appear to be extensively used in film speech; in fact, they occur more frequently in films than they do in spontaneous English (*ibidem*: 23)⁶⁵. According to the study, this can be attributed to the fact that demonstratives qualify as verbal resources able to intensify viewers' engagement with the disclosing of events. Deictic features tend to be utilized in film speech to direct viewers' focus on a specific shot or item, as well as on a specific diegetic situation or fact, thus intensifying their engagement with the overall fictional context (*ibidem*: 35).

Hence, fictional dialogue presents some of the key conversational elements of real-life spontaneous speech, but it also has its own linguistic specificities: while some features like deictics are employed in greater amounts than they normally are in spontaneous conversation, other elements are inevitably omitted. As pointed out by Alvarez-Pereyre (2011), in these types of texts "[t]he texture of speech is affected, with a propensity to minimise phenomena such as backchannels (e.g. *uh-huh*), repairs, overlapping, simultaneous speech and unclear words" (*ibidem*: 58). Considering the entertaining nature of telecinematic dialogues — i.e. their dramatic and narrative core — audiovisual texts go through a general process of normalization and standardization to favour clarity. Therefore, the language of films and TV series is said to be less vague than real-life talk, as well as more emotional and emphatic (Baños Piñero & Chaume 2009; Quaglio 2009; Bednarek 2012; Baños Piñero 2024).

With regard to television dialogues, a major benchmark for the investigation of prefabricated orality is Quaglio's (2009) work on the language of the American series *Friends*. Quaglio observes the degree of similarity between television dialogue and spontaneous speech, taking into account various linguistic features typical of impromptu conversation: hedges, vague coordination tags and nouns of vague reference to describe vagueness; intensifiers, stance markers, expletives and emphatic *do* to describe emotional language; expletives, slang terms and vocatives to describe informal language; the use of passives and nominalizations, as well as contractions and demonstrative pronouns. The author concludes that the language of the television series *Friends* shares the core linguistic features of face-to-face interactions (*ibidem*: 68, 139); indeed, similar patterns of linguistic features occur in the two contexts such as contractions, first- and second-person personal

⁶⁵ The data obtained for spontaneous spoken English was retrieved from Bibet *et al.* (1999) and Strauss (2002).

pronouns, and private/mental verbs (*ibidem*: 139-140). Nevertheless, some differences occur: natural conversation usually presents more features of vague language, while the dialogues of *Friends* appear to be more emotional and emphatic, as well as more colloquial than spontaneous conversation (*ibidem*: 139).

Quaglio's results are in line with Bednarek's (2012) findings on keywords and trigrams in television dialogue. Through the investigation of 32 episodes of seven TV series belonging to different genres, Bednarek (2012)⁶⁶ observes the frequency of single words (like *hey* and *man*) and n-grams (like *I told you*) considering their potential role as arrays of informality and emotionality in her corpus of television dialogue (*ibidem*: 37). Results show that the language of television is characterized by a higher degree of emotionality compared to spontaneous conversation. This seems to transcend individual series and genres. Most notably, emotionality appears to be attributable to the dramatic and conflicting nature of television products (*ibidem*: 59). On the other hand, what seems to vary across individual series depending on genre is the level of informality (*ibidem*).

As a matter of fact, extra-linguistic factors can influence the type of orality employed in the text. This has also been pointed out by Valdeón (2009) in the investigation of four dramatic series: *Judging Amy*, *The Practice*, *Teachers* and *Queer as Folk*. The author notices that two elements seem to influence the frequency and the type of orality features occurring in television dialogue: the degree of (in)formality of a scene, on the one hand, and the audience for which a product is intended, on the other (*ibidem*: 220). Most notably, in a legal drama like *The Practice*, which mainly takes place in a formal context, features of orality and colloquiality are less present than they are in series like *Judging Amy*, where the events unravel in a less formal work environment (*ibidem*: 219). As for the type of orality features included in the dialogues, the selection is also dependent on the audience for which a product is intended, i.e. in television series created for an adult audience, like *Teachers* and *Queer as Folk*, orality is predominantly conveyed through taboo words and swearwords (*ibidem*: 218).

As will be examined henceforward, the peculiarities discussed here have a direct impact on the translation of audiovisual texts for an international audience.

⁶⁶ Bednarek (2012) studies thirty-two episodes extracted from seven television series representing five different genres, namely *NCIS* (crime), *Supernatural*, *Lost* (mystery), *House* (medical drama), *How I Met Your Mother*, *My Name is Earl* (comedy), and *Desperate Housewives* (drama) (*ibidem*: 38-39).

2.5. Dubbing Prefabricated Orality in Audiovisual Texts

The distinctive traits of audiovisual texts, i.e. the way dialogue originates, its linguistic specificities and the pursuit of spontaneous-sounding interlocutions, have a direct impact on their translation process since not only the original scriptwriters, but also the translators and adapters of the audiovisual product must strike a balance between the written and spoken mode. Translated dialogues must indeed sound natural and improvised in the target language, that is, credible to the ears of the target audience (Baños Piñero & Chaume 2009; Romero-Fresco 2009a; Chaume 2012: 81; Bruti 2019; Baños Piñero 2024: 43).

Naturally, since translation is said to oscillate between two poles, namely “its adequacy to the source text and its acceptability within the target culture” (Marzà & Chaume 2009: 33), the oral features able to convey naturalness and spontaneity in the translated text will differ depending on both language and cultural constraints. In other words, each language has its own linguistic instruments to simulate spontaneous conversation based on specific cultural norms, which determine the linguistic and iconic adequacy of an audiovisual text regarding spokenness (Chaume 2012: 81). As pointed out by Baños Piñero & Chaume (2009), “translators must be aware that the original script they receive from the translation agency or dubbing studio [...] has been written to be spoken in such a way as to appear spontaneous”, taking into account the complex semiotic nature of the specific type of texts observed here, i.e. audiovisual texts (Chaume 2012: 82). As reported by Spiteri Miggiani (2019: 234), “[t]he search for natural-sounding dialogue can also be considered as a non-technical constraint [...] when this is set as an intention by the dialogue writer”.

Since “[a]n audiovisual text is a semiotic construct woven by a series of signifying codes that operate simultaneously to produce meaning” (Chaume 2012: 100), when translating the original script, these multiple signifying codes must be taken into account (Baños Piñero & Chaume 2009; Chaume 2012: 82). In other words, when translating films and/or TV series, attention must be paid to the various codes involved in the audiovisual product, namely both the acoustic channel — including the verbal and linguistic component, as well as sounds and music — and the visual channel — conveying images and kinetic information (see Díaz Cintas 2009: 9). The meanings and information conveyed in the multimodal text through these codes have to be examined concurrently and collectively by the translator; hence, considering both linguistic and non-linguistic signs (Chaume 2012: 100). As a matter of fact, this is the only way to investigate the meaning-making resources

of the multimodal text, i.e. to fully accept and understand translational choices and their consequent impact (Pavesi 2019a: 157). To summarize:

AVT is always constrained by the presence of the original production, which lives on semiotically through images (and sound) in the adoptive culture. Dialogue exchanges do not just happen in a vacuum but they always take place in a given context, which, in the case of AVT, is a concrete situation in time, captured and frozen by the camera. The fact that target viewers have the same access as source viewers to the visuals of the programme [...] has vast implications for the way the translation can be carried out (Díaz Cintas 2009: 9).

It is important to remark that the present discussion focuses on the adaptations made for audiences who can rely on both the visual and the acoustic code during the screening. Consequently, the present discussion does not take into account the adaptations made for audiences with sensory and linguistic barriers such as people with hearing and/or visual impairments. These adaptations must be carried out on their own terms, considering that translators must replace the source text with an equivalent text in the target language while also making this translation accessible to the needs of the audience⁶⁷. The process through which accessibility occurs in AVT is beyond the scope of this investigation; therefore, it should be noted that the discussions addressed here are limited to the adaptations made through dubbing for Italian hearing viewers.

2.5.1. The Dubbing Process: Technical and Non-Technical Constraints

The adaptation of an audiovisual text is a complex process during which translators must cater for the choices made by the original scriptwriters and actors, while also paying attention to factors like genre conventions, industry guidelines or restrictions imposed by the commissioner (Baños Piñero 2024: 64) and, concomitantly, comply with a number of technical constraints imposed by the medium itself.

As previously noted, this study focuses on cinematic products adapted for an Italian audience through dubbing. Dubbing is a translation mode that “involves replacing the original soundtrack containing the actors’ dialogue with a target language recording that

⁶⁷ For a detailed discussion on accessibility in AVT see, among others, Romero-Fresco (2013, 2019); Greco (2019); Jiang *et al.* (2023). For studies on audio description see, among others, Salway (2007); Di Giovanni (2018); Perego (2019); Chmiel & Mazur (2021); Mazur (2023). For studies on SDH subtitling, see Neves (2009, 2019); McIntyre & Lugea (2015); Romero-Fresco (2015); Taylor (2015); Tamayo & Chaume (2017); Bruti & Zanotti (2017, 2018).

reproduces the original message” (Díaz Cintas 2009: 4-5). When adapting audiovisual products through this translation mode, the target-language dialogues replacing the original ones must respect some restrictions and requirements imposed by the original soundtrack; that is, they must respect synchronization (Baños Piñero & Chaume 2009; Chaume 2004c, 2012, 2020).

In dubbing, three main types of synchronization can be established: lip synchrony, kinesic synchrony and isochrony. Lip synchrony consists of adapting the text to the articulatory movements of the lips of the speaker; kinesic synchrony consists of adapting the text to the body movements of the actors on the screen; isochrony, finally, consists of making the duration of the translated statement coincide with the duration of the original one (Chaume 2004c: 43-44, 2012: 67-69, 2020: 111-113)⁶⁸. The level of adherence between the words uttered and the images on the screen is essential in determining the success or failure of audiovisual products (Chaume 2012: 14), especially in countries like Italy, where the public is used to an excellent tradition of dubbed foreign films (Pavesi 2005: 16). In dubbing, it is central to ensure that “the target language sounds and the actors’ lip movements are synchronised, in such a way that target viewers are led to believe that the actors on screen are actually speaking their language” (Díaz Cintas 2009: 5)⁶⁹.

Producing a text whose enunciation satisfies the synch criteria discussed above is one of the key features in creating high-quality dubbed products; at the same time, credibility and naturalness are also very important, insomuch that, in some cases, it is preferable to prioritize realistic dialogue over perfect synchronization. Most notably, audiences are more likely to tolerate dialogue that is natural but not perfectly synchronized compared to a perfectly synchronized text that sounds artificial (Chaume 2012: 86). In this regard, Chaume (*ibidem*: 83) reports that “a major benchmark for good dubbing quality is that the target language sounds realistic, credible, and plausible; i.e. that it does not distract us from the storyline”. This is achieved by the translator staying true to the prevailing norms existing in the target culture regarding the oral register and, consequently, credible prefabricated oral register. If these norms are not respected, the dubbed performance will sound forced and

⁶⁸ These three types of synchronization are the most popular and well-established forms of synchrony for dubbing; in AVT, however, various classifications have been proposed by scholars over the years. See, among others, Whitman-Linsen (1992), Agost (1999) and Spiteri Miggiani (2021).

⁶⁹ As for the dubbing of original English movies, Maffia & Pellegrino (2016) have highlighted some of the specific rhythmic characteristics and articulation rates observable in dubbing, along with some of the strategies used to guarantee synchronization between the original and the dubbed adaptation made for various European languages such as Italian, French, Spanish, German and Portuguese. For instance, in all languages, with the exception of Portuguese, there is a tendency to reduce the original silent pauses, which are indeed fewer and shorter in dubbing (*ibidem*: 118).

artificial to the audience, impacting the overall success of the cinematic or television product in the international market. As phrased by Whitman-Linsen (1992: 118), “[a]rtificiality is one of the main faults pilloried in denouncements of dubbed versions: the audience can hear that it is not an original”. Hence, dubbing primarily involves the domestication of a foreign product so that viewers do not recognize, or simply forget, that what they are watching on the screen is a translation (Chaume 2020: 113).

Before diving into the language of dubbing, looking at the specific features that guarantee the aforementioned sense of realism, it is important to stress that dubbing itself is a complex process in which various people can leave their authorial mark. Indeed, there is no single person responsible for the final product; on the contrary, there are a number of agents who partake such as translators, dubbing directors, voice talents (i.e. dubbers), sync assistants, dialogue writers and so on (Chaume 2016). More specifically, the dubbing process is carried out in the following way (Chiaro 2008, 2009; Chaume 2012: 29-30; Zanotti 2014a: 111; Bosseaux 2018): once the work has been commissioned to a dubbing company, the translator produces an initial literal translation of the script, which is then edited by the dialogue writer (sometimes the dialogue writer coincides with the translator, i.e. they are the same person). At this stage, a series of changes are made to the translation from a stylistic point of view in order to obtain a discourse that is credible and truthful to the source text, linguistically and grammatically correct, and which fits the required prosody, while also matching the actor’s facial expressions, gestures, mouth articulation and lip-movements (Spiteri Miggiani 2019: 12). Subsequently, dubbing actors record the new soundtrack under the guidance of the dubbing director, who may intervene on the text if the dubbing script is perceived as inappropriate to the on-screen context or to the characterization of the characters; at the same time, dubbing actors may also modify the script in order to adapt it to their own acting style (Chiaro 2008: 247, 2009: 145-146; Chaume 2012: 36-37; Baños Piñero 2014). Finally, during the post-production process, sound engineers reassemble the audio tracks and edit them, adding sound effects (Chaume 2012: 37)⁷⁰.

According to Chaume (2016), one way in which the various agents involved in the process can maximise the success of audiovisual products is to follow specific conventions that help the audience recognize the product, its genre, and watch it in a particular way. These conventions include: conforming with the synchronization norms (especially in close-

⁷⁰ For a more detailed description of the dubbing process and the agents involved, see: Whitman-Linsen (1992); Paolinelli & Di Fortunato (2005); Chiaro (2008, 2009); Chaume (2012: 29-30); Zanotti (2014: 111); Bosseaux (2018); Spiteri Miggiani (2019); Minutella (2021a: 19-67).

up shots and in detailed lip shots), creating credible and realistic dialogue lines, ensuring coherence between what is seen on the screen and what is heard, and providing the target audience with a translation faithful to the source language for what concerns content, form, function and effect. In addition, clear sound quality and good acting are also mentioned (Chaume 2012: 14-20, 2020: 110-111). The former includes sound mixing, mastering and the editing process. Most notably, the editing process is particularly important for the final product since the way in which the dubbed recorded voices are mixed into the original soundtrack can have a massive impact on the credibility of the overall dubbed version (Spiteri Miggiani 2019: 7, 33). Good acting specifically refers to the voice talent, i.e. the dubber's performance and dramatization of the dialogue, which should sound neither faked (overacted) nor monotonous (underacted) (Chaume 2016). It is the dubbing director who selects the voice talent, i.e. the dubbing actor whose voice will replace the one of the original actor (Spiteri Miggiani 2019: 9). This is a critical aspect since vocal performances will impact the target audience's perception of characters, their conversations, and overall, the appraisal of the final product (see Bosseaux 2015).

Sánchez-Mompeán (2020a) carried out a groundbreaking investigation on the naturalness of prosodic and paralinguistic features in Spanish dubbing. Since voice is central in defining the identity of characters on-screen, it is of pivotal importance to identify voice talents whose voice and vocal performance correspond to what viewers expect to hear from actors on-screen and what they usually associate with a character's identity and personality (Chaume 2012: 69). Performance and intonation, i.e. the way in which actors perform their lines, influence the realism and credibility levels of the dubbed dialogues. However, much depends on the way the target language script has been written and translated (Baños Piñero & Chaume 2009; Spiteri Miggiani 2019: 34; Sánchez-Mompeán 2020a: 204). The aim is to submit a "natural-sounding dialogue, so that the text presented to dubbing actors will instigate their best performances" (Spiteri Miggiani 2019: 235).

Apart from the technical constraints that need to be observed during dialogue writing, Spiteri Miggiani (2019: 30-31) also recognizes some non-technical constraints, among which are language issues, translation and blacklisted expressions due to legal and commercial restrictions (*ibidem*: 30-31). In particular, the author claims that these language constraints "also imply respecting established familiar, and viewer-accepted dubbing tendencies" (*ibidem*: 31), which means conforming the language of dubbing to established standardization, naturalization and explicitation norms (*ibidem*). Further details on the typicalities and norms established for the language of dubbing are provided hereunder.

2.5.2. The Language of Dubbing: Dubbese and Dubbitis

The language of dubbing has been referred to as “dubbese” (Myers 1973; Pavesi 1996), that is “a culture-specific linguistic and syntactic model for dubbed texts” (Chaume 2012: 87); “a third norm” (Pavesi 1996), describable “as a self-standing, culture-specific variety of language” (Pavesi 2019a: 158), which is close but not identical to spontaneous spoken discourse. As also reported by Bruti (2019: 205), “dubbese adheres neither to the source nor to the target language, but to a third language that strengthens formulaic language and translational clichés”.

Various scholars have investigated the language of dubbing with regard to translation norms and strategies. Romero-Fresco (2009b), for instance, carried out a description of the language typically found in dubbing, focusing on both the fictional and the translational dimension of Spanish dubbese. Chaume (2004b), instead, proposed a model structured around the four levels of the language, i.e. the prosodic, morphological, syntactic and lexical level. The author notices that while a certain degree of standardization can be noticed at all levels, this tendency is considerably substantial in dubbing at the morphological and at the phonetic level. This is no surprise since, generally speaking, in dubbing there is a well-established propensity to make the verbal code more standard, reducing interpersonal features and relying on the partial and selective replication of lexical and syntactical features of spoken language (see Baños Piñero & Chaume 2009).

As for Italian dubbing, a tendency to normalize and standardize the language of dubbed movies, particularly with regard to phonetic aspects, as well as the lexical and morphosyntactic level of the language, has been highlighted by Pavesi (2005: 45), who defined the group of linguistic features systematically selected in dubbing to evoke spokenness as “privileged carriers of orality” (Pavesi 2008). With regard to Italian dubbing, orality is mainly conveyed using syntactic structures, marked word order, personal pronouns and weak connectors (*ibidem*). Through the investigation of these features, Pavesi (2008) carried out a quantitative and qualitative study where the Italian dubbing of original English films is compared with the language of spontaneous Italian. In particular, the author found that “dubbed Italian appears to result from the interaction of target language norms (which play the most significant role), source language interference (selectively and to a restricted extent), and formulaic language, a feature which has been recognized as typical of film language” (*ibidem*: 94). The language of audiovisual translation is therefore able “to

reproduce some of the major patterns of orality of the target language” (*ibidem*); nevertheless, it also amplifies the use of translation routines, that is, “recurrent solutions to translation problems which tend to become overextended” (*ibidem*).

To some degree, formulaicity in dubbing reflects the formulaicity of the original source language. However, certain dubbese features are so imbedded in this specific type of language that they appear in dubbing even when they could be avoided, that is, when the dubbing constraints would not necessarily require the selection of dubbese-specific linguistic items. This is highlighted by Romero-Fresco (2012) in Spanish dubbing with reference to markers of agreement used as transition markers. More precisely, Romero-Fresco (2012) carried out a quantitative and qualitative analysis of discourse markers having the function of transition markers in three corpora: a parallel corpus containing the transcription of the American TV series *Friends* and their Spanish dubbed version; a corpus containing the language of Spanish scripted texts, retrieved through the transcription of the original Spanish sitcom *Siete vidas*; and, finally, a corpus of spontaneous spoken Spanish (CREA, compiled by the Royal Spanish Academy) (*ibidem*: 184-185). Results show that the solutions found in dubbing differ from the way transition markers are usually employed in spontaneous Spanish; the translator, in fact, tends to opt for unnatural features, i.e. renderings that are far from what occurs in spontaneous conversation (*ibidem*: 199-200). Interestingly, in the texts analysed, these results cannot be attributed to dubbing constraints, since unnatural features are selected even in those cases in which the translator had the freedom to opt for other renderings (*ibidem*: 200). According to the author, a possible explanation lies in what can be defined as “suspension of linguistic disbelief” (*ibidem*: 201).

As a matter of fact, even in dubbing there is a tacit agreement between the creators of the dubbed products and the target viewers, who seem to unconsciously “absorb and accept dubbese as a natural-sounding language” (Spiteri Miggiani 2019: 35). In the context of prefabricated orality, there are features of the language reproduced in dubbing that do not sound natural; nevertheless, by watching dubbed movies and TV series, viewers gradually “learn the dubbing conventions, get used to them and avoid questioning them in order to enjoy the viewing” (Romero-Fresco 2009a: 68)⁷¹. Hence, the audience is likely to suspend

⁷¹ In this respect, important data regarding viewers’ preferences and strategies to immerse themselves in the story are provided by the experiments carried out by Di Giovanni & Romero-Fresco (2019) and Romero-Fresco (2020). Di Giovanni & Romero-Fresco (2019), for instance, used eye-tracking technology to compare the eye movements of viewers watching the original English version and the Italian dubbed version of clips extracted from the film *The Grand Budapest Hotel* (Anderson, 2014). The findings obtained from the analysis allowed viewers to confirm the existence of a so-called dubbing effect, “a mechanism whereby dubbing viewers avoid

its disbelief even if the target text does not sound as natural as expected (Romero-Fresco 2009a: 68; Pérez-González 2014: 120). In this respect, as pointed out by Schettino *et al.* (2022: 197), “naturalness is not an inherent property of speech, but is specified with reference to the communicative context of application”. This is in line with Romero-Fresco’s (2009a: 51) definition of naturalness in dubbing as the “nativelike selection of expression in a given context”. Drawing on Warren’s (2005) definition of naturalness — according to which it involves “knowing which particular combinations are conventional in a language community although other combinations are conceivable” (*ibidem*: 40) — Romero-Fresco (2007) points out that in the context of audiovisual products, rather than focusing on what is correct or grammatically possible, naturalness and idiomaticity pertain to what is conventional among the many grammatically possible options in a specific situation (*ibidem*: 51). As a result, in the context of AVT, it may be possible for the audience to accept certain linguistic features reproduced in cinematic discourse as natural and to recognize dialogues as realistic even though they are actually not (Baños Piñero 2024: 224). Thus, the orality features included in the translation of an audiovisual text may be far from the orality features native speakers would normally select in the same conversational context depicted on the screen. As reported by Chaume (2004d: 854): “This genre [i.e. AVT] is already accepted to be less coherent than other genres, the audience accepts deficiencies in cohesion and coherence in audiovisual translations as part of their inherent characteristics”.

Interestingly, similar considerations have also been made with regard to the prosodic and the phonetic level of dubbing since dialogue must sound credible and realistic not only from a linguistic point of view, but also from a prosodic and acting perspective (Chaume 2012: 19). Generally speaking, very few features of real-life oral speech are reproduced in dubbing since voice actors usually present very clear articulation and intonation (*ibidem*: 87), far from the way speech is articulated in everyday life. In this regard, Sánchez-Mompeán (2020a, 2020b) has investigated the naturalness of dubbed speech by focusing on the prosodic level. More specifically, the author examined the tonal patterns of Spanish dubbing

looking at the characters’ mouths and focus mainly on their eyes, in contrast with the natural and idiosyncratic way in which original viewers normally watch reality and films” (Di Giovanni & Romero-Fresco 2019: 142). In other words, viewers adopt some strategies that help them ignore the inevitable lack of lip synchrony and favour their suspension of disbelief (*ibidem*). This has also been corroborated by Romero-Fresco’s (2020) investigation, where he compares the eye movements of a group of native Spanish speakers with the eye movements of two other groups of viewers, namely a group of native English speakers watching the same clip in English and a group of Spanish native speakers watching an original and comparable clip in Spanish. Even in this case, findings suggest that viewers do perform some unconscious eye movement strategies, which allow them to be transported into the diegetic world.

in order to ascertain whether intonation can be considered as one of the characteristics of dubbed prefabricated orality (Sánchez-Mompeán 2020b: 285). Indeed, also in dubbing, voice talents must be able to conceal the fact that they are reading a script (*ibidem*: 286; see also Sánchez-Mompeán 2023b). Intonation plays a pivotal role in recreating the natural sound of impromptu conversation and in enabling the correct interpretation of meaning since, as pointed out by Sánchez-Mompeán (2020b), “the voice becomes the only available instrument for conveying the vocal nuances and the pragmatic load behind the original actors’ words” (*ibidem*: 287). In her investigation of the Spanish dubbed adaptation of the American TV series *How I met your mother*, Sánchez-Mompeán (2020a) notices that “dubbed prosody is characterised by specific trends that might curtail natural dialogues, thus affecting both the quality of the dubbed text and the viewers’ perception” (*ibidem*: 143). Dubbing actors usually resort to a combination of opposite traits: monotonous and artificial features resulting from reading a script, on the one hand, and overacted and over-involved prosodic delivery, on the other. This combination creates a specific melody, characteristic of dubbed products, defined by Sánchez-Mompeán as “dubbitis” (*ibidem*: 148). In other words, “dubbing intonation seems to be characterized by a combination of spontaneous and non-spontaneous orality” (Sánchez-Mompeán 2020b: 296), creating some melody patterns that are recognized by the general public as typical of dubbed dialogue.

As reported by Spiteri Miggiani (2019: 37), “just as dubbese is unconsciously accepted by the audience, so is intonation, as long as it falls within credible parameters, beyond which, one would be left with a parodic effect”. According to Sánchez-Mompeán (2020b), the audience seems to recognize the lack of naturalness in prosodic patterns, but they accept it as legitimate within the context of dubbing (*ibidem*: 296). The author defines this phenomenon as the “suspension of prosodic disbelief” (Sánchez-Mompeán 2020a: 194); in short, viewers accept the fact that on-screen characters may not sound as people do in real-life situations. Rather, protagonists on the screen sound like characters commonly do sound in dubbed products (Sánchez-Mompeán 2020b: 296). Nevertheless, more attention should be paid to intonation and tonal patterns so as to obtain more natural-sounding performances capable of reproducing the attitude and the pragmatic nuances of the original version (see Sánchez-Mompeán 2019, 2020a: 229). This is particularly relevant in relation to the fact that acoustic features, like intonation and prosody, do not have a universal value. The meanings conveyed through voice and prosody can vary across languages and cultures (Bosseaux 2015: 73), carrying over different connotations and associations. Hence, when

adapting movies or TV series through dubbing, the target language may require a different prosody.

As already noted by Zabalbeascoa (1997: 339), next to the verbal component of the original soundtrack, more attentiveness should be paid to non-verbal elements in AVT. In fact, these features are used in dialogues to transmit information that can complement the meanings conveyed verbally. In dubbing, however, the information transmitted through the non-verbal component runs the risk of being altered or lost since it cannot necessarily be retrieved from verbal features. It is interesting to notice that even when it comes to factors such as intonation, loudness, rhythm, tempo and tension, dubbing seems to resort to recurrent patterns, which define dubbed dialogue as a characteristic style of speech associated with repeated deliveries and typifying features. Sánchez-Mompeán (2023b) notices this in relation to Spanish dubbing; similar considerations, however, have also been made with regard to the overall dubbing industry. Ranzato & Zabalbeascoa (2022: 17), for instance, allocate the fixed prosodic patterns typical of dubbing to the strong and well-established norms settled within the industry.

Therefore, routinization in dubbing concerns both verbal and non-verbal features. As for the verbal features, Pavesi (2008, 2016) states that some of the typical characteristics of dubbese actually enhance both translators' and viewers' processing of the dialogue; in addition, "they trigger target audiences' feelings of shared identity and belonging to the same lingua-cultural community" (Pavesi 2016: 101). Translated texts, in fact, exhibit a high degree of formulaicity, resulting from the repetitive and conventional linguistic patterns originally found in audiovisual texts, which are in turn rendered into the target language through stock equivalent solutions (i.e. translation routines). These translation routines generally include not only individual lexemes, inserts and formulaic sequences, but also semantic, structural and pragmatic calques, as well as loan creations deriving from the original source text (*ibidem*: 102).

Over the last few decades, researchers have focused on the typicality of audiovisual dialogues and the strategies adopted in AVT, focusing on various linguistic phenomena such as discourse markers (Chaume 2004d), interjections (Bruti & Pavesi 2008), question-tags (Chiaro 2000; Bonsignori 2009a), dislocations and cleft sentences (Pavesi 2005), but also personal pronouns (Pavesi 2009a), intensifiers (Baños Piñero 2013), phrasal verbs (Valentini 2013) and general extenders (Zanotti 2014b, 2014c). Most notably, besides translation-driven investigations carried out through single case studies, the employment of larger, multi-purposed and self-standing databases has proven to be of cardinal importance. AVT

research has indeed made significant steps forward thanks to the collection of large databases containing the transcriptions of telecinematic dialogues to be studied empirically through corpus-based methodologies (Forchini 2013: 505), i.e. researching into the recurring and distinctive patterns of telecinematic dialogue and addressing translational issues, shifts and strategies (Pavesi 2019b, 2022).

2.6. Corpus Linguistics Approaches to AVT: Systematic Linguistic and Translational Behaviours

Since the 1990s (Baker 1993, 1996), Corpus-based Translation Studies (CBTS) has been employed as an approach to research on translation, bringing together the theoretical and methodological frameworks of Descriptive Translation Studies (DTS) and Corpus Linguistics (CL). This approach has enabled scholars to validate the systematic linguistic and translational behaviour of AVT registers through empirical data (Pavesi 2019b: 315). As a result, a vast bulk of studies has been conducted through a corpus-based approach in the field of AVT (see Baños Piñero *et al.* 2013; Pavesi 2019b, 2022; Bruti 2020 for a review)⁷². To this end, translation-driven corpora⁷³ have been compiled through the collection of dialogue transcriptions, purposefully designed to investigate the translation of various linguistic phenomena in both films and television series (see also Bednarek 2015). These databases allow for contrastive analyses and generalizations, detecting “patterns that would be difficult to identify through manual analysis” (Baños Piñero *et al.* 2013: 483-484), like the simplification, explicitation and normalization of translated texts (*ibidem*: 484). Among these databases, it is possible to find, most notably, the *Forlì Corpus of Screen Translation* (Forlìxt 1) (see Valentini 2008; Heiss & Soffritti 2008; Valentini & Linardi

⁷² Considering the interests of this research, i.e. dubbing, the present section mainly focuses on reviewing corpus-based translation studies in the context of AVT carried out on the translation mode of dubbing. Nonetheless, it is important to draw attention to the fact that significant research has also been conducted by means of this methodological approach in relation to other translation modes such as subtitling, fansubbing and audio description; see Baños Piñero *et al.* (2013) and Bruti (2020) for a review.

⁷³ Three types of corpora are usually employed to research AVT: parallel, monolingual and bilingual comparable corpora. Parallel corpora consist of a set of source language texts accompanied by their translation into a given target language. Monolingual comparable corpora include two sets of texts in the same language, where the first set contains translated texts and the second set contains comparable non-translated texts (domestic texts). Bilingual comparable corpora, instead, contain non-translated texts belonging to different languages but sharing the same textual and stylistic features, as well as discursive and pragmatic functions (Pavesi 2019b: 316; Bruti 2020: 383). Hence, while parallel corpora accommodate the scrutiny of translation strategies, comparable corpora include texts that are similar in genre or text type. In addition, reference corpora can be added to the list, that is, corpora used for comparison purposes (Baños Piñero *et al.* 2013: 484).

2009), as well as the *Pavia Corpus of Film Dialogue* (PCFD) (see Freddi & Pavesi 2009; Freddi 2013; Pavesi 2014; Zago 2019)⁷⁴.

Forlìx 1 is an electronic database of films and transcripts of film dialogues/subtitles, containing both audiovisual and textual data, developed for the comparison of original and dubbed versions of films in various European languages such as Italian, German and French (Valentini 2008). The extent of usefulness of this multimedia corpus is exemplified by Valentini (2013) in the investigation of linguistic traits typically found in spontaneous and colloquial language, namely phrasal verbs as fulfilling adverbial, interjectional and phrasal functions. The study is conducted through a corpus of three types of texts: original Italian films, original French films and their translation from French into dubbed Italian. By comparing these three types of texts, it emerges that the verb-particle constructions under investigation are more frequent in original films, both French and Italian, than they are in dubbing (*ibidem*: 548). This may be attributed either to the influence of the source language on the translation process or to the language restrictions underlying the translation of texts from French into Italian (*ibidem*: 549). Moreover, although phrasal verbs are reproduced in Italian dubbing, they are often indistinctively used in the target text to translate various French registers, namely colloquial registers, foul language, and police jargon. This seems to confirm the tendency observed in dubbing to neutralize social varieties (*ibidem*: 552-554; see also Romero-Fresco 2009b). In fact, “if on the one hand the use of phrasal verbs in dubbed dialogues contributes to draw dubbed language closer to informal spontaneous conversation, on the other hand, it favors the levelling out of registers, since different social varieties in the source text [...] are translated by making use of the same lexical resource” (Valentini 2013: 554). As for translation strategies, the original French verb-particle constructions are usually rendered through the consistent use of lexical constructions, identifiable as translation routines. This is the case, for instance, of the verb-particle construction *andare via*, used to translate a variety of features occurring in the original French version; among others, the verbs *partir* and *filer*, as well as the construction *y aller* (*ibidem*). Moreover, translation also occurs through the use of calques, which reproduce in Italian the syntactic structure of the original French phrasal verb, as is the case of *s’enrouler autour* translated as *arrotolarsi intorno* (*ibidem*). As for the multimedia component of the

⁷⁴ More information about the characteristics and the state of the corpus, updated to its most recent version (2023), can be retrieved online at the following link: https://linguisticapavia.cdl.unipv.it/sites/cdl69/files/2023-10/Testi%20Pavia%20Corpus%20of%20Film%20Dialogue_sett_23.pdf.

corpus, Valentini investigates whether any correlation can be found between the abovementioned translation strategies and visual semiotic constraints. The author highlights that there is a correlation between translation strategies and non-verbal features. More specifically, the linguistic choices made in translation in relation to the use of verb-particle constructions are rarely influenced by factors like prosody and facial expressions; what seems to influence the translation in dubbing is body language (*ibidem*: 555). In this regard, the author claims that in both original and dubbed dialogues, the use of phrasal verbs “could be regarded as a residual element of the writing process” (*ibidem*: 557) in which these constructions are used to describe the direction of a gesture of the actors on screen, thus bringing film dialogues closer to writing than to orality.

The PCFD consists of a parallel and comparable spoken corpus, defined by Bruti (2020: 386) as “one of the most fully fledged corpora of AVT”. First developed to explore the language of films in original and dubbed Anglophone products, it then turned into a relational database, allowing for more in-depth translational analyses of both single items and discourse sequences. The corpus also includes a comparable component, that is to say, original Italian productions, which make it possible to compare dubbed Italian dialogues with original Italian dialogues and to reveal what “typifies dubbese as a third language” (Pavesi 2014: 43). Focusing on the English-Italian language pair, the PCFD allowed to carry out research on a wide range of topics: Pavesi, for instance, observed formulaicity in original and dubbed audiovisual products, as well as the translation norms of Italian dubbing with regard to various phenomena such as personal pronouns (2009), demonstratives (2013) and conversational routines (2016). The results obtained from these investigations allowed to delineate some of the defining features of Italian dubbing such as, for instance, the general reduction that occurs in dubbed Italian of certain features typical of impromptu conversation compared to the occurrences noted both in the original source language dialogues and in spontaneous Italian⁷⁵.

Pavesi highlights this tendency in relation to both third-person singular subject pronouns (Pavesi 2009a: 128) and demonstratives (Pavesi 2013: 108, 111). These results seem to suggest that Italian dubbing is target oriented (Pavesi 2009a: 129) and in line with the prefabricated language of Italian products (Pavesi 2013: 110). More specifically, in the investigation of third-person singular subject pronouns, Pavesi’s (2009a) findings suggest

⁷⁵ The comparisons with spontaneous language are usually carried out using the corpus of spoken Italian *Lessico di frequenza dell’Italiano Parlato* (LIP) (De Mauro *et al.* 1993).

that Italian dubbing does maintain some of the tendencies found in spontaneous conversation; in fact, the less frequent occurrences of third-person pronouns demonstrate the lack of interference of the source language on the adaptation. If English did influence dubbing, the linguistic items observed by Pavesi would have resulted way more frequent in Italian adaptations (*ibidem*: 128-129). For this reason, Italian dubbing is said to be oriented toward the target linguistic system, because it reflects some of the tendencies found in spontaneous Italian (see also Pavesi 2005). Along these lines, Italian dubbing also shows some adherence to original Italian film speech. Pavesi (2013) observes this through the analysis of demonstratives, confirming the fact that “the language of audiovisual dialogue dubbed from English into Italian tends to gravitate towards the simulated orality of original Italian productions” (*ibidem*: 110).

As for the translational approaches, despite the general alignment between dubbed Italian and spontaneous Italian discussed so far, there are cases in which the translation presents marked solutions, reflecting the influence of the source English language. With regard to the translation of pronouns, for instance, Pavesi (2009a) underlines cases in which, under the influence of the source text, explicit subject pronouns are used in dubbing even though the employment of a null pronoun was possible, if not even more suitable in the situation depicted on the screen (*ibidem*: 132). At the same time, the considerable transformations observed in dubbing can also be attributed to the syntactic and pragmatic differences between the source and the target language, as is evident, for instance, when dubbing demonstratives (Pavesi 2013: 115). In this case, translators usually resort to translation strategies such as total omission, substitution with a different type of pronoun or a non-deictic lexical element, as well as explicitation using a noun phrase (*ibidem*: 111). Hence, it is possible to draw the conclusion that while English dialogues provide explicit linguistic cues (i.e. demonstratives) to understand dialogue and the situational context, Italian dubbing resorts to weaker reference forms such as clitic pronouns, thus relying “more on the interactant’s mutual understanding and the iconicity of the multimodal product” (*ibidem*: 128), where iconicity is defined as “an intrinsic quality of films, which — compared to other forms of narration — primarily show rather than tell” (*ibidem*).

To sum up, Pavesi underlines how the combination of target- and source-oriented perspectives in corpus-based translation studies can contribute to gaining insights on the translation of naturalness in AVT. The mismatch existing between English and Italian, in fact, represents a problem in translation, often solved through the exploitation of routine translations. This is also pointed out with regard to the adaptation of certain conversational

routines. More specifically, Pavesi (2016) investigates English demonstrative cleft formulas — like *that's what I mean, this is how it started* and *that's where I'll go* — in 24 Anglophone films retrieved from the PCFD. As pointed out by the author, “demonstrative clefts are typified by a degree of fixedness as they come in a recurrent pattern which allows only limited variation” (*ibidem*: 105) and fulfil a range of discourse-structuring functions; namely, they are used to provide an explanation, to conclude a discourse, as well as to point to, hold on and highlight relevant items (*ibidem*: 105-106). The fixedness these features present in spontaneous conversations is also faithfully reproduced in English dialogues, where they occur with a similar frequency, a similar pattern — namely, *that* followed by a present tense verb — and also perform equivalent discourse-organizational functions (*ibidem*: 106-107). As for Italian dubbing, these elements pose a problem in translation considering their rare occurrence compared to the English language (*ibidem*: 108). As a result, in dubbing, these elements are either omitted or translated through demonstrative *it*-clefts, that is, “clefts whose clefted constituent is a demonstrative pronoun”, such as *è questo che* used to translate *that's what's* (*ibidem*: 109). Considering their structural and functional similarity to the original constructions, the use of *it*-clefts can be regarded as a routinised translation solution (*ibidem*: 111-113), thus demonstrating the recurrence of translational behaviours involving a restricted group of options. To phrase it differently, in dubbing, “translators identify formulas in the target language that come close to the original formulas and employ them systematically to represent orality on screen” (*ibidem*: 115).

To summarize, by reviewing Valentini's (2013) and Pavesi's (2009a, 2013, 2016) research, it was possible to highlight some of the systematic linguistic and translational behaviours identified for Italian dubbing through empirical data, such as the following:

- The influence of the source language.
- The target language orientation of Italian dubbing.
- The tendency to neutralize the translated text.
- The tendency to employ translation routines and calques.

On these premises, Pavesi (2018) proposed a model to investigate the features of dubbing based on three dimensions: “naturalness and register specificity, target language orientation and source language interference, and routinisation and creativity” (*ibidem*: 101). A more exhaustive outline of the strategies adopted in dubbing to translate oral features is provided in the following section, observing the rendering of various phenomena typical of impromptu conversation.

2.7. The Translation of Orality Features in Dubbing

The representation of spoken discourse in audiovisual products translated for an international audience has been a prolific field of research, aimed at investigating regularities both within and across languages (Freddi & Pavesi 2009; Baños Piñero *et al.* 2013; Bruti 2019; Pavesi 2019a). Empirical studies on AVT have described the language of dubbing as being characterized by a sense of orality usually transmitted by the means of a selection of spoken language markers, evoking naturalness and colloquiality (Pavesi 2019a: 158). Accordingly, in dubbing, spontaneity is conveyed by virtue of features like discourse markers (Chaume 2004d; Romero-Fresco 2007) and interjections (Bruti & Pavesi 2008), intensifiers (Romero-Fresco 2009a; Baños Piñero 2013), vocatives and familiarizers (Forchini 2013), swear words and taboo words (Valdeón 2015, 2020), but also by reproducing the syntactic features of spoken language, for instance, through dislocations and cleft sentences (Pavesi 2005, 2008), questions (Ghia 2014, 2019), as well as through conversational routines (Bonsignori *et al.* 2012; Bonsignori & Bruti 2014).

2.7.1. The Omission of Conversational, Informal and Pragmatic Markers

Authors have noted a progressive increase of spoken informal elements in film speech, witnessing a wider use of linguistic features traditionally regarded as marginal (Valdeón 2008: 120), while also recognizing its artificiality and the systematic deletion of orality features due to lip-sync constraints, on the one hand, and translation routines and universals, on the other (Bruti 2019: 201; see also Pavesi 2005: 48). Valdeón (2008), for instance, observes the translation into Spanish of six types of inserts — namely, interjections, greetings and farewells, attention signals, hesitations, polite formulae, and discourse markers — reviewing their occurrences in the dubbing of three British films, i.e. *Four Weddings and a Funeral* (Newell, 1993), *Bridget Jones' Diary* (Maguire, 2001) and *Notting Hill* (Michell, 2000). The author underlines the lack of a consistent approach when translating these elements, which are, for the most part, ignored or overlooked, even though they would in fact require specific treatment within the transformative dubbing process (*ibidem*: 133).

As already pointed out, disfluencies, discourse markers and low register items (Valdeón 2009, 2011) are included to guarantee informality and colloquiality. In this respect, the translation of pragmatic markers is necessary to guarantee the audience's immersion in the fictional world. Moreover, they not only function as marks of naturalness and orality, but

they also fulfil a central role in supporting the processing of information in conversation, as well as in providing essential clues to interpret the interlocutor's attitude, intention and perspective "with respect to the propositional content of an utterance" (Galiano 2024: 502). As pointed out by Chaume (2004d: 844), when it comes to orality features, "[t]ranslators must understand the pragmatic meaning [...] since, most of the time, their translations are expected to produce the same effect on the addressees of the target text as the source text produced on its own addressees". More specifically, Chaume (2004d) investigated discourse markers in the Spanish translated version of the movie *Pulp Fiction* (Tarantino, 1994), considering the role these elements fulfil in verbal exchanges, i.e. to produce a coherent discourse while also making the speaker's intentions clear (*ibidem*: 845). The author observes different Spanish translations of the film, namely the written translated version, the dubbed translated version and the subtitled translated version (*ibidem*: 844). Generally speaking, results reveal no one-to-one correspondence between two languages when translating discourse markers; in most cases, these elements are simply omitted, affecting the balance between interpersonal meaning and semantic meaning. In particular, it is the interpersonal meaning that risks being lost (*ibidem*: 854)⁷⁶.

Similar considerations have also been made with regard to Italian dubbing. Bruti (2019)⁷⁷, for instance, observes that translated dialogues generally present fewer conversational features such as discourse markers, interjections, hesitations, false starts and disfluencies, as well as neater verbal exchanges (*ibidem*: 193). The extensive reduction of orality features is also demonstrated by the general trend remarked by Chiaro (2009) in dubbing, later corroborated by Bonsignori (2015), according to which the target dubbed text presents a lower number of words compared to the original version. Some features such as interruptions and unfinished sentences are more faithfully rendered in dubbing due to isochrony constraints, which motivate the use in the target language of conversational structures similar to the ones employed in the source language (Valdeón 2011: 227); on the other hand, elements like repeats are often omitted or standardized (*ibidem*: 231), with the risk of affecting characterization and altering the pragmatic effect (Bruti 2019: 201). In

⁷⁶ Nevertheless, Chaume (2004d) points out that even if the Spanish texts are less coherent than the original one, the Spanish audience is expected to rely on their linguistic and textual competences to understand less cohesioned texts, as well as on the visual mode to capture the intention of the source text (*ibidem*: 847). Thus, images can solve the possible ambiguities that may arise from less cohesioned texts (*ibidem*: 854) and help viewers retrieve interpersonal meanings in their entirety (*ibidem*: 855).

⁷⁷ Bruti (2019) observes structural and expressive aspects of conversation in both dubbing and subtitling; however, for the interests of the present discussion, the review of Bruti's considerations is limited to dubbing.

relation to discourse markers, Bruti (2019) observes the film *Love Actually* (Curtis, 2003), highlighting the propensity in Italian dubbing to eliminate features like *you know* and *I mean*. When translated, the former is mainly rendered through the literal translation *sai*, while the latter is rendered using a variety of pragmatic equivalents, like *insomma*, *voglio dire*, *cioè*, or *un momento* (*ibidem*: 203).

Bruti's (2019) considerations are perfectly in line with Galiano's (2024) more recent findings. Galiano (2024) compared the distribution of pragmatic markers in film dialogue, focusing on two films where these elements are particularly frequent, namely *Notting Hill* (Michell, 1999) and *Boyhood* (Linklater, 2014). The aim of the investigation was to establish the translation strategies used to render pragmatic markers (PMs)⁷⁸ when dubbing original Anglophone movies into Italian. According to the comparison of the original English dialogues with the dubbed Italian dialogues, omission seems to be the preferred strategy in translation. Indeed, in the film *Boyhood* more than half of the original PMs (54.57%) are omitted in Italian, compared to the 22.82% of PMs omitted in the dubbing of *Notting Hill* (*ibidem*: 515). It is interesting to point out that these elements are most frequently omitted in dubbing when they "contextually work as pause fillers and are less essential from a procedural and intersubjective point of view" (*ibidem*: 516). This is particularly relevant considering the object of study of the present investigation, i.e. hesitations, which, as noted in Chapter One, can work as fillers. As pointed out by Galiano, the omission of pragmatic markers still represents a loss with respect to orality and spontaneity, as well as textual coherence (*ibidem*). Even in this case, the analysis of the translation strategies reveals that, when translated, a variety of elements are used in dubbing to render the original PM (*ibidem*: 517); at the same time, stable routine translations are used with regard to the most frequent ones, namely: *well*, *you know*, *I mean* and *(al)right*. Most notably, *well* is translated as *bene*, *be'/beh*; *you know* is translated as *sai*; *I mean* is translated as *insomma* or *voglio dire*; *(al)right* is translated as *va bene* or *certo* (*ibidem*: 518-120).

On the whole, the studies discussed here allow to confirm some of the considerations previously made regarding the linguistic and translational behaviour of Italian dubbese, as well as to extend those considerations to various conversational features and pragmatic markers. Moreover, while it can be confirmed that the source text does to some extent

⁷⁸ Galiano (2024) includes 16 pragmatic markers, namely, among others, *well*, *(al)right*, *you know*, *sort of*, *sure*, *okay then* as for the English language (*ibidem*: 507) and *allora*, *dai*, *su*, *guarda*, *comunque*, *d'accordo* as for the Italian language (*ibidem*: 508). Other features such as interjections, like *oh*, *ah*, *mhm*, and general extenders, like *and things like that* were excluded from the analysis (*ibidem*: 505).

influence the target text, it is also important to notice that certain features are more subject to this influence than others. This is further discussed in the following section.

2.7.2. The Influence of the Source Language and the Standardization of Orality Features

Research has shown that, in dubbing, the impact of the source language is usually less strong at the syntactic level than it is, for instance, at the morphological level. This has been noted in various languages, including Spanish (Baños Piñero 2014, 2024) and Italian (Pavesi 2005, 2008; Bonsignori *et al.* 2012; Ghia 2014, 2019).

In her recent book, Baños Piñero (2024) observes the Spanish dubbing of the American TV series *Friends*. In so doing, the author consolidates previous findings obtained from the investigation of *Friends* (in its original and Spanish dubbed version) compared to the Spanish TV series *Siete vidas* (see Baños Piñero & Chaume 2009, Baños Piñero 2014). The author remarks that, although there is an attempt to evoke naturalness and spontaneity, the Spanish dubbed dialogues of the series are not as spontaneous-sounding as might be expected. Most notably, in the Spanish dubbed version of *Friends*, spontaneity is mainly conveyed through the imitation of orality at the lexical-semantic level. This level, however, appears to be ‘overloaded’ compared to the other linguistic levels, where orality features are either limited (at the phonetic-prosodic level, for instance) or completely non-existent (at the morphological level) (Baños Piñero 2024: 208-209). In other words, “the intersection between natural conversation and prefabricated discourse [...] reaches its peak at the lexical semantic level, is remarkable at the syntactic level, less significant at the phonetic-prosodic, and minor at the morphological level” (Baños Piñero 2014: 430). Even though there are cases in which orality markers are added in the translated text, their inclusion in the Spanish soundtrack seems to be influenced by the original English text, often resulting in the insertion of stereotyped elements that end up making the exchanges sound unnatural in the target language (Baños Piñero 2024: 209). At the phonetic-prosodic level, for instance, emphatic pronunciations are sometimes used even though they do not respect Spanish pronunciation conventions (Baños Piñero 2014: 415). Nevertheless, this also happens with other elements such as vocatives (*Señores Geller*) or interjections (*¡ups!* to translate *whoops!*) (Baños Piñero 2024: 194-196).

According to the author, this may also occur as a consequence of the neutralization process that commonly happens in translation due to linguistic standardization and synchronization constraints (Baños Piñero 2014: 431). Baños Piñero (2013) also highlights

this aspect when investigating the same dubbed product (i.e. the Spanish dubbing of *Friends*), focusing on adverbial intensifiers such as *very*, *so*, *really*, *totally* or *pretty*. Findings suggest that these features are commonly omitted or, more frequently, translated through degree adverbs (*ibidem*: 534), the latter being usually standardized. Thus, instead of elements mirroring colloquial conversation, the final product exhibits elements typically found in written texts. This is exemplified by the selection of features like *muy/tan* or *mucho/tanto* to translate the original English element *so*, leading to a Spanish dubbed version that is comparatively less emphatic and colloquial. There are cases in which more colloquial elements are used in translation to mirror the original register, as is the case of the English *so* translated using the adverbial phrase *un montón* (*ibidem*: 535-536), also resorting to elements that reinforce the sense of orality in the dubbed text, among which morphological resources (*aburridísimo*), idiomatic lexical bundles (*estoy como una cabra*) and exclamative clauses (*¡Qué detalle!*) (*ibidem*: 537-539). From the translational point of view, even though there is an attempt to move toward naturally-occurring conversation, standardization and conventionalism seem to be the preferred strategies. With regard to the translation of emphatic language, the author concludes that “Spanish dubbese is not fully based on the reproduction of spontaneous conversation but rather on the use of specific orality markers aimed at eliciting a particular type of response in the target audience” (*ibidem*: 540).

Similar results have also been pointed out by Romero-Fresco (2009a) in relation to both intensifiers and discourse markers. Romero-Fresco (2009a) conducted his investigation on the same case study illustrated by Baños Piñero, i.e. the Spanish dubbing of *Friends*. More specifically, the author compares the original and translated dialogues of the American TV series with the untranslated dialogues of an original Spanish product (*Siete Vidas*), as well as with spontaneous Spanish speech (retrieved from CREA, compiled by the Royal Spanish Academy). Findings show that while there is a similarity between the prefabricated language of *Siete Vidas* and real-life Spanish, the same cannot always be claimed for dubbed scripts. Indeed, while some dubbing choices reflect natural, conversational Spanish, it is also common to encounter elements that appear neither in the prefabricated language of *Siete Vidas* nor in the colloquial spoken data from CREA. As a result, unnatural renderings like *de veras* (*¿de veras?*) and *veamos* are used to translate *really* (*really?*) and *let’s see*, respectively. The former is typically used in writing rather than in oral speech (in fact, it occurs more commonly in the written section of CREA) (*ibidem*: 62), while the latter is indeed used in oral speech, but it is rather employed in more formal, non-colloquial settings (*ibidem*: 65). Thus, as reported by Romero-Fresco (2009a: 67), “orality does not suffice to

make dubbing dialogue idiomatic. It has to be the right type of orality, in this case colloquial conversation”; for this reason, the author distinguishes between orality and naturalness/idiomaticity (*ibidem*: 58; see also Romero-Fresco 2007, 2012). In the case of *veamos*, for instance, other more natural and colloquial solutions would be possible, like *a ver* (*ibidem*: 65). Most notably, the author remarks that more than 90% of the scenes observed in the analysis would have allowed a more natural translation (*ibidem*: 67). In some cases, the solutions found in dubbing cannot even be explained in terms of synchronization. For example, the use of *¿de veras?* to translate *really?* represents a less ‘ideal’ solution with respect to lip synchrony, adding a superfluous bilabial sound, which could be easily avoided by resorting to renderings like *¿en serio?* or *¿ah, sí?* (*ibidem*).

Similar translational behaviours have also been observed by Zanotti (2014b) in relation to the Italian dubbing of general extenders. Notably, when discussing the translation of the English extender *or something* through lexical units, the author points out that dubbing professionals frequently resort to translation routines like *o cose del genere/o qualcosa del genere* (*ibidem*: 123), regardless of the higher number of phonological syllables contained in the Italian rendering (*ibidem*: 136). Thus, similarly to what happens in Spanish dubbing, also in Italian translations the influence of the original English language can be noticed. In Italian dubbing, however, this mainly concerns pragmatic elements. This has been observed by Minutella (2021b) in her investigation of Anglicisms, calques and translation routines in 18 English language animated films produced between 1994 and 2019. Findings show that, in Italian, dubbing professionals aim at avoiding repetitiveness through variation and creativity in translation; therefore, the influence of the source language through direct borrowing is quite limited. Nevertheless, some translation routines are used when rendering frequent orality features such as *già* to translate *yeah* and *amico* to translate *man, buddy, dude, mate* and *pal* (*ibidem*: 94-96). Hence, the influence of the English language mainly concerns what can be defined as pragmatic Anglicisms, typically in the form of discourse markers and interjections (*ibidem*: 95).

In this regard, the translation of interjections in the context of Italian dubbing has been studied in detail by Bruti & Pavesi (2008). The authors compared the occurrence of interjections in 12 contemporary American and British films dubbed into Italian, also observing the way these elements are used in audiovisual dialogues as opposed to

spontaneous spoken English and Italian⁷⁹. Through the comparison of film dialogue with spontaneous language, results show that prefabricated dialogue does present a limited number of interjections compared to their occurrences in spontaneous interactions (*ibidem*: 216). Moreover, the authors remark that, in dubbing interjections, the source language seems to influence the target version. More specifically, the most frequent interjections of spontaneous Italian (*eh/eh?*) tend to be underused in dubbed film speech (*ibidem*: 216-217), while the English interjection *oh* is preferred (*ibidem*: 217). Thus, “Italian interjections which exhibit a degree of similarity with English ones tend to be over-represented in dubbing, whereas interjections which are specific and restricted to Italian tend to be under-represented” (*ibidem*: 119-120). As a matter of fact, *oh* is the most frequent interjection in both British and American English; its occurrence in Italian dubbing can be attributed to the impact of the original English text on the dubbed translated version (*ibidem*: 218-219).

2.7.3. The Use of Translation Routines

Contrary to what has been witnessed in relation to lexis and pragmatics, the language of dubbing seems to be quite independent from the original text when it comes to syntax. On this matter, Pavesi (2005, 2008) observed marked word order constructions such as right and left dislocations, as well as cleft sentences in Italian dubbing.

Through the investigation of five films⁸⁰, the author notices that the differences emerging between dubbed film speech and spontaneous speech are quite limited and less influential than the points of agreement found between the two (Pavesi 2005: 135). In particular, findings suggest that “most marked word order constructions are “added” to the translations” (Pavesi 2008: 92) as a reflection of the norms of the target language (see Pavesi 2005: 136). Most notably, the author points out that even though left dislocations are more frequent in spontaneous Italian, right dislocations are preferred in dubbing, probably due to the fact that they can easily be inserted in the text through translation routines without having to extensively rewrite the original text (*ibidem*: 100-101) — for instance, *do you know (that)* translated as *lo sai che* (Pavesi 2008: 94). The influence of the source language more

⁷⁹ Bruti & Pavesi (2008: 212) based their account of the frequency of interjections in spontaneous spoken English on Biber *et al.*'s (1999: 1096-1098) grammar. As for the Italian language, the authors used the data obtained from the corpus *Lessico di Frequenza dell'Italiano Parlato* (LIP) (De Mauro *et al.* 1993).

⁸⁰ The results described by Pavesi (2008) are drawn upon a previous investigation presented in her monograph *La traduzione filmica: aspetti del parlato doppiato dall'inglese all'italiano* (Pavesi 2005), where she observes the following titles: *Sliding Doors* (Howitt, 1997), *Notting Hill* (Michell, 1999), *Finding Forrester* (Van Sant, 2000), *Dead Man Walking* (Robbins, 1995) and *Secrets and Lies* (Leigh, 1996).

commonly occurs with left dislocations, which have a substantial discursive and interactional value, i.e. involving larger portions of text (for example, involving multiple conversational turns) (see Pavesi 2005: 92-93). Apart from that, the translated texts display renderings that are faithful to spontaneous Italian. Albeit prototypical, the solutions found for left dislocations are independent from the source language (*ibidem*: 94)⁸¹. The propensity to move towards the norms of the target language is also supported by the data obtained regarding the translation of cleft sentences. Cleft sentences are usually more susceptible to source language interferences (*ibidem*: 134); in the Italian dubbed text, this is reflected by the occurrence of calques from English. Nevertheless, it is interesting to notice that the translation strategies employed in dubbing do not ‘force’ the language by inserting linguistic structures that are completely unrelated to the Italian system; rather, translators intensify the frequency of constructions that are possible, although less frequent, in spontaneous spoken Italian (*ibidem*: 127-128) — for example, *certo che* to translate the original *of course* (*ibidem*: 117). Hence, it is possible to conclude that, as for the syntactic structures analysed, in Italian dubbing “the naturalness of the target language solutions is not carried over from English but comes from within the target language” (Pavesi 2008: 93), also resorting to formulaic language and translation routines. The emergence of translation routines in Italian dubbing has also been highlighted in other constructions such as conversational routines (Bonsignori *et al.* 2012; Bonsignori & Bruti 2014) and interrogative constructions (Ghia 2014, 2019).

Conversational routines, for instance, have been observed in both films (Bonsignori *et al.* 2012) and TV series (Bonsignori & Bruti 2014), with the intention of examining whether social dynamics in dubbing are coherent to the narrative of the original English films. Most notably, Bonsignori *et al.* (2012) investigate greetings and leave-taking as conversational openings and closings, as well as in relation to the character’s idiolect, the use of slang, verbal habits and so on. More specifically, the authors examined a self-constructed corpus of 15 contemporary films dubbed from English into Italian and compared it with a reference corpus of five original Italian movies (*ibidem*: 358)⁸². Data obtained from

⁸¹ An example of left dislocation added to the Italian text independently from the original English version is provided by the translation of the line “*Elastic Thrombosis*. They’re guilty of lyric poaching”, extracted from *Sliding Doors* and rendered in Italian as “*Gli Elastic Thrombosis*. *Questi sono colpevoli di plagio*” (Pavesi 2005: 89, emphasis in the original).

⁸² The corpus of British or American productions dubbed into Italian consisted of fifteen titles, namely: *The English Patient* (Minghella, 1996), *Oscar and Lucinda* (Armstrong, 1997), *Sliding Doors* (Howitt, 1997), *Bridget Jones’s Diary* (Maguire, 2001), *Love Actually* (Curtis, 2003), *Green Street Hooligans* (Alexander, 2005), *Match Point* (Allen, 2005), *Becoming Jane* (Jarrold, 2007), *Music and Lyrics* (Lawrence, 2007), *In Her*

the analysis reports that, in general, slang expressions are commonly neutralized in dubbing, especially when there is no corresponding element in Italian (*ibidem*: 367). Moreover, there are cases where the influence of the source text undermines the naturalness of the target text as, for instance, in the case of *that's all* routinely translated with *è tutto*, which is identified as a pragmatic Anglicism (*ibidem*: 366). Nevertheless, translators seem to regard these features with great consideration, maintaining more or less the same number of occurrences as the source text (*ibidem*: 372-373) through the addition of greetings, leave-taking and good wishes to compensate for those instances in which these elements are omitted in the translation (*ibidem*: 363, 372). Accuracy seems to have become even more evident in more recent films in terms of coherence in register and in characters' portrayal; according to the authors, when more substantial variations occur, these are attributable to the adapters' misinterpretation of the mimicry of the film, making dubbed characters sound more straightforward than they are in the original version (*ibidem*: 368).

As for television dialogue, instead, Bonsignori & Bruti (2014) found that formulaic language is more common in openings than it is in closings; in the latter, more variety among the patterns used in translation has been observed. The authors investigated a self-compiled parallel corpus consisting of two Anglo-American TV series: *Brother & Sisters* and *Gilmore Girls*. Generally speaking, in both series, openings are not only more numerous than closings, but they are also granted more space (*ibidem*: 98). Contrarily, the overall number of leave-takings is higher in dubbing, where they are almost always translated (*ibidem*: 105). As for the translation strategies adopted, greetings are usually rendered through the informal greeting *ciao* so as to mirror the same level of intimacy existing between characters in the original version (*ibidem*: 100). Some general variation, however, can be noticed in the target text. Sometimes, this results in a more unnatural effect in dubbing, as is the case of *hello* and *hi* being translated with *buongiorno* — which belongs to a more formal register (*ibidem*: 101) — or in the use of the more neutral and 'passé-partout' expression *salve* used to register "a change in the ongoing situation" (*ibidem*: 103). Leave-takings present more variety in translation, ranging from the informal *ciao* to the more formal *arrivederci* and *addio*, as well as to the uncommon expression *ti saluto* (*ibidem*: 105). The wide range of features used to close verbal exchanges — such as *a più tardi*, *a dopo* and *notte/buonanotte* — can be

Shoes (Hanson, 2005), *The Devil Wears Prada* (Frankel, 2006), *The Holiday* (Meyers, 2006), *The Duchess* (Dibb, 2008), *Revolutionary Road* (Mendes, 2008), *Sex and the City - The Movie* (Patrick King, 2008). The corpus of Italian productions consisted of five films, namely: *L'ultimo bacio* (Muccino, 2001), *Ma che colpa abbiamo noi* (Verdone, 2003), *Io e Napoleone* (Virzi, 2006), *Manuale d'Amore 2 (capitoli successivi)* (Veronesi, 2007) and *Tutta la vita davanti* (Virzi, 2008) (Bonsignori et al. 2012: 358).

attributed to the fact that these elements are usually added in the translated text, resulting in a higher number of occurrences compared to the source text (*ibidem*: 107). For example, the leave-taking expression *ci vediamo* is commonly added in dubbing, even when no leave-taking is present in the source text and the verbal exchange is closed by an expressive act (*ibidem*: 108) such as a compliment or a word of thanks (*ibidem*: 91).

The adaptation of audiovisual products in relation to the interactional and social dynamics displayed on the screen has also been observed with respect to interrogatives as orality markers with a pragmatic value (Ghia 2014, 2019). As reported by Ghia (2014), direct questions represent an interpersonal phenomenon widely spread in spontaneous speech, where they fulfil various functions — among which, the search for information, the management of turns, as well as managing the relationships among speakers (*ibidem*: 58-59) — and carry out various speech acts such as requesting, inviting and offering, as well as complaining and showing (dis)agreement (*ibidem*: 59). Drawing on the PCFD, the relevance of these features in screen interactions is underlined by the fact that questions are way more present in fictional dialogues than they are in spontaneous conversation. This is probably due to the diegetic function these elements fulfil, namely conveying information about the character and the plot (*ibidem*: 66). Interestingly, the same occurs in dubbing; in fact, interrogatives are more present in translated texts than they are in the original film dialogue (*ibidem*: 72). This is partly due to the tendency observed in dubbing to add what Kozloff (2000: 47) defines as “verbal wallpaper”, i.e. incidental dialogue (*ibidem*: 99). From the qualitative point of view, particular attention seems to be paid in Italian dubbing to the preservation of naturalness, introducing ellipsis, marked word orders and weak connectors as compensation strategies for those English questions that do not have an equivalent rendering in Italian (*ibidem*: 77). Therefore, although Italian dubbing differs to some extent from the source text, it faithfully reproduces Italian conversational patterns, opting for translations that convey a sense of realism and spontaneity, while also retaining most of the original diegetic functional meanings (*ibidem*: 77). The same is also confirmed in a later study on direct questions as a means to express affiliation and disaffiliation (i.e. positive or negative personal stance) in original English dialogues dubbed into Italian (Ghia 2019). Even in this case, the investigation of the PCFD displays a certain degree of autonomy of the dubbed text with respect to the source text. More specifically, in the source text, disaffiliation is mainly conveyed through questions including inserts, emotionally-loaded chunks and non-canonical word order (*ibidem*: 587), while no specific pattern emerges as a marker of affiliation (*ibidem*: 594). In Italian dubbing, the same pragmatic values are rendered

differently: disalignment is mainly conveyed through weak connectors like *ma* (*ibidem*: 593), while affiliation is rendered through various syntactically marked structures, including cleft constructions and dislocations (*ibidem*: 591-592).

To conclude this section, the data presented so far has allowed to pinpoint some of the features of spontaneous language reproduced in audiovisual dialogues and how said features are translated in dubbing. As reported by Pavesi (2008: 90):

Internal dubbing norms may in fact account for reliance on some features to represent orality [...] the deletion of others to keep within the constraints set by films' compressed time scale [...] and the choice of only some features to convey and emphasize the interactionality of face-to-face communication witnessed from a distance [...].

More specifically, the studies reviewed in the present discussion have underlined some of the typicalities of Italian dubbing, the translation routines identified for specific markers of orality and the way translators resort to specific translation strategies — sometimes influenced by the ST — and/or to more creative solutions depending on factors such as the story, the characters, the relations between characters, and the type of register. Along with syntactic features and constructions, various conversational features and pragmatic markers have been observed: discourse markers, interjections, intensifiers and so on. Nevertheless, among the pragmatic markers reproduced in prefabricated orality, translators also operate with hesitation markers and the variety of pragmatic meanings conveyed through them. As discussed above, research has shown that pragmatic features tend to be overlooked and eliminated in dubbing, particularly when they are employed as fillers. This study aims to investigate whether the tendencies identified in previous research with respect to pragmatic markers, such as omission and standardization, also apply to hesitations in dubbing, and whether the strategies used to render these features have evolved over time. An overview of the studies conducted on hesitation markers in dubbing is provided in the following section.

2.8. Dubbing Hesitations

A brief overview of how hesitations are treated in fictional texts (Jucker 2021; Tonetti Tübben & Landert 2022) has been provided in Chapter One (see §1.8.2). The discussion has underlined how authors insert hesitations in fictional texts for specific purposes. Most notably, some of the functions and characteristics hesitations display in spontaneous conversations are also rendered in audiovisual texts; in other cases, hesitations are

reproduced in fiction with the aim to evoke some general stereotypes commonly associated with these elements — for characterization purposes, for instance — that do not fully reflect the way they are actually used in impromptu conversation (Jucker 2015b: 65). Nevertheless, having established the role recognized to these elements in fictional texts, it is important to focus on their translation; more specifically, how are hesitations treated in AVT and what are the procedures used in dubbing to translate hesitation markers. As already observed in the first chapter of this investigation (see §1.3), there is no universally agreed-on definition of hesitations; thereby, a wide variety of linguistic and paralinguistic features can be labelled under this term, also varying across languages. Hence, in AVT, hesitations have been studied by scholars through a variety of features such as discourse markers playing the role of hesitations and (self-)repair markers (Romero-Fresco 2007) or in combination with other features typical of oral speech, like interjections and attention signals, as well as disfluencies such as interruptions and repeats (Valdeón 2008; Bruti 2019).

Romero-Fresco (2007) attempted to observe the degree of naturalness of Spanish dubbing through the investigation of hesitation markers. In particular, the author compared the prefabricated orality of the Spanish dubbing of the American sitcom *Friends* to the prefabricated dialogue of an original Spanish sitcom (*Siete Vidas*, non-translated), as well as to Spanish spontaneous conversation (CREA). Romero-Fresco specifically focused on seven discourse markers used to render self-corrections and pauses in Spanish (*ibidem*: 189) — namely, *bueno*, *o sea*, *vamos*, *pues*, *es que*, *claro*, and *hombre* — drawing attention to the fact that every linguistic item contained in the dialogue — including disfluencies and errors — fulfils a specific function in the “overall communicative goal of the dramatic dialogue” (Baumgarten 2005: 86 cited by Romero-Fresco 2007: 188). A first quantitative analysis showed that the abovementioned discourse markers fulfilling the role of hesitations are less frequent in dubbed Spanish (dubbed *Friends*) compared to their occurrences in the original Spanish dialogue (*Siete Vidas*); at the same time, the dialogues of *Siete Vidas* presented fewer discourse markers compared to spontaneous language (CREA) (*ibidem*: 193). As for the qualitative analysis, the investigation showed a general tendency to overlook these elements in dubbing. Indeed, hesitations are mainly omitted or translated through three privileged carriers of orality, i.e. *bueno*, *en fin* and *verás*, of which *en fin* and *verás* tend to be used in an unnatural way, far from the way they are utilized in both spontaneous and prefabricated conversations (*ibidem*: 199). The dubbed translation not only reveals a general lack of variation when it comes to linguistic items conveying the sense of hesitation but also the propensity to opt for “more monotonous, less colorful, less idiomatic and, overall, less

natural” solutions (*ibidem*) compared to the language of Spanish spontaneous colloquial dialogue.

Similar results have also been noted by Valdeón (2008) in his analysis of inserts and disfluencies in Spanish, including hesitations. Differently from Romero-Fresco’s study, where hesitations are observed by the means of discourse markers, Valdeón focuses on vocalizations such as *uh* and *er*, i.e. on the same phenomena considered in the current investigation. Nevertheless, in Valdeón’s study, hesitators are examined among a group of other orality features such as interjections, greetings and farewells, attention signals, polite formulae and discourse markers in three British films dubbed into Spanish, namely *Four Weddings and a Funeral* (Newell, 1994), *Bridget Jones’ Diary* (Maguire, 2001) and *Notting Hill* (Michell, 1999). With regard to the adaptation of hesitators, the author notices that there are various phenomena in Spanish capable of fulfilling similar functions to the original gap-filling elements (*uh* and *er*), i.e. lengthened vowels, repeated syllables, gap-filling words like *bueno*, as well as unintelligible words or noises. Despite the wide range of possible solutions, hesitations are most frequently omitted in Spanish dubbing (*ibidem*: 127). Generally speaking, the author highlights the lack of a consistent approach in translating inserts and, more specifically, hesitations (*ibidem*: 133).

These findings are in line with the considerations made as for the Italian dubbing industry. Bruti (2019), for instance, observes hesitations among a group of orality markers such as interruptions and unfinished sentences, interjections and formulaic language. With regard to hesitation markers, the author investigates the dubbed version of the film *Notting Hill* (Michell, 1999), where pauses filled by vocalizations in the source text are not always preserved in the target text. More specifically, original hesitations are generally substituted by a variety of items used to convey naturalness and spontaneity in dubbing; namely, next to vocalizations such as *eh* and *ehm*, filler phrases like *in fondo*, *be’* or *a dire il vero* are used (*ibidem*: 202-203).

In general, hesitations tend to be neglected in AVT. When adapting dialogues for dubbing, the tendency to add unnatural linguistic features can be noticed (Romero-Fresco 2007: 199), as well as the lack of a consistent approach in translation (Valdeón 2008: 131). As a matter of fact, hesitations are simply omitted from the text (Romero-Fresco 2007: 199) or substituted with indistinct noises that may compromise the naturalness of the dubbed version (Valdeón 2008: 127). As pointed out by Baños Piñero (2014: 439), the strategies employed by translators could be associated with the general tendency to prioritize clarity, thus explicitly verbalizing hesitations and pauses in dubbing in the same way as lexical

precision is favoured over vague language (Chiaro 2000; Zanotti 2014b). Nevertheless, if hesitations are omitted from the target text, changes may occur in the way viewers experience the narrative development of the story, altering the protagonists' characterization as well as interpersonal dynamics (Valdeón 2008: 125; Bruti 2019: 197). This is particularly relevant when hesitations are edited out of the speech of a shy and insecure character (Bruti 2019: 197). Generally speaking, the omission of features like hesitations does not seriously alter the structure of the story since the deletion of these elements does not manipulate the denotative meaning of the sentence (Chiaro 2009: 151); yet, as pointed out by Bruti (2019), there may be consequences on the original interpersonal dynamics between characters and how they are perceived by the target audience (*ibidem*: 197).

In this regard, Calogero (2024) specifically observed vocalizations and pauses of silence with the aim to ascertain whether the way these elements are rendered (or not rendered) in dubbing can change the perception of characters, leading them to a performance that does not correspond to the original one. Notably, the investigation was conducted on the translation of hesitations as filled and unfilled pauses, observing the role they play in the characterization of Hugh Grant's characters in four British romantic comedies, i.e. *Four Weddings and a Funeral* (Newell, 1994), *Notting Hill* (Michell, 1999), *Bridget Jones' Diary* (Maguire, 2001) and *Love Actually* (Curtis, 2003). The analysis was carried out through a comparison of the original English dialogues with Italian dubbed dialogues. Moreover, when possible, the adaptations made for other languages such as French and Spanish were also considered. Results showed that the translation strategy deployed in Italian dubbing can indeed alter characterization, making the original shy and awkward protagonist sound more talkative, confident and assertive in the target version (*ibidem*: 76). In addition to what has been noted by Bruti (2019) with regard to the translation of orality features in Italian dubbing (i.e. the tendency to resort to omission and substitution strategies), Calogero's (2024) multilingual comparative analysis highlights the propensity not only to use indistinct (nasal) noises in lieu of the original vocalization, but also the general preference to render hesitations through the lengthening of vowel sounds at the end of a word (*ibidem*: 76)⁸³. On the one hand, Calogero's investigation offers a first attempt to study hesitations as a phenomenon in its own right, that is, focusing uniquely on the translation of hesitations as filled and unfilled

⁸³ In this regard, lengthenings have been observed as an intrinsic characteristic of Spanish dubbese (Baños Piñero 2014: 415) and dubbitis (Sánchez-Mompeán 2020a: 116-117), hence as a conscious technique used by dubbing actors to enhance synchronization with the movements of characters' mouths. This finding seems to suggest that the use of prolongations at the end of a word in order to render hesitations could be understood as another characteristic of dubbing language.

pauses in dubbing, without taking into account other concomitant linguistic phenomena such as disfluencies (for instance, false starts or repeats) or discourse markers. On the other hand, similarly to the other studies reviewed here, Calogero's investigation constitutes a case study based on a small corpus of films, which focuses on the stereotypical ways in which scriptwriters and actors use these elements as a shortcut to characterization (Jucker 2015b); thereby, it does not allow to make generalizations on hesitations in film speech and the way they are treated in Italian dubbing.

For the abovementioned reason, extensive work on hesitations in film dialogue is necessary. This study thus aims to fill this void. Drawing on the corpus-based research reviewed here, the current investigation aims to examine whether any recurrent pattern can be established for hesitations in film speech while also observing the way these features are tackled in Italian dubbing. In addition, this study is carried out through a diachronic perspective, which will allow to shed some light on the evolution of film speech and dubbing over time; that is, any change regarding how orality is codified in film speech and how it is treated in dubbing.

Considering the diachronic approach employed in the analysis, it is worth focusing on previous studies conducted on the evolution of film speech and dubbing. A review of the main investigations conducted along these lines is provided hereinafter.

2.9. Diachronic Studies on Film Speech, Prefabricated Orality and Dubbing

2.9.1. The Evolution of Film Speech over Time: Previous Research

Some of the scholars interested in attesting the extent to which audiovisual dialogue approximates spontaneous conversation have provided a first glimpse into the evolution of telecinematic language over time, mainly focusing on film speech and highlighting some of the changes it has undergone in relation to its adherence to impromptu conversation.

One of the most extensive studies on the language of films has been carried out by Veirano Pinto (2014). In her investigation of the *North American Movie Corpus* (NAMC), Veirano Pinto (2014) examined multiple variables such as genre, ratings and, among other factors, the year of release. Through the analysis of 640 films distributed between 1930 and 2010, the author draws attention to the degree of closeness of movie language to spontaneous

speech⁸⁴. In particular, results show that the language of films from the 1970s and 1980s is, generally speaking, closer to spontaneous conversation; on the contrary, the mean score of the movies distributed between 1930 and 1950 suggests that the movies created in these decades contain more features of planned speech. Furthermore, the films from the 1960s, 1990s and 2000s present features of both poles, placing themselves at the centre of the continuum going from impromptu conversation, on the one hand, and planned speech, on the other (*ibidem*: 123).

The results obtained by Veirano Pinto (2014) are further corroborated by Zago's (2016) findings on film remakes. In particular, Zago (2016) focuses on film speech with the aim to evaluate whether the language of films has become more colloquial over time, where colloquiality is described as "the language of immediacy/closeness in non-binding contexts" (*ibidem*: 31). In order to do so, the author compares original American movies from the 1950s and 1960s with their remakes from the 1990s and 2000s, respectively. As reported by Zago (2016: 72), "the remake approach is adopted as the testing ground for colloquialisation of film dialogue [...] because [...] what changes — the language — can be highlighted against the backdrop of what remains practically the same — the situational contexts". In general terms, Zago's findings seem to offer some support to the claim that film dialogue has become more colloquial over time (*ibidem*: 111). This is supported by the positive scores obtained from the analysis of remakes regarding features typically associated with colloquiality such as interactions invoking personal facts and private identities, intimacy and familiarity between the interactants, and the focus on interpersonal relationships rather than on the presentation of facts and events (*ibidem*: 113). Comparing the scores extracted from Veirano Pinto (2014) and from Zago (2016) regarding the spontaneity and colloquiality of American films distributed during the same decades (1950s/1960s and 1990s/2000s) (see Table 2.1), it is possible to establish that more recent films seem to present a higher score compared to older films (Zago 2016: 161)⁸⁵.

⁸⁴ It must be pointed out that Veirano Pinto (2014) studied the language of American films by resorting to movie subtitles.

⁸⁵ More specifically, this comparison was carried out by Zago (2016) by confronting his Factor 1 scores concerning colloquiality (*ibidem*: 102) with Veirano Pinto's (2014) Factor 2 scores concerning spontaneous discourse (*ibidem*: 122); in other words, by limiting Veirano Pinto's findings to the four decades studied by Zago, factor analyses carried out on similar linguistic features were compared. Results show that more recent films exhibit more positive factor scores than older films (Zago 2016: 160-161).

Table 2.1. Comparable Data from Veirano Pinto (2014) and Zago (2016)		
Zago's (2016: 110) Factor 1 Scores		
Film	Original	Remake
<i>Father of the Bride</i> (1950/1991)	-4.01661011	-2.19943777
<i>Born Yesterday</i> (1950/1993)	6.11023575	4.44672307
<i>The Narrow Margin</i> (1952/1990)	0.17757128	1.89476969
<i>The Desperate Hours</i> (1955/1990)	-1.73470780	0.93834409
<i>Ocean's Eleven</i> (1960/2001)	-5.37388876	-2.88000707
<i>Charade/ The Truth about Charlie</i> (1963/2002)	-0.95228165	-1.97071827
<i>Guess Who's Coming to Dinner/Guess Who's</i> (1967/2005)	1.82947130	3.11660765
<i>Yours, Mine and Our</i> (1968/2005)	-1.27758710	0.99823602
Veirano Pinto's (2014: 122) Factor 2 Scores		
Decade	Score	
1950	-3.03	
1990	1.78	
1960	-1.43	
2000	1.91	

Furthermore, Zago's research not only provides interesting data regarding the evolution of film speech over time, but it also allows to pinpoint some fascinating differences between Italian and American film speech during the same decades. In fact, Zago points out that "colloquiality is not a prerogative of the remakes", meaning that "back in 1950 American film dialogue was already capable of a good approximation to spontaneous conversational style" (Zago 2016: 148); drawing on Rossi's (1999) investigation of Italian films (see §2.4.2), the same cannot be claimed for 1950 Italian film dialogue. Although colloquiality is present, the language of Italian films is overall more formal than the language of American films from the same period (Zago 2016: 159).

Finally, to conclude this diachronic overview of studies conducted on the evolution of audiovisual dialogue, it is important to mention Werner's (2021) recent study on telecinematic language. More specifically, Werner investigates the language of both films and TV series from six English-speaking countries in two large-scale corpora, namely the *TV Corpus* (from 1950 to 2017) and the *Movie Corpus* (from 1930 to 2018) (see Davies 2021). More specifically, the author observes specific items relating to emotionality and informality; the former being investigated through features like emotive interjections, emotion verbs, intensifiers and swear words (i.e. *wow*, *oh my god*, *geez*, *I can't believe*, *really*, *shit*), and the latter being investigated through features like greetings, vocatives and non-standard features (i.e. *hi*, *hey*, *farewell*, *cheers*, *honey*, *sweetie*, *dude*, *gonna*, *wanna*, *gotta*). In line with the abovementioned results (Veirano Pinto 2014; Zago 2016), Werner's (2021) findings offer support to the fact that the language of both films and television series has increasingly become more emotional and more informal over time (*ibidem*: 63). The

higher degree of informality in film speech has also been highlighted by Jucker & Landert (2023). The authors investigated Davies's (2021) *Movie Corpus* with the intention of exploring politeness and impoliteness values in films. Through the analysis of a number of features, among which title nouns (i.e. *Sir* and *Madam*) and formulaic expressions (i.e. *could you/would you*), the authors ascertained a marked increase of swearwords in film speech over the last five decades (from the 1970s/1980s to recent years) (Jucker & Landert 2023: 137), thus offering support to the claim that telecinematic language has become more colloquial and informal over time.

Having discussed the evolution of telecinematic language over the years, the following section will present diachronic findings concerning Italian dubbing.

2.9.2. Diachronic Studies on Dubbing and Orality Features

Diachronic studies conducted on Italian dubbing have investigated the adaptation of both cinematic and television productions (Motta 2015), observing various phenomena, namely word formation processes (Bonsignori 2015), familiarizers (Forchini 2013) and expletives (Formentelli & Ghia 2021) with the intention of ascertaining the evolution of dubbese over the years (Minutella 2021b). Many of these investigations have been conducted comparing the original version with spontaneous conversation in the source language, on the one hand, and the original English version with the Italian dubbed version, on the other. Therefore, in addition to examining the evolution of Italian dubbing, some of the considerations made here also fit into and complement the discussions reported in the previous section (§2.9.1).

Motta (2015), for instance, investigated Italian dubbing in relation to original American television serials, taking into account the evolution of the transmission stations in Italy. In particular, the author observes dubbing through the years, distinguishing between paleotelevision (paleotv) and neotelevision (neotv), the former consisting of the national broadcasting stations and the latter being characterized by the influence of private broadcasting stations (*ibidem*: 951-952). Through a diachronic perspective, the author aims to ascertain any differences in the modalities used by translators to adapt Anglo-American texts. The study was carried out observing different genres of American serials produced between 1959 and 2003 — namely, science fiction, family and teen, police drama, and legal drama — and their Italian versions (dubbed between 1960 and 2003) (*ibidem*: 955-956). On the one hand, Motta emphasizes the presence of some stable features in the evolution of dubbing such as the use of focus constructions, like dislocations and cleft sentences (*ibidem*:

959). On the other, many of the changes attested by Motta in the language of dubbed television are attributable to the evolving nature of the texts being dubbed; that is, as American productions started to embed characters from different social classes and speaking different linguistic varieties, morpho-syntactic traits characteristic of spoken language tended to be increasingly incorporated in dubbing as well. This allowed to compensate for the sociolinguistic levelling in standard Italian that typically occurs in dubbing. At the same time, this resulted in a higher number of swearwords and explication of double meanings that were otherwise not rendered in Italian (*ibidem*: 960-961). The most noticeable difference, however, regards the presence of the English language in dubbed products. According to the author, English has always been present; nevertheless, in older titles, it was usually restricted to character names. In more recent dubbings, instead, English has become remarkably more present: it occurs in external elements, like titles and captions (*ibidem*: 960-961), as well as in internal elements. As for the latter, more recent dubbed serials present, for example, a higher number of English loans, lexical calques and syntactic structures that reproduce the original ones (*ibidem*: 964-965). Interestingly, “[t]he effect, though, is not that of an unnatural language with a high rise in formality [...] or that of a pseudo translation” (*ibidem*: 968); on the contrary, Motta claims that the presence of English “is linked to a sense of efficiency that renders Italian more permeable to a higher frequency of abbreviations, inscriptions, and especially calques, which deeply model contemporary Italian” (*ibidem*).

The increasing impact of the English language on Italian dubbed products has also been noted by Minutella (2021b) in the adaptation of animated films. Most notably, Minutella (2021b) investigates Anglicisms and calques in the Italian dubbing of original English language animated films. In so doing, she pinpoints some interesting data regarding Italian dubbese and the evolution of the relationship between Italian dubbing and the English language over the span of 27 years (1994-2019). In particular, among the films analysed, Minutella observes movies belonging to a saga (i.e. prequels and sequels) in order to highlight the different degrees of Anglicisations over time. In this regard, the film *Toy Story* (Lasseter, 1995) and its sequels were taken into consideration (*ibidem*: 78). The analysis shows that the year of production seems to influence the occurrence of the interjection *okay* in dubbing. More specifically, the first two movies of the saga, *Toy Story* (Lasseter, 1995) and *Toy Story 2* (Lasseter, 1999), contain fewer occurrences of this interjection in their Italian dubbed version compared to the films released in the new millennium, i.e. *Toy Story 3* (Unkrich, 2010) and *Toy Story 4* (Cooley, 2019) (*ibidem*: 83). In particular, in the two most

recent titles, the author notices a pervasive presence of the Anglicism both as a response elicitor and as a turn-taking tool (*ibidem*: 84). In other words, the interjection *okay* “performs a variety of pragmatic functions in the Italian dialogues of the more recent films”; as a result, “[t]here seems to be an increase and an overextension of the Anglicism, which may mirror the large use of this discourse marker in contemporary Italian conversation, especially among young people” (*ibidem*). A similar tendency has also been pointed out by Minutella concerning the emotive interjection *wow*, used to express surprise and awe (*ibidem*: 84). This feature is absent in older movies such as *Mulan* (Bancroft & Cook, 1998) and *The Emperor’s New Groove* (Dindal, 2000), while it has become more pervasive in the Italian dubbing of more recent titles such as *Zootopia* (Howard & Moore, 2016) and *Incredibles 2* (Bird, 2018) as the result of an overextension process (*ibidem*: 84-85). Thus, Italian dubbing seems to have become more absorbent of the source language over the years, especially when it comes to linguistic features fulfilling pragmatic functions.

An opposing trend, however, has been noticed by Formentelli & Ghia (2021) in relation to expletives in non-fictional movies. The authors investigate emotionality expressed through intensifying expletives, namely *bloody*, *fucking*, *(god)damn*, *the fuck* and *the hell* with the intention of answering the following two questions: how are they used in Anglophone films and how are they dubbed into Italian? Drawing on the PCFD, the investigation has been carried out through a short-term diachronic study, observing film dialogue over the span of approximately twenty years (from 1995 to 2017). In line with Jucker & Landert’s (2023) findings, Formentelli & Ghia (2021) outline how expletives almost doubled in the source language over time, highlighting the central role these elements fulfil in expressing emotionality in audiovisual products (*ibidem*: 56).

As for the Italian rendering, four translation strategies were considered: omission, close rendering, lexical/syntactic replacement and compensation (*ibidem*: 53). Over the time span analysed, an increase in omission is noticed, accompanied by a decrease in the use of the other translation strategies, which results in the lower occurrences of expletives like *bloody* and *fucking* in the translated text. In the case of *bloody*, this equates to lowering the pragmatic force of the dialogue (*ibidem*: 57), while for *fucking* other strategies like replacement and compensation (usually through the use of the interjection *cazzo*) manage to preserve the emphatic strength of the original text (*ibidem*: 59). The only expletive exhibiting a marked increase in dubbing over time is *the hell*, also characterized by a lower percentage of omission (*ibidem*: 62). This expletive is usually rendered in Italian using a well-established translation routine that mimics the original formulaic expressions in which this

expletive occurs. Nevertheless, even though the Italian rendering calques the original structure, the translation of *the hell* with *che diavolo* softens the original offensive force. In this respect, target-oriented solutions are rare and limited to older films (*ibidem*: 63). The opposite occurs, for instance, in the case of *the fuck*, which remains quite stable through the years, being translated through stock phrases built around the target expression *cazzo* (*ibidem*: 61-62), which instead maintains the original vulgar force (*ibidem*: 67). In the first decade analysed, routine formulae are widely used in Italian dubbing; *bloody* and *fucking*, for instance, are rendered using great variation in older movies, creating clusters of patterns and recurring expressions in the target language (*ibidem*: 60), which then diachronically disappear, shifting toward target language models (*ibidem*: 67). Translation routines thus change diachronically, moving towards a tendency to delete and mitigate expletives in dubbed Italian. These results prove Italian dubbing to be moving in a different direction compared to other languages like Spanish, where swearwords and taboo words are said to have become exponentially more frequent (Valdeón 2015, 2020). The same does not apply to Italian dubbing, which “resorts to its own combinations of linguistic devices to transfer the emotional and pragmatic force of the original dialogues” (Formentelli & Ghia 2021: 69), while also displaying a still existing inclination to a partial (self-)censorship of expletives and taboo words. This is also corroborated by Zamora & Pavesi’s (2021) investigation of swearwords in a parallel corpus of American films dubbed in both Spanish and Italian. The authors remark that in both languages, swearing expressions tend to be mitigated; however, while the reduction is particularly important in Italian, Spanish dubbing seems to be more inclined to reproduce the vulgarity of the source text.

Conversely, a different trend has been pointed out by Bonsignori (2015) with regard to teen talk in Italian. In this case, dubbing turns out to be moving toward a more colloquial language by means of creative solutions that convey a sense of naturalness to the target language, with compensation strategies being preferable to mere omission to translate a wide range of phenomena. More precisely, Bonsignori (2015) investigates the language of adolescence and the role of language as an identity marker in three American movies produced in different decades, namely *Sixteen Candles* (Hughes, 1984), *Clueless* (Hecherling, 1995) and *Juno* (Reitman, 2007). More specifically, the study revolves around one of the characterizing traits of teen talk, namely word formation through morphological processes (i.e. affixation, compounding, conversion, clipping, blending, acronym/initialism

and reduplication)⁸⁶, observing whether the peculiarities of teen talk are reproduced in Italian dubbing (*ibidem*: 54). Results show an increasing range of phenomena in English film speech over time. As a matter of fact, the two more recent movies, *Clueless* and, more notably, *Juno*, present more lexical innovation compared to the film from 1984. In the latter case, the language is rather standard, with only one attested instance of blending. Moreover, in *Sixteen Candles*, some of the word-formation phenomena attested in more recent movies are completely absent, as in the case of reduplicatives (*ibidem*: 67). In addition, the analysis pointed out a closer adherence of teen talk to spontaneous conversation over time. As a matter of fact, the most recent film analysed (*Juno*) presents more features of orality, like false starts, self-corrections, hesitations and discourse particles, than the two movies from previous decades. The orality features cited above are very few in *Clueless* and nearly absent in *Sixteen Candles* (*ibidem*). Therefore, changes in film speech can be noticed with dialogues being acted in a more genuine manner, i.e. closer to real-life conversation, by trying to reproduce as many features of spontaneous speech as possible (*ibidem*).

As for Italian dubbing, a growing attention to lexical innovation can be noticed over time; indeed, a wider range of translation strategies are used in dubbing with regard to word formation. More specifically, when morphological processes like compounding, blending, clipping or reduplication are used in the original dialogue, Italian dubbing commonly resorts to translation strategies such as normalization and explicitation with a variation in meaning (*ibidem*: 59), literal translation (this specifically occurs when translating blends) (*ibidem*: 64), reformulation with standard expressions (*ibidem*: 62-63) and explicitation through reformulation (as in the case of acronyms and initialisms) (*ibidem*: 64). This is usually due to the existing mismatch between English and Italian, meaning a lack of expressions corresponding to the original English ones in Italian. Nevertheless, an increase of attention to lexical innovation can be observed in the most recent film. In the Italian dubbed version of *Juno* there is an attempt to create new terms and constructions in line with the source text, for example, through the use of affixation (the prefix *stra-* with the function of a superlative

⁸⁶ Affixation involves creating new words by adding prefixes or suffixes to an existing base. Compounding involves the creation of a new word by combining two or more existing lexemes together. Conversion involves creating a new word by simply changing its word class without any formal change. Clipping involves creating a new word by deleting part of the lexeme and retaining only part of the stem. Blending involves the creation of a new word by combining the parts of two already existing words. Acronyms and initialisms involve using only the initial letter of the words of a phrase to form a new word; in initialisms the letters of the alphabet are pronounced separately, as a sequence of letters (DVD), while in acronyms they are pronounced as a regular word (NATO). Finally, reduplication involves the repetition of sounds, syllables, or words either totally (*boo-boo*) or partially with some vowel or consonant variation (*chit-chat*) (Jackson & Ze Amvela 2007: 81-103; Plag *et al.* 2007: 94-109; Mattiello 2013; Bonsignori 2015; Pulcini 2018: 183).

typical of colloquial speech) (*ibidem*: 57), creative compound words, colloquial terms and/or phrases (*ibidem*: 59, 61), or through the creation of new expressions (*ibidem*: 64, 66). All things considered, Bonsignori's (2015) study allowed the detection of two trends: film speech becoming closer to real-life conversation over time, on the one hand, and Italian dubbing becoming more creative and colloquial, on the other.

The first claim is also supported by Forchini (2013) in her investigation of familiarizers in original American films distributed between the 1970s and the 2000s. As for the language of dubbing instead, Forchini's investigation puts forward the hypothesis that, compared to the past, a growing attention is attested in dubbing as for the translation of the pragmatic meaning of certain linguistic elements (*ibidem*: 522). More precisely, Forchini carried out a corpus-based investigation of four familiarizers in original American films dubbed into Italian. In particular, the author considers the four most frequent familiarizers in both spontaneous speech and in TV series conversation (Quaglio 2009), namely *guys*, *man*, *buddy* and *dude* (Forchini 2013: 507). These familiarizers have been studied in six original American movies, three of them being released before the year 2000 (between 1974 and 1985) and the other three being released after the year 2000 (between 2001 and 2008)⁸⁷ (*ibidem*: 507). To begin with, a first quantitative analysis shows that film speech is increasingly more similar to spontaneous face-to-face conversation with regard to the vocatives observed in the analysis (*ibidem*: 513). After focusing on the functional nature of these familiarizers (observed through the analysis of collocations, colligations and lexical bundles), the author investigates whether the same functions are rendered in translation. The analysis of the dubbed component of the corpus reveals that familiarizers are usually omitted in dubbing, especially after the year 2000. This tendency to opt for a zero-translation strategy, however, contrasts with the increase in usage established in the English language (*ibidem*: 514). Omission occurs even in those cases where "there are appropriate functional equivalents between languages" (*ibidem*: 520), that is to say, when there would be an equivalent rendering in Italian fulfilling the same function of the original English expression. However, even though zero translation is the most used strategy to dub familiarizers, it should be noted that when these elements are indeed dubbed, the more recent translations mirror in a more precise way the function these features fulfil in the original language

⁸⁷ More specifically, the films *Young Frankenstein* (Brooks, 1974), *The Blues Brothers* (Landis, 1980), and *Back to the Future* (Zemeckis, 1985) constitute the before 2000 component of the corpus; *Shallow Hal* (Farrelly & Farrelly, 2001), *The Pursuit of Happyness* (Muccino, 2006) and *Twilight* (Hardwicke, 2008) constitute the after 2000 component of the corpus (Forchini 2013: 507).

compared to the translations of older movies (*ibidem*: 521). In other words, even if less frequent, more recent translations are qualitatively more functionally accurate compared to the way familiarizers used to be translated in earlier movies. According to the author, this could reflect “a rise in interest among translators, who seem to be paying more attention to pragmatic equivalence in more recent movies rather than in older ones” (*ibidem*: 522).

The discussions presented in the sections above have tried to showcase the usefulness of diachronic investigation for the exploration of film speech and dubbese. The review of studies presented here has allowed to pinpoint the evolution of Italian dubbing as more creative and colloquial (Bonsignori 2015), more absorbent of the English language (Motta 2015; Minutella 2021b), and more attentive to the translation of pragmatic meanings (Forchini 2013), while still showing some degree of (self-)censorship when it comes to informality being conveyed through expletives and taboo words (Formentelli & Ghia 2021). Nevertheless, as pointed out by Werner (2021), diachronic studies are still underrepresented, especially as regards the evolution of translational norms and patterns in AVT (Zanotti 2022: 449). As pointed out by Pavesi (2019b: 329), “The broader implementation of diachronic corpora would be an additional asset to research allowing analysts to probe challenging questions including how translational norms become established and how they evolve in time in different linguacultural settings and across spatial divides”. This study aims to do so with regard to orality features by examining a phenomenon hitherto insufficiently researched, i.e. hesitations.

The results outlined above illustrate a heterogeneous scenery with respect to Italian dubbing; in particular, the increased colloquiality observed in Italian translations seems to vary according to the phenomenon (i.e. expletives vs lexical innovation). It is thus difficult to formulate hypotheses regarding the diachronic evolution of hesitations in Italian dubbing. Drawing upon the previous research according to which Italian dubbing has become more absorbent of pragmatic meanings and more qualitatively accurate in achieving pragmatic equivalence (Forchini 2013; Minutella 2021b), an increase in the attention paid to the pragmatic markers studied here could be hypothesized. Nevertheless, considering the tendency of Italian dubbing to overlook said features (see §2.8), these elements are expected to be less present in dubbing compared to the original English text.

The methodology adopted for the diachronic investigation of hesitations in original Anglophone films dubbed into Italian will be provided in the third chapter of this dissertation.

Chapter Three

Corpus and Methodology

This chapter aims to describe the criteria and methodology used to build and examine the corpus of film speech developed for the analysis. Firstly, the corpus will be presented by focusing on the rationale followed for the selection of films. Subsequently, the methodology used to survey hesitations within the corpus will be described, explaining the way these features have been allocated to different positions within the turn, interpreted based on the Functional Framework developed for the investigation and examined from a translational perspective. Finally, insights are provided concerning software and tools used to carry out the inquiry.

3.1. The Analysis of Hesitations in Cinematic Dialogues

This study aims to diachronically investigate the occurrence of hesitations in original and dubbed film speech. As reviewed in Chapter One, studies on hesitations conducted through a diachronic perspective have mainly focused on written texts, examining the language of both fictional and non-fictional texts such as literary fiction, magazines, newspapers and non-fiction books (Jucker 2015a; Rühlemann & Hilpert 2017; Tottie 2017). Likewise, in the context of AVT, diachronic studies carried out through large-scale corpora also rely on written materials (i.e. subtitles) rather than verbatim transcriptions of filmic dialogue. These extensive databases built on subtitle files include, for instance, the *Movie Corpus* (Davies 2019), examined by Werner (2021) and Jucker & Landert (2023), and the *North American Movie Corpus* (NAMC), compiled by Veirano Pinto (2013, 2014). Research carried out on manually transcribed texts usually involves smaller corpora, like the ones compiled by Forchini (2012) and Zago (2016) in their analyses of film speech, which comprise 11 and 16 films, respectively. With respect to diachronic investigations of Italian dubbed film speech (see §2.9.2), scholars have mainly focused on specific phenomena in a limited number of titles distributed over an equally limited period of time: Forchini (2013), for example, investigated familiarizers in six films over the span of approximately 30 years (1970s-2000s); Bonsignori (2015) studied word-formation processes in three titles distributed over

a 23 years period (from 1984 to 2007); Minutella's (2021b) analysis revolves around Anglicisms and calques in animated movies spanning 27 years (1994-2019), while Formentelli & Ghia (2021) examined the *Pavia Corpus of Film Dialogue* (PCFD) to investigate expletives in movies produced over 22 years (1995-2017).

Based on these considerations, the present investigation aims to focus on a phenomenon hitherto insufficiently researched in AVT — especially as a category in its own terms — in a corpus of fictional dialogues conveyed to the audience through the spoken mode, either in the original or dubbed version. Drawing on previous research (Veirano Pinto 2014; Zago 2016), this is achieved through a diachronic corpus compiled ad hoc, covering a rather broad period (70 years) and including a number of titles (28 movies) that allows to obtain a diachronic overview of film dialogue within the limits of creating a corpus of verbatim transcriptions⁸⁸. This corpus will be studied by resorting to both quantitative and qualitative methods with a view to ascertaining whether recurrent patterns emerge from the way hesitations are employed in film speech and examining their function within the turn. In so doing, the present study seems to be in accordance with Rühlemann & Hilpert's (2017) considerations concerning the need to move toward an integrated approach between corpus linguistics and pragmatics: a burgeoning field referred to as “Corpus Pragmatics”, i.e. the rapprochement of the two abovementioned disciplines and “an integration of their key methodologies” (Rühlemann & Aijmer 2014: 23). As stated by Rühlemann & Hilpert (2017: 106):

Corpus linguistics typically scans texts vertically (Tognini-Bonelli 2010) looking for association patterns within narrow co-textual windows that can be described in quantitative terms, while pragmatic analysis typically processes texts horizontally, thus investigating, in much larger co-textual windows and more varied dimensions of context, functions of language that can be described in qualitative terms.

Bringing together these methodologies, the analysis of hesitations in film speech has been carried out by extracting frequency data from both original and dubbed dialogues over the years. At the same time, the quantitative analysis has been accompanied by a careful examination and interpretation of the phenomenon in context, that is, observing the way vocalizations are employed in the scene depicted on-screen and the functions they fulfil. This

⁸⁸ Further details on the limits of creating a large-scale corpus of verbatim transcriptions will be provided in §3.2.7.

was achieved first through a perceptual method, i.e. by manually annotating the corpus based on the way these elements are used in context; then, occurrences of hesitations in Anglophone film speech have been investigated through corpus-linguistic tools to compute their collocates and verify whether any recurrent pattern emerges through concordancing. Next, hesitations have been examined from a translational point of view to ascertain whether the phenomenon is reproduced in Italian dubbing and, more specifically, the type of strategies employed in the dubbed adaptations (see §3.7). By doing so, this study aims to shed some light on the evolution of film dialogue in both English originals and Italian dubbing with respect to orality features by focusing on a category of pragmatic markers still neglected in AVT research, i.e. hesitations.

3.1.1. Objectives and Research Questions

Although prefabricated orality has been a topic of research for a few decades now — the term “prefabricated orality” was first used by Chaume in the early 2000s (Chaume 2004b) — the phenomenon of hesitations is still quite overlooked with respect to the degree of similarity between film speech and impromptu conversation, as well as from a translational point of view. As reviewed in §2.8, in AVT, hesitations have been traditionally studied in co-occurrence with or among a wide variety of features of spontaneous language such as interjections and attention signals, as well as disfluencies like interruptions and repeats (Romero-Fresco 2007; Valdeón 2008; Bruti 2019; see also Zago’s 2016 multi-dimensional analysis on colloquiality)⁸⁹. Moreover, investigations conducted on these pragmatic features have so far revolved around case studies based on a small corpus of films, focusing on the stereotypical ways in which these elements are used as a shortcut to characterization (Jucker 2015b; Calogero 2024). Although these studies have shed some light on the way hesitations are strategically employed in film speech to achieve specific effects, they do not enable us to make generalizations on the occurrences of hesitation markers in film speech and the way they are treated in the adaptation process when dubbing from English into Italian. This investigation aims to fill this void. In particular, the objective of this study is to examine the degree of similarity between film speech and impromptu conversation, raising awareness on the role played by these often-overlooked features in pre-planned conversations, while also verifying whether there is an evolution in the way orality features are codified in film speech

⁸⁹ Zago’s (2016) investigation does not specifically focus on hesitation phenomena; nevertheless, in his multi-dimensional analysis, he considers a wide variety of features, among which also hesitations.

(RQ1-3). At the same time, this research aims to examine the phenomenon from a translational point of view, i.e. offering a picture of the procedures chosen to render hesitations in Italian dubbing over the years (RQ4-5). Most notably, this study aims to address the following research questions:

RQ1: What are the mechanisms of codification of orality in English-language film speech with respect to hesitation markers?

RQ2: What functions do hesitations fulfil within the mechanisms of interaction in the cinematic context?

RQ3: Does the way hesitations are used in film speech change over time? If so, does this provide some information regarding the evolution of film dialogue with respect to orality features? Does the genre of the film influence the frequency of occurrence of the phenomenon?

RQ4: How are hesitations treated in Italian dubbing? Can regularities in translation behaviour be observed?

RQ5: Does the way hesitations are rendered (or not rendered) in Italian dubbing change over time? If so, does this provide some information regarding the evolution of Italian dubbing with respect to orality features? Does the genre of the film influence the frequency of occurrence of the phenomenon in dubbed Italian?

To answer these questions, a collection of texts is gathered to empirically ascertain and describe systematic behaviours in original and translated audiovisual registers; more specifically, a parallel corpus of original English and Italian dubbed dialogues has been compiled. The rationale pursued to develop the corpus is presented hereinafter.

3.2. Building the Corpus: Steps and Criteria

As underlined by Baker *et al.* (2006), in building a corpus, decisions must be made concerning “the type, content, structure and size of the corpus” (*ibidem*: 55). In fact, “[t]hese decisions all help to ensure that the corpus fulfils its purpose, allowing the analyst to answer the types of questions set out earlier with confidence” (Gillings 2022: 594). In order to investigate the phenomenon of hesitations, a parallel corpus of English-Italian dialogues has been compiled. In particular, the database developed for the analysis corresponds to a translation-driven corpus, i.e. “a principled set of source language dialogues together with

their translations into a given target language” (Pavesi 2022: 549). The corpus comprises a collection of dialogues originally created in the English language and then dubbed into Italian; these texts were specifically selected with a double aim: to examine the way pragmatic features are treated in original Anglophone film speech and “to zoom in on choices made in distinct acts of translation” for dubbing (Bernardini 2022: 489).

In building the corpus, the first step was to circumscribe the sampling frame, i.e. to restrain what movies to consider for the study. For this purpose, the already existing *Roma Tre Film Corpus* was taken as a point of departure. The *Roma Tre Film Corpus* is a database developed at the University of Roma Tre that includes the orthographic transcriptions of titles distributed between 1930 and 2023 nominated for an Academy Award in the Best Picture category. First, the list of titles nominated for the above-mentioned category has been gathered to verify the extent of the database (i.e. *Roma Tre Film Corpus*). This was achieved by visiting the *Winners and Nominees* section of the *Academy of Motion Picture Arts and Sciences* webpage, where it is possible to browse Oscar winners and nominees since 1929, also filtering the search by award category⁹⁰. By limiting the investigation to Best Picture⁹¹, the list of nominees was created and compared with the list of movie transcriptions included in the *Roma Tre Film Corpus*. From this catalogue of 351 transcribed films, titles have been sampled based on a number of criteria: language, country, year of release, film genre and movie length, as well as some narrative features. Further details on said criteria are provided below.

3.2.1. Language

Since the analysis of hesitations focuses on a parallel corpus of English-Italian dialogues, the first condition that had to be respected was for a film to be originally produced in the English language with dialogues being uttered entirely, or for the vast majority, in English. In other words, movies where actors speak languages other than English were not taken into account — in *Inglourious Basterds* (2009), for example, dialogues involve various languages such as English, French, German, and Italian. Likewise, movies where characters

⁹⁰ This section of the *Academy of Motion Picture Arts and Sciences* webpage can be retrieved at the following link: <https://www.oscars.org/oscars/ceremonies/2025>.

⁹¹ It is important to remark that the name of this award category has changed over the years; thus, the research was carried out by making sure that, when filtering the category section, the correct name was being selected. More specifically, before the 1930s, the category was called “Outstanding Picture”; from 1931 to 1941, it was called “Outstanding Production” and “Outstanding Motion Picture” between 1941 and 1944, while from 1945, it was called “Best Motion Picture”; from 1963, it has been simply called “Best Picture”.

use a fictional or invented language were excluded — for example *Avatar* (2009), where characters speak the Na’vi language, or films from *The Lord of the Rings* (2001-2003) saga, where the elvish language is used. On the other hand, movies where a language other than English is used to a very limited extent have been considered — for instance, a few words uttered sporadically by secondary characters, who do not speak as much as the protagonists, like the Hispanic gardener in *The Kids Are All Right* (2010). As will be discussed in §3.2.6, this does not apply to characters using the Italian language in the source text.

3.2.2. Country

The selection of titles to include in the corpus was limited to North American productions. This decision was made to enable generalizations regarding the evolution of Anglophone film speech over time. More specifically, it was important to guarantee that the number of films produced in a selected country could remain stable throughout the entire period while also ensuring that, for the films produced in said country, the other criteria discussed here would also apply. Given the role of the United States in the international film landscape, the majority of titles nominated for an Academy Award in the Best Picture category are, indeed, USA productions⁹². Therefore, by restricting the corpus to films produced in the USA, it was possible to guarantee a wider range of material to choose from and consistency in the application of the selection criteria. In 2009, Díaz Cintas wrote:

The United States is the first largest exporter of audiovisual productions in the world dominating, not exclusively but mostly, the film industries of many other countries. [...] Since a large percentage of films and television programmes consumed by viewers worldwide are originally produced in the United States it seems legitimate to expect that they will exert a certain degree of influence both in the language — usually via translation — and in the attitudes of millions of people across the globe (Díaz Cintas 2009: 7-8).

Over the last ten or fifteen years, the situation has changed: even though the English language is still dominant, streaming platforms have widely modified the production market of audiovisual products with a growing number of films and TV series translated into English (Sánchez-Mompeán 2021, 2023a)⁹³. Nevertheless, the central role of American film

⁹² This information was retrieved through a search using the *Find* tool on Microsoft Word. This is the software used to make the list of films nominated for the Best Picture Academy Award. By conducting a query writing “USA” in the search bar, 533 results are obtained with regard to country of production as against 115 when searching “UK”.

⁹³ On the international increasing use of dubbing, see also Chaume (2019).

production and its impact on the international market over the years is unquestionable (see, among others, Rossi 2006)⁹⁴. As pointed out by Gennari (2009: 89), “[e]ver since the First World War, Europe has been one of the principal markets for Hollywood films. [...] After the Second World War, however, Italy became one of the largest foreign markets for the American film Industry”⁹⁵. In this respect, considering the diachronic perspective adopted in this investigation, it is methodologically significant to pay attention to what has influenced the history of Italian (dubbed) cinema, in spite of the most recent developments.

In addition, as noted in Chapter One (§1.6), language variation can influence the occurrence of hesitations (Clark & Fox Tree 2002; Tottie 2014b). Although it is not possible to overlimit the corpus — an American movie may involve actors speaking different English varieties (i.e. American, British, Australian, Indian, etc.) —, the inclusion of films produced in various Anglophone countries would establish an additional variable for the analysis. In other words, this would require investigating whether differences in hesitations occur among productions of different countries. While this is an interesting avenue for future developments, it goes beyond the scope of the present study. As a result, within the corpus, the sampling process has been limited to North American movies.

3.2.3. Year of Release

The corpus comprises films distributed over a time span of 70 years, starting from 1950. This criterion was established in consideration of the dubious availability of the original dubs produced during the decades spanning 1930 to 1940.

The development of dubbed cinema in Italy was undoubtedly linked to the Fascist regime, as dubbing represented a practice that allowed the domestication of foreign films, which were considered to be a threat to the Italian language, culture and mindset (Chaume 2012: 13). At first, dubbings were made by the American majors in the USA (Mereu Keating 2019a: 151-152); then, in the early 1930s, dubbing studios were relocated to Italy (Pavesi & Zanotti 2023: 176). By establishing dubbing studios in Rome, it was easier to monitor the adaptation process, that is, to verify that dubbed titles would comply with the censorship requirements imposed by the government (Mereu Keating 2016, 2019a). As pointed out by Mereu Keating (2016: 63):

⁹⁴ Besides cinema, authors like Motta (2015) have also underlined the central role played by Anglo-American dubbed TV serials in Italian television, from both the quantitative and the qualitative points of view.

⁹⁵ After the war, people wanted to be entertained; in addition, they wanted to catch up with what had happened in the world during the war, i.e. a period when connections had been limited. Hollywood major companies set out to do both; hence, the surge of American films (D'Amico 2008: 51).

If the majority of foreign-language talkies produced before mid-1932 could only be released in their silenced version, or in spoken Italian versions prepared abroad, and exclusively in those urban cinemas already wired with the requisite sound equipment, after that date Italian dubbings and versions made-abroad were not simply discouraged but prohibited by Mussolini's administration, in the ambivalent attempt to sustain the domestic film industry as well as to reinforce the nationalistic cultural and linguistic propaganda promoted by the regime.

Nevertheless, dubbings from the 1930s and the 1940s pose considerable challenges regarding access (O'Sullivan & Cornu 2019a: 24), as it is often difficult to establish the exact nature of dubbed films from this period and to retrieve original copies (see Pavesi & Zanotti 2023: 176-177). As pointed out by Cornu (2019a: 215), "[t]he lack of reliable information and viewing material has major implications for research". This is especially true for early dubbings whose commercial treatment poses concerns regarding "the provenance, identity, and reception of such films" (*ibidem*). In this respect, authors like O'Sullivan & Cornu (2019a) have underlined how "[t]he history of the early stages of AVT raises specific problems as researchers cannot just rely on the few films available on DVDs which often don't carry the dubbed or subtitled versions made for their initial theatrical release" (*ibidem*: 24)⁹⁶; they commonly carry a redubbed version.

Redubbing is a retranslation practice (Chaume 2007)⁹⁷ that has become quite common with the advent of DVDs. In the Italian context, Zanotti (2015: 112-113) highlights how DVD editions commonly present a dubbed soundtrack that is not the original one; most notably, commercial editions often "feature an Italian soundtrack that is different from the version that was shown on Italian screen" at the time of its release (Pavesi & Zanotti 2023: 177). This practice usually concerns older dubbings, a large number of which have been redubbed after the war (*ibidem*: 182); thereby, it can be difficult to track down and acquire their original dubs (*ibidem*). For these reasons, even though dubbing is a practice that developed in Italy since the 1930s (Mereu Keating 2019a: 155), the decision to leave out movies produced before the 1950s was made in order to guarantee the availability of original dubs. In other words, considering "[t]he difficulty of identifying and classifying retranslated films" (Zanotti 2018: 59), the risks of working with redubs were minimised by focusing on

⁹⁶ On the difficulties in distinguishing early dubbings from their redubs on DVD editions, see also O'Sullivan & Cornu (2019b).

⁹⁷ In the context of AVT, redubs constitute a form of retranslation, i.e. a second or succeeding translation of the same source text in the same target language (Chaume 2007: 50), attributable to a number of linguistic, historical and economic factors (*ibidem*: 56-61, 2018b: 18-23).

titles originally produced in the decades following the war. This was done to guarantee that the dubbed version analysed is the one contemporary Italian viewers enjoyed at the time the film was originally distributed in Italy.

On the other end of the chronological range, the corpus includes movies distributed until 2019. Even though the *Roma Tre Film Corpus* comprehends films spanning 1930 to 2023, the first three years of the 2020s did not include enough titles within the list of movie transcriptions for them to be considered in the diachronic investigation and respect all the selection criteria.

3.2.4. Film Genre

As claimed by Kozloff (2000: 26), “genre conventions have been a powerful force in shaping film dialogue”⁹⁸. In this regard, the decision of what genres to include in the analysis of hesitation has represented a central step in building the corpus.

Traditionally, corpus-based research on the language of cinematic dialogue has been carried out by resorting to a compilation of movies representative of spontaneous interactions; for example, the *Pavia Corpus of Film Dialogue* (PCFD) (Freddi & Pavesi 2009; Pavesi 2014; Zago 2019) includes “conversational films” (Freddi & Pavesi 2009), i.e. film genres that portray situations set in contemporary times and reproduce contemporary dialogues, hence real-life situations that trigger spontaneous-sounding spoken language. In particular, scholars like Pavesi (2005) have underlined the fact that the degree of naturalness of dialogue depends on the type of film and purposes of the cinematic production. Thereby, naturalistic dialogue would presumably be more relevant in certain film genres, such as romantic and psychological films, as opposed to, for example, musicals and auteur films. In the latter, the degree of spontaneity of dialogue is subordinated to other factors that play a more significant role (*ibidem*: 10).

Since this study focuses on film speech and prefabricated orality, similar considerations have also been made with respect to the genre criterion discussed here. More specifically, certain genres had to be excluded from the analysis: musicals, for example, were disregarded due to the fact that they feature less spoken dialogue since part of the narrative is conveyed through music in the form of songs. At the same time, since the aim of the

⁹⁸ The same has also been acknowledged with respect to TV dialogue (Bednarek 2010, 2012). Bednarek (2010, 2012), for instance, sees genre as an intermediate level between the register of fictional television and the linguistic profile of a specific TV series; that is to say that, in addition to having its own peculiarities, a given TV series will exhibit specific linguistic features associated with a specific genre as well as the more pervasive linguistic features of TV dialogue as a register.

investigation is to be as representative as possible of the film speech register, it would be important to sample films from more than one genre. For this purpose, previous studies conducted on film speech in relation to genre have been reviewed.

On the one hand, scholars like Veirano Pinto (2014) carried out an extensive investigation on 640 North American movies produced between 1930 and 2010, taking into account various situational variants, among which genre (including comedies, dramas, action/adventure, horror/suspense/mystery). Results suggest that “the language of action/adventure movies has more marks of planned speech than that of the other genres. This might be due to the fact that in many action/adventure movies the dialogue tends to present more integrated language to verbalize the events that serve a narrative purpose” (*ibidem*: 123). On the other hand, the importance of genre has been downplayed by Forchini (2012). The author compares 11 American films with the *Longman Spoken American Corpus* (LSAC) by means of Biber’s (1988) five dimensions of variation (involved vs informal production; narrative vs non-narrative concerns; explicit vs situation-dependent reference; overt expressions of persuasion and abstract vs non-abstract information). Forchini found that film dialogue and non-spontaneous conversation were remarkably similar in their being involved, non-narrative, situation-dependent and not highly persuasive; the only dissimilarity between the two was that film dialogue was found to be more abstract than spontaneous conversation (this finding was due to the higher frequency of conjuncts in films). This applied to her entire corpus, with no substantial genre-related differences. In particular, the author categorized the films in her corpus as “comedies”, “non-comedies” (included: drama, sci-fi, thriller, action, adventure, crime and fantasy movies) and borderline films (*ibidem*: 54). Comedies were found to be slightly more similar to spoken language than non-comedies based on the higher presence of conjuncts in the former. The difference, however, is minimal and it does not trigger any qualitative disparity (*ibidem*: 101).

Bringing together Forchini’s (2012) and Veirano Pinto’s (2014) findings, genres such as westerns, action and adventure films, as well as period dramas were excluded from the analysis; that is, less realistic genres were neglected so as to favour fictional products where dialogue is supposed to approximate more to real-life situations (Zago 2016: 77-78). In addition, in order to guarantee the largest possible amount of data, it was important to focus on genres that generally include more dialogue compared to others. According to the investigation carried out by *Polygraph* on 2,000 screenplays, movie scripts contain on

average 9,000 words⁹⁹; in particular, it seems that crime movies include the highest number of words, followed by comedies and drama films, whereas comedies contain the most words per minute of runtime, followed by crime and drama films (see Figure 3.1 retrieved from Pavlik 2017: online)¹⁰⁰.

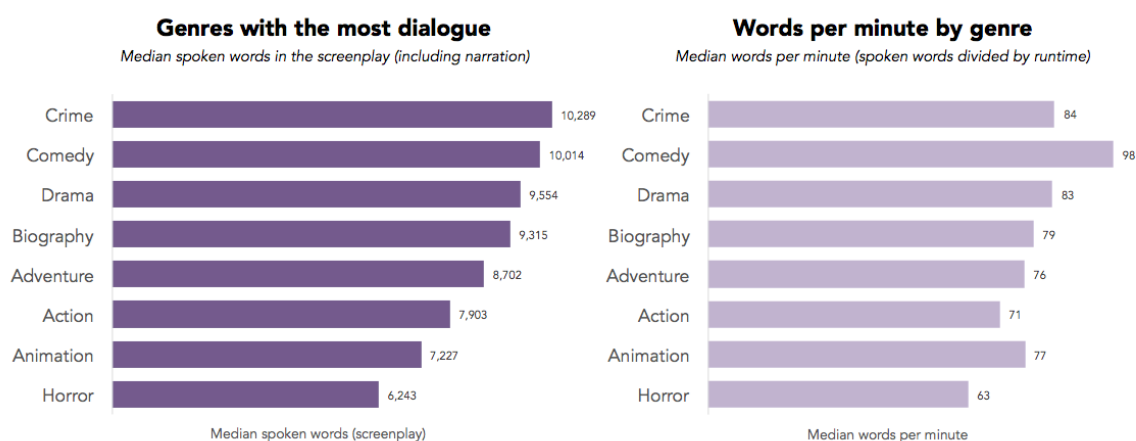


Figure 3.1. Spoken Words by Genre

As briefly mentioned above, in order to build a comprehensive and representative corpus of film speech, it was important to include more than one genre. More specifically, the first three genres identified as containing the most dialogue in Figure 3.1 were taken into account, namely crime, comedy and drama. As pointed out by Zago (2016: 82), “genres are fluid concepts whose delimitation is not always clear-cut”. Therefore, similarly to Forchini (2012), in the present analysis, films have been divided into two categories: comedies and non-comedies, where the latter includes mainly drama (and its subtypes) and crime films (and their subtypes).

These conditions were also established taking into account some of Zago’s (2016) considerations regarding the criteria to build his corpus of remakes for the investigation of colloquiality. Most notably, the decision to focus the analysis of hesitations on the genre categories mentioned above “was partly motivated by the programmatic exclusion of other less realistic genres” (*ibidem*: 80), such as western and action films. Moreover, the abovementioned genre categories offered films that would comply with the chronological

⁹⁹ Further information on the Polygraph’s Film Dialogue Dataset is available online at the following link: <https://github.com/matthewfdaniels/scripts/?ref=kaylinpavlik.com>.

¹⁰⁰ Image retrieved online at <https://www.kaylinpavlik.com/long-winded-actors-and-movies-with-the-most-dialogue/>.

requirement set for the analysis (*ibidem*: 80-81), i.e. a balanced number of comedies and non-comedies spanning over seven decades (1950s-2010s).

3.2.5. Movie Length

Movie length was taken into account based on Veirano Pinto's (2014) findings concerning the impact the duration of a film may have on the naturalness of the dialogue. More specifically, the author found that very long and very short movies present more features of planned speech compared to standard-length movies (i.e. up to 2 hours), which exhibit more features of spontaneous conversation (*ibidem*: 124).

In light of these findings, a minimum and a maximum duration were established for a film to be included in the corpus: movies had to be between 90 and 150 minutes long. In other words, films had to last at least one and a half hours; at the same time, movies longer than two and a half hours were disregarded. This restriction was set to favour the selection of movies where it is more likely to find features of spontaneous-sounding dialogue. Moreover, this condition was established in the interest of balance within the corpus; more specifically, a comedy film of three hours would presumably present more dialogue than a two-hour drama film, thus risking throwing off the equilibrium between the seven subsections of the corpus and the multiple criteria set for the analysis such as the combination of genre, dialogue, and the chronological limits, i.e. to balance the amount of dialogue contained in each decade.

3.2.6. Narrative Aspects

Some narrative and plot-specific factors were also considered in the movie sampling process. In particular, films depicting characters with speech disorders such as stuttering have not been taken into account since, in those cases, hesitations do not necessarily constitute pragmatic markers aimed at simulating spontaneity; they rather indicate a condition that causes speakers to have difficulty expressing themselves — for example, the character of Billy Bibbit in *One Flew Over the Cuckoo's Nest* (1975).

In addition, movies involving protagonists of Italian origins — Italian-American characters, such as Marty Piletti and his family in *Marty* (1955), Enrico “Rico” Rizzo in *Midnight Cowboy* (1969) or Loretta Castorini and her family in *Moonstruck* (1987) — have been excluded from the corpus. Since the collection of texts discussed here corresponds to a parallel corpus of English-Italian dialogues, this was to prevent the Italian language from

influencing the occurrence of hesitations in the source text. More specifically, it was necessary to avoid the risk of Italianness being rendered through an overemphasized characterization of Italian characters, for example, by resorting to dialect variations; as remarked in Chapter One, this could potentially impact the occurrence of hesitation phenomena both in the original English versions (Clark & Fox Tree 2002) and in the Italian dubbed adaptations (Giannini 2003a).

Finally, before concluding the current section, another aspect must be remarked: this study does not encompass animated movies. Animated movies have been examined by scholars with regard to their dubbing process (Sánchez Mompeán 2015), vocal performances (Montgomery 2017), humour (Minutella 2015), accents and dialects (Minutella 2016), linguistic variation and multilingualism (Minutella 2021a), specific features such as vocatives (Vignozzi 2021), and also from a diachronic perspective (Minutella 2021b). This investigation, however, aims to observe the phenomenon of hesitations in relation to (dubbing) actors' performances and technical factors such as the synchronization constraints imposed by the audiovisual medium. In animated movies, for instance, synchronization can be more flexible (Chaume 2012: 75-76), as opposed to non-animated products, where near-perfect lip synchrony constitutes a factor of primary importance for the success of the film (see Whitman-Linsen 1992). Moreover, animated films “also welcome exaggerated vocal performances, which would be out of place in most live-action productions, as a way to enhance the intended impressions of the characters” as well as to match the animation style (Crochiquia *et al.* 2024: 26). Therefore, non-animated films serve better for the purposes of the study, i.e. examining the evolution of orality in original English-language productions and in their Italian dubbed versions.

3.2.7. The Diachronic Corpus of Film Dialogue

Based on the abovementioned criteria, the final corpus consists of 28 American productions originally developed in the English language and then dubbed into Italian, released between 1950 and 2019, lasting between 1.5 and 2.5 hours (for a total screening time of approximately 54 hours — 3238 minutes — for each language), including genres where dialogue and spontaneous-sounding speech presumably have a central role. More specifically, the corpus comprises seven decades, four movies for each decade: two comedies and two non-comedies.

As briefly noted above, a pivotal aspect while building the corpus and setting the sampling criteria was for the database to be as representative as possible. It could be argued that 28 titles (four films per decade) are not enough to attain illustrative data concerning orality features in film speech as a whole. In this respect, it is undoubtable that in order to obtain generalizable results, it would be paramount to sample a significant number of films from various genres. Nevertheless, because of the limits of researching film speech through verbatim transcriptions, it is not possible to incorporate the same number of pictures considered, for example, by Veirano Pinto (2014), i.e. 20 movies per genre, for a total of 80 movies per decade. As reported by Zago (2021), large-size corpora of film discourse, like those compiled by Davies (2019) and Veirano Pinto (2014), usually consist of subtitle files rather than verbatim transcriptions of the soundtrack. In fact, even though word-for-word transcripts allow the transcriber to ensure that there is correspondence between the words in the corpus and the dialogues registered on the soundtrack, the manual transcription of the words actually uttered by actors while watching a movie is a laborious and time-consuming activity that can only be carried out for a limited number of films (Zago 2021: 172). The author claims that “the data contained in large filmic corpora are more representative but less “fine-grained” than those included in small filmic corpora, and vice versa” (*ibidem*). In particular, large-size corpora inevitably imply the user to be less familiar with the content of the corpus and a greater down-sampling of the countless results obtained from the searches. This is especially relevant in studies conducted on telecinematic language, where the inclusion of many titles makes it difficult to examine the linguistic features detected through corpus methods by confronting it with the actual video and soundtrack, a comparative process that is central for the investigation and correct interpretation of this particular type of discourse (*ibidem*: 170-172).

In this regard, Zago’s (2021) recent investigation provides insightful information concerning the stability of film discourse — i.e. film discourse represents a register characterized by regular and conventionalized patterns — through the comparison of studies conducted on film speech using both similar and different corpus methodologies on both a small- and a large-scale corpus (*ibidem*: 174). Results show similar data across the different corpora investigated; in particular, these affinities seem to suggest that film discourse has stable, repetitive and conventionalized lexico-grammar (*ibidem*: 179)¹⁰¹. Most notably, Zago claims that “film discourse is a stable register whose fundamental vectors of variation

¹⁰¹ In this respect, it is possible to conclude that film speech is indeed quite similar to the kind of conversation it tries to emulate, i.e. spontaneous conversation (Zago 2021: 179).

distinctly emerge when even a small number of observations (i.e., of films) are inspected” (*ibidem*: 175). Since the fundamental features of film discourse become noticeable even when investigating a more limited number of films, smaller-size corpora like the one compiled for this study can be considered useful and reliable tools for linguistic analysis (Zago 2021: 179).

In particular, in the present study, it was decided to prioritize accuracy in transcribing, annotating and interpreting the dialogue rather than expanding the number of films analysed. The level of accuracy and depth required for this investigation was achieved by means of manual audio analysis, although this inevitably limited the dataset. Tools such as AI-driven Automatic Speech Recognition (ASR) software would have indeed enabled the processing of a larger dataset; standard ASR systems, however, often filter out ‘noise’, including the very phenomena that constitute the core of this investigation. Given that hesitations constitute the central data of this research, manual annotation and the human ear remain the only reliable tools to transcribe all linguistic and paralinguistic elements relevant to the present study, interpret which of these features can be categorized as hesitations, and annotate them according to their location (§3.5) and function within the turn (§3.6). In other words, the advantages of a manual analysis were considered pivotal, despite the resulting limitations of the final dataset.

The final version of the corpus is reported in Table 3.1, which comprises 28 titles spanning seven decades (i.e. seven sub-sections), each of which includes four films: two comedies and two non-comedies. The information listed below has been retrieved from Morandini *et al.* (2024). Since genre categories can be ambiguous, in those cases where further details were necessary, the classification offered by the *Internet Movie Database* (IMDb)¹⁰² has been employed; this website, in fact, generally provides a wide list of tags to describe the genre of the film. Even though the majority of non-comedy movies include drama films, titles categorized as thrillers (and crime in the IMDb) are also included; hence, the more accurate “non-comedy” label has been used. As for the country of production, the films *Juno* (2007) and *Joker* (2019) present both the USA and Canada; Canada has been included among the macro-category of North American productions.

¹⁰² *Internet Movie Database* (IMDb) available online at <https://www.imdb.com/it/>.

Table 3.1. The Diachronic Corpus of Film Dialogue					
Year	Original Title	Italian Title	Country	Duration	Genre (Morandini)
1950	<i>Father of the Bride</i>	<i>Il padre della sposa</i>	USA	93'	Comedy
1955	<i>Mister Roberts</i>	<i>La nave matta di Mr. Roberts</i>	USA	123'	Comedy
1957	<i>Twelve Angry Men</i>	<i>La parola ai giurati</i>	USA	95'	Drama
1958	<i>Separate Tables</i>	<i>Tavole separate</i>	USA	99'	Drama
1960	<i>The Apartment</i>	<i>L'appartamento</i>	USA	125'	Dramatic Comedy ¹⁰³
1961	<i>The Hustler</i>	<i>Lo spaccone</i>	USA	135'	Drama
1967	<i>Guess Who's Coming to Dinner</i>	<i>Indovina chi viene a cena</i>	USA	108'	Comedy
1967	<i>The Graduate</i>	<i>Il laureato</i>	USA	106'	Drama
1974	<i>The Conversation</i>	<i>La conversazione</i>	USA	113'	Drama
1977	<i>Annie Hall</i>	<i>Io e Annie</i>	USA	94'	Comedy
1978	<i>Heaven Can Wait</i>	<i>Il paradiso può attendere</i>	USA	100'	Comedy
1979	<i>Kramer vs. Kramer</i>	<i>Kramer contro Kramer</i>	USA	104'	Drama
1982	<i>Tootsie</i>	<i>Tootsie</i>	USA	116'	Comedy
1982	<i>The Verdict</i>	<i>Il verdetto</i>	USA	129'	Drama
1988	<i>Mississippi Burning</i>	<i>Mississippi Burning - Le radici dell'odio</i>	USA	125'	Drama
1989	<i>Driving Miss Daisy</i>	<i>A spasso con Daisy</i>	USA	99'	Comedy
1991	<i>The Silence of the Lambs</i>	<i>Il silenzio degli innocenti</i>	USA	118'	Thriller
1994	<i>Forrest Gump</i>	<i>Forrest Gump</i>	USA	142'	Comedy
1996	<i>Jerry Maguire</i>	<i>Jerry Maguire</i>	USA	138'	Sentimental ¹⁰⁴
1999	<i>American Beauty</i>	<i>American Beauty</i>	USA	122'	Drama
2001	<i>In the Bedroom</i>	<i>In the Bedroom</i>	USA	131'	Drama
2004	<i>Sideways</i>	<i>Sideways - In viaggio con Jack</i>	USA	123'	Comedy
2007	<i>Michael Clayton</i>	<i>Michael Clayton</i>	USA	119'	Thriller
2007	<i>Juno</i>	<i>Juno</i>	USA/ Canada	96'	Comedy
2010	<i>The Kids Are All Right</i>	<i>I ragazzi stanno bene</i>	USA	106'	Comedy
2012	<i>Silver Linings Playbook</i>	<i>Il lato positivo</i>	USA	122'	Comedy
2016	<i>Manchester by the Sea</i>	<i>Manchester by the Sea</i>	USA	135'	Drama
2019	<i>Joker</i>	<i>Joker</i>	USA/ Canada	122'	Drama
				Total 3238'	

Having described the criteria established to sample films, the database has been compiled through a simplified prosodic transcription of the English and Italian soundtrack of each

¹⁰³ On IMDb, *The Apartment* (1960) is categorized as “Comedy”, “Satire”, “Romantic Comedy”, “Dark Comedy”, etc.; hence, in the current corpus, the film has been included in the comedy category.

¹⁰⁴ On IMDb, *Jerry Maguire* (1996) is categorized as “Romantic Comedy”, “Sport”, “Comedy”, “Romantic” and “Drama”; hence, considering the prevalence of tags associated with comedies and romantic comedies, the film has been included in the comedy category.

movie. Further details regarding the transcription process are provided henceforward. Nevertheless, before discussing transcription criteria, it is important to address issues concerning fair use and copyright of the material included in the corpus. Most notably, the data collected for this study are used for research and non-profit purposes. Drawing upon previous work created for the examination of film speech, such as the PCFD¹⁰⁵, the database only includes simplified prosodic transcriptions of dialogues, excluding the soundtrack and movie picture. Hence, the original cinematic product is only partially represented in the corpus; the corpus does not want to substitute for the original product and it does not prejudice its market. Moreover, when cited as illustrative examples, all films are cited and acknowledged (see the extracts below, as well as in Chapter Four and Chapter Five)¹⁰⁶.

3.3. Criteria for Soundtrack Transcription

As briefly mentioned in the previous section, the analysis of hesitations in the parallel English-Italian corpus has been carried out by looking at the simplified prosodic transcription of each title. A transcript can be described as the graphic representation of speech (Cameron 2001), “which thus implies diamesic variation from oral to written language” (Bonsignori & Filmer 2023: 76).

The *Roma Tre Film Corpus* comprises orthographic transcriptions of films; nevertheless, as pointed out by Bonsignori (2009b), orthographic transcriptions usually overlook information concerning background noises and paralinguistic elements, non-standard pronunciations, prosodic features related to intonation, stress and pausing, as well as “peripheral items” (*ibidem*: 194) like fillers, backchannels and hesitations. Considering the interests of this investigation, it is pivotal to include these elements in the examination of film speech. For this reason, the orthographic transcriptions of the *Roma Tre Film Corpus* have been revised with a double objective: to verify whether the words present in the transcribed text actually coincide with the words uttered in the soundtrack and to adapt the database to the needs of the study. In other words, it was essential to ascertain whether the transcriptions included the necessary information for the investigation of hesitations. Thereby, written texts have been adapted to a simplified prosodic transcription, where the term “simplified” is used to underline the fact that not all the prosodic information conveyed

¹⁰⁵ This information was retrieved online on the official page of the *Pavia Corpus of Film Dialogue*, available at the following link: https://allegatific.unipv.it/sibal/file/Pavia%20Corpus%20of%20Film%20Dialogue%20-%20official%20page_final%2008_04_21.pdf.

¹⁰⁶ See section 107 of the US Copyright Act (1976) and the Italian Law no. 633 Art. 70 (1bis 2008).

in the English and Italian soundtracks is reproduced; some elements have necessarily been neglected.

Scholars like Johnstone (2002: 21) have pointed out how “there is no single generally accepted way to represent speech on the page”; for that matter, the transcription procedure involves a complex process in which the researcher needs to carry out an interpretative act (see Lapadat & Lindsay 1999; Tilley 2003), depending on the specific purposes of the research project (Bonsignori 2009b: 185). Hence, it is the transcriber who determines which aspects need to be considered with more attention and how detailed the transcription needs to be (Tilley 2003; Bonsignori 2009b: 185; Bonsignori & Filmer 2023: 76). To put it in Bonsignori’s (2009b: 199) words, “[t]ranscription involves the re-elaboration of oral data: they are analysed, selected and interpreted by the transcriber/researcher, who decides which elements to represent and how, on the basis of the aims of the research project”. In the corpus examined here, for example, songs and background television noises have not been reported in the transcriptions.

More specifically, song lyrics have not been transcribed; if necessary for the understanding of the text, labels such as *((song playing))* and *((song continues playing))* have been used¹⁰⁷. Similarly, these descriptors are used when characters are singing; for example, at the end of the film *Juno* (2007), the label *((Juno and Bleeker play the guitar and sing a song))* is used. A different case is the one at the end of the movie *Joker* (2019), where the song *That’s Life* by Frank Sinatra starts playing and the character repeats the lyrics of the song; in that case, the descriptive label *((softly repeats the lyrics of the song))* is used and the words uttered by the protagonist are included in the transcription of both the English (central column) and the Italian (left column) version of the film (see Extract 1).

Extract 1 - *Joker* (2019)

Arthur	<i>((softly repeats the lyrics of the song))</i> that’s life .. and as funny as it may seem .. some people get their kicks . stomping on a dream . but I don’t let it let it get me down- - <i>((song continues))</i>	<i>((softly repeats the lyrics of the song))</i> that’s life .. and as funny as it may seem .. some people get their kicks . stomping on a dream . but I don’t let it let it get me down- - <i>((song continues))</i>

A similar case also concerns voices coming from the TV; more specifically, background voices coming from the television or the radio have been transcribed when clearly audible

¹⁰⁷ As will be explained below, the use of double round brackets to signal labels is in line with the transcription conventions described by Bonsignori (2009b).

and taking up space within the narrative. When these two conditions are not present, labels such as ((*tv chattering*)) have been employed. Along these lines, occurrences of hesitations uttered by background voices, for example, in TV programs or commercials, have not been considered in the quantitative and qualitative analysis. This also occurs in *Forrest Gump* (1994), where real historical footage is inserted in the film (i.e. dialogue that is not purposefully scripted for the movie), like George Wallace’s speech at the University of Alabama. On the other hand, lines coming from the TV have been carefully transcribed and considered for the analysis when they are uttered by characters as part of dialogues that contribute to the diegetic level, that is, to the unfolding of the events relevant for the story (i.e. not as background noises). For example, this is the case in the following extract from *Silence of the Lambs* (1991), where the television newscaster provides information about the victim of the serial killer (Buffalo Bill) that the protagonist, Clarice Starling, is trying to catch (Extract 2).

Extract 2 - *Silence of the Lambs* (1991)

Tim Langhorn on TV	and I’m Tim Langhorn, our top story for this morning . Catherine Martin, the twenty five year old daughter of senator Ruth Martin listed first as a missing person, is now believed to have been kidnapped by the serial killer known only as Buffalo Bill, Memphis police sources indicate that the missing girl’s blouse	e io sono Tim Langhorn, questa la notizia più importante della mattinata . Catherine Martin la figlia venticinquenne del senatore Ruth Martin, inclusa da prima tra le persone scomparse, si crede ora sia stata rapita dal pluriomicida noto solo come Buffalo Bill, fonti di polizia di Memphis indicano che la giacchetta della ragazza scomparsa

Since no standard guidelines exist for the transcription of oral data (both verbal and non-verbal codes) “but only reference models that can be varied and adjusted during the transcription process itself” (Bonsignori 2009b: 200), Bonsignori’s (2009b) conventions for the creation of the PCFD have been followed in this study¹⁰⁸.

As can be inferred from the examples above, each transcript is organized in three columns: the left-hand column includes information regarding the speaker, the conversational medium (i.e. phone calls) and narrative information, for example, whether off-screen narration is being used, i.e. “an extradiegetic offstage character narration

¹⁰⁸ A list of the symbols used in the prosodic transcription based on Bonsignori’s (2009b) conventions is also reported in the Appendices.

technique” (Spiteri Miggiani 2019: 114), also referred to as voice-over (V.O.); the central column comprises the transcript of the English soundtrack, while the right-hand column contains the transcript of the corresponding dubbed version (Bonsignori 2009b: 187).

If the speaker is a minor or an unknown character, he or she is referred to with the general term *Man/Woman* or with their profession such as *Lawyer* or *FBI agent* (Extract 3). If the Italian line is uttered by a different character from the one who said it in the English version, this information is also reported in the left-hand column (*ibidem*: 187-188).

Extract 3 - *Manchester by the Sea* (2016)

Lawyer	well but-- Joe has provided for Patrick's upkeep, food, clothes et cetera, and the house and the boat are owned outright	ma Joe-- ha provveduto al mantenimento di Patrick, cibo, abiti, e la casa e la barca sono sue chiaramente
Lee	I can't commute from Boston every day until he turns eighteen	non posso fare il pendolare fino a quando avrà diciott'anni

As for columns including dialogue, it is of pivotal importance to carry out a well-defined division of turns. This is achieved, for instance, by specifying if, within the same scene, the same character speaks to different interlocutors; when this occurs, the “=” symbol is used to signal latching (Bonsignori 2009b: 189). This symbol is used both at the beginning and at the end of the turn of the character switching interlocutors in succession. To avoid equivocation, the new interlocutor is also specified in the left-hand column by inserting *to* (*ibidem*), i.e. *Hoke to Boolie* in extract 4.

Extract 4 - *Driving Miss Daisy* (1989)

Daisy	yes thank you Hoke
Hoke	yessum =
Hoke to Boolie	= I'll be right with you mister Werthan

Both columns containing the English and the Italian soundtrack also include labels used to describe paralinguistic, kinetic and narrative information, for example, background noises and gestures that are relevant to the understanding of the verbal exchange (Extract 5). Following Bonsignori's (2009b) conventions, these labels are reported in double round brackets (*ibidem*: 189). Most notably, these descriptive labels are written in English in the transcription of both the Anglophone and the Italian soundtrack to make sure that the same number of words is used with regard to these labels in both versions.

Extract 5 - Forrest Gump (1994)

Doctor	all right Forrest, you can open your eyes now ... ((puts Forrest down)) let's take a little walk around	ecco fatto Forrest, ora puoi aprire gli occhi ... ((puts Forrest down)) facciamoci un giretto

Likewise, double round brackets are used to convey information about sounds such as sighs, as well as laughter or chuckles (Bonsignori 2009b: 190). If a portion of what the speaker says is uttered in a specific manner — i.e. laughing or whispering, as in extract 6 — the text in question is highlighted in the following way: the descriptive label is inserted before the section of speech involved; in addition, the plus symbol (+) is used at the beginning of said segment and the asterisk symbol (*) at the end (*ibidem*).

Extract 6 - Guess Who's Coming to Dinner (1967)

Joey	for a whole week I've been nervous	sono stata sulle spine tutta la settimana
John	no . ((chuckling)) +I don't believe it*	no . ((chuckling)) +non posso crederci*

The plus and the asterisk symbols are also used to mark when overlapping occurs, that is, two characters talking at the same time (Extract 7). In this case, the plus symbol signals the beginning of the overlapping and the asterisk signals the end (Bonsignori 2009b: 191).

Extract 7 - American Beauty (1999)

Brad	exactly .. +we're--*	esatto .. +perciò noi--*
Lester	+like* the time that mister Flournoy used the company Mastercard to pay for that hooker and ((chuckles)) she used the card numbers and stayed at the St Regis for what was it? like three months?	+quella volta* che il signor Flournoy usò la Mastercard della ditta per pagare una puttana e poi ((chuckles)) lei usa i numeri della carta e restò al St Regis per-per quanto è stato? per tre mesi?

Moreover, labels have also been used to structure transcriptions when changes of scenes or changes of shots occur (Extract 8). Oftentimes, these labels are used to signal the end of a previous segment and the start of a new one (Extract 9).

Extract 8 - Juno (2007)

Mark	I'm glad you're here I actually have something for you come here ... ((change of shot)) almost there	lieto di vederti sai ho una cosa per te vieni ... ((change of shot)) ci siamo quasi
Juno	whoa! Mark is this the baby's room? it's beautiful!	cavolo Mark è la camera del bambino? da panico!

Extract 9 - Annie Hall (1977)

Annie	were you always funny?	tu fai ridere sempre?
Alvy	hey what is this? an interview? we're supposed to be making love	ehi cos'è questa un'intervista? non siamo qui per fare l'amore?
	((change of scene))	((change of scene))
Comic's agent	this guy is naturally funny I think he can write for you	quest'uomo è un buffo congenito credo possa scrivere per te

As can be inferred from the elements described above, prosodic transcription does not follow conventional orthographic norms; namely, the traditional symbols of the written language are used to convey a variety of other meanings. In this regard, lowercase letters are generally used in the corpus with the exception of the first-person singular English pronoun *I* and the initial letter of proper names, countries, cities and workdays. In other cases, capital letters are used to convey prosodic information such as a character speaking with a loud voice (i.e. yelling) (Extract 10), or to convey emphatic intonation, i.e. stressing a specific word or a selection of words, like *bad* in extract 11 (Bonsignori 2009b: 194).

Extract 10 - The Graduate (1967)

Elaine	GET OUT! GET OUT! OUT! GET OUT!	VA VIA! VA VIA! VIA! VAI VIA!

Extract 11 - Silver Linings Playbook (2012)

Pat	Danny was in for assault because of crystal meth and alcohol	Danny era dentro per aggressione di anfetamine e alcol
Danny	BAD combination	pessima combinazione

As for other orthographic symbols, a single hyphen is used for repeats, i.e. the repetition of the same bit of language (a word, a syllable or a sound), “producing a momentary ‘stutter’ effect” (Biber *et al.* 1999: 1055) (Extract 12).

Extract 12 - Annie Hall (1977)

Alvy	+no* I'm sorry I can't do it we-we've blown it already I--you know I-I-I-- I can't go in in the middle	no mi dispiace ma non posso il-- cinema è già saltato io lo sai io non posso entrare a metà

Double hyphens are used for false starts and self-corrections (Extract 13) (Bonsignori 2009b: 190), as well as to signal interrupted speech (*ibidem*: 193), i.e. speech left unfinished, speech

interrupted by another character (Extract 14), or a sentence momentarily suspended and then continued or reformulated (Extract 15).

Extract 13 - *Tootsie* (1982)

Sandy	he-- will he be back?	uh ma tornerà?

Extract 14 - *The Kids Are All Right* (2010)

Joni	I don't know .. I just wish that you could've been--	io non lo so .. avrei voluto che tu fossi un po'--
Paul	what?	cosa?

Extract 15 - *Forrest Gump* (1994)

Forrest	there's only one thing I can say about-- . the war in Vietnam	c'è solo una cosa che posso dire sulla- - . guerra nel Vietnam

Since the hyphen symbol is used to convey the information described above, hyphenated compound words have been transcribed as open compounds, i.e. separated by a space. Likewise, numbers have been written as single words, omitting the hyphen (Extract 16) (Bonsignori 2009b: 199).

Extract 16 - *Twelve Angry Men* (1957)

Juror no.3	I run a messenger service ((sniffs)) the Beck and Call Company, the name is my wife's idea got thirty seven men working, started with nothing	io, ho un'agenzia di espressi per città, chiamate e rispondiamo, il nome è un'idea di mia moglie, ho trentasette dipendenti . e ho cominciato dal niente

In addition, other symbols have been used to represent prosodic features and specific sounds. In this regard, the “%” symbol is used to signal a glottal stop, that is, a glottal fricative plosive sound, which generates an expiratory breath from the larynx (Extract 17). As pointed out by Baños Piñero & Chaume (2009), glottal occlusions are among the specific features linked to phonetic articulation in Spanish dubbese; Chaume (2012: 89) also includes these elements among the main characteristic features, at a phonetic level, of a descriptive, not language-specific, analytical model for the study of dubbese. Moreover, the colon symbol (:) is used to signal prolonged sounds. In line with the International Phonetic Alphabet (IPA), this symbol is inserted after the lengthened sound, i.e. a sound of an unusual length (Extract 18).

Extract 17 - *The Apartment* (1960)

Miss Kubelik	well to make the figures come out even if I have no colds a year some poor slob must have five colds a year	beh% dato che io non ho mai il raffreddore qualche povero diavolo deve avere cinque raffreddori all'anno

Extract 18 - *Mister Roberts* (1955)

Doc	well I heard you were working cargo today so I thought I'd get ready . in days when there's only work to be done I can <i>a:lw</i> ays count on a big turnout at sick hall	ho saputo che hai giornata di carico e mi sono preparato . quando c'è da faticare l'equipaggio si dà appuntamento in infermeria, tutti malati

While carrying out the transcription, it was important to distinguish marked prolongations from unmarked ones. The former could be described as slightly withheld sounds, influenced by the contextual intonative outline; for example, a sound used to withhold a sentence momentarily suspended, but which will be continued after a momentary pause. The latter coincide with what Maclay & Osgood (1959: 24) defined as “non-phonemic lengthenings of phonemes”. In this case, the duration of segmental prolongation markedly exceeds its typical length in the same context of occurrence (see Gallo *et al.* 2024). In other words, it was crucial “to distinguish disfluent prolongations from other cases of lengthening, due to accentuation or in utterance-final or prepausal position” (Cataldo *et al.* 2019: 208). This was achieved through a “pragmatic approach” (Lickley 2015: 457) based on the subjective perceptive judgment of the transcriber linked to duration markedness of sounds; nevertheless, in cases of doubt, the speech analysis software Praat was employed for disambiguation purposes, as also in cases that required a more detailed investigation (see §3.8.1 for further details on Praat).

As for intonation, question marks and exclamation marks are used to signal variations from the declarative tone, that is, an interrogative tone in the former case and an exclamatory tone in the latter (Bonsignori 2009b: 193). Commas (,) are used to convey a simple intonation break, while full stops are used to signal actual pauses. The number of full stops indicates the length of the pause; most notably, one dot is used for a very short pause of a maximum of 1 second (.), two full stops are used for a medium pause of approximately 2 seconds (..), while three full stops are used to transcribe a long pause of 3 or more seconds (...). The length of the pause has been registered ‘manually’ by looking at the timecode while watching and transcribing the soundtrack. However, as will be further explained in §3.8.1, when

necessary, the exact duration of the silence has been measured using the speech analysis software Praat.

Importantly, non-standard pronunciations (*gonna*, *wanna*) and items such as fillers, interjections, filled pauses, backchannels (*mh*, *mm*, *mm-mh*, *uh-huh*, *hm*), discourse markers, invariant tags and contracted forms (*ain't*) are also replicated (*ibidem*: 193-195). In this regard, prosodic transcription is aimed at reproducing how words are actually uttered and pronounced by speakers; thereby, elements like acronyms and initialisms are distinguished. More specifically, when pronounced letter by letter, letters are separated by a space in the transcription (Extract 19) (*ibidem*: 199).

Extract 19 - *Silver Linings Playbook* (2012)

Danny	that's when I was x ray technician and I-I u:h had uh my A D D and my anxiety [...]	questo quando ero un tecnico radiologo e-e-e-e con la-la-la mia iperattività e la mia ansia [...]

In some cases, it has not been possible to fully hear and understand what a character says; when this occurs, subtitles constitute a helpful tool — especially in the case of closed captions, which usually include a more faithful rewriting of the spoken text — as do movie scripts available online. Nevertheless, when doubts persist, the missing item is expressed in the transcribed text by three capital Xs in round brackets (Bonsignori 2009b: 198), as can be observed in extract 20.

Extract 20 - *Driving Miss Daisy* (1989)

Hoke	what you talking about? I'm talking this woman had some lungs! she'd have been a whole church choir by herself if they'd let her . (XXX) . and fat too! woman had a but bigger than that stove over here	e come cantava! tu la dovevi sentire era uno spettacolo! aveva due polmoni che erano due mantici, avrebbe fatto il coro della chiesa da sola se la lasciavano fare, davvero! e poi la ciccia! era grossa pressappoco come quella stufa laggiù

By following the conventions described above, the English component of the corpus comprises 347,396 words, while the Italian corpus comprises 317,392 words — in line with the trend registered in dubbing, according to which the target text generally presents a lower number of words compared to the source text (Chiaro 2009; Bonsignori 2015). The distribution of the transcribed text in the seven sub-sections of the corpus (i.e. the seven decades) is reported in Table 3.2.

Table 3.2. Sub-Sections of the Diachronic Corpus of Film Dialogue		
Decade	Number of Words in the English Transcriptions	Number of Words in the Italian Transcriptions
1950s	55,071	47,890
1960s	48,631	43,404
1970s	50,770	46,909
1980s	49,875	46,010
1990s	46,832	44,709
2000s	45,274	42,035
2010s	50,943	46,435
Total	347,396	317,392

Having reviewed the general conventions employed in the prosodic transcriptions of the titles listed in Table 3.1, particular attention needs to be paid to the way hesitations have been transcribed.

3.3.1. Transcribing Hesitations

As remarked in Chapter One (§1.3), there is no universal agreement on how hesitations ought to be transcribed. Thus, similarly to Bonsignori (2009b: 195), different transcriptions have been used for hesitations in both English and Italian. In particular, the use of various transcriptions was aimed at carefully reproducing in written form the phonetic realization of vocalizations, namely whether a central schwa or another vowel sound was uttered. More precisely, vocalizations have been transcribed mainly resorting to American conventional forms such as *uh/um* and *uhm* (also in their prolonged version *u:h/u:m/u:hm*) (Extract 21); at the same time, the British transcription *ehm* has been used, also in the nasalized forms of *em* and *ahm*. As for the Italian language, hesitations were mainly rendered using the transcription *eh*, also in its nasalized versions *em/ehm* (Extract 21); nevertheless, *uh/uhm* have also been used in order to distinguish hesitations pronounced using a schwa sound ([ə]) from vocalizations produced uttering a half-open front vowel sound ([e]) (see extract 22).

In addition, there may also be circumstances where the sound uttered by the (dubbing) actor is not clearly distinguishable; when this occurs, descriptive labels are used, including *((indistinct sounds))* or *((indistinct hesitating sounds))* (Extract 22). However, it is important to remark that not all indistinct sounds have been categorized as hesitation markers; in this regard, it has been of pivotal importance to interpret what is heard based on the situation depicted on-screen. This is particularly relevant for the investigation conducted on the Italian soundtrack. As pointed out by Tissi (2000: 113), brief hesitations sounding like schwa need to be distinguished from *eh* sounds; in turn, the *eh* hesitating sound needs

to be distinguished from the interjection *eh*. On this matter, it should be noted that “a transcription itself does not constitute the data, but it is merely ‘a representation of the data’ or ‘a referential tool’ (Hutchby and Wooffiitt, 2008, p. 70). The actual data remain the film scenes and their respective film dialogues” (Chepinchikj & Thompson 2016: 43). In this sense, the analysis of the simplified prosodic transcription of each title in both the English and the dubbed Italian version has been carried out by examining the text as against the corresponding soundtrack and visuals. This comparative analysis has been carried out with the purpose of recognizing and isolating each occurrence of hesitation while also interpreting their role within the turn.

Extract 21 - *The Hustler* (1961)

Eddie	well hey uh I'll tell you what I'll play you jack up pool just keep one hand in my pocket	beh ehm em facciamo una bazzichetta e io gioco con una mano in tasca eh?

Extract 22 - *Sideways* (2004)

Jack	((indistinct hesitating sound)) I've just been doing some thinking about it	uh vedi-- d-diciamo che ci sto riflettendo un po'

Further details on how hesitations have been studied with respect to the co-textual and contextual environment are provided in the following paragraphs.

3.4. Investigating Hesitations in a Corpus of Film Dialogue

The soundtracks of the films included in the corpus have been transcribed on Word documents; during the transcription phase, each occurrence of hesitation has been annotated as such, using the <hes> label. More specifically, while watching and re-watching the film, instances of <hes> were manually identified and coded according to where they appeared in a turn (i.e. initially, medially, finally, or on their own). Moreover, each occurrence of hesitation has been qualitatively commented upon (adding comments to the document) with notes describing the way the phenomenon was used within the turn. Then, after the transcription process, each occurrence of hesitation has been isolated and organized in a database purposefully created for the study¹⁰⁹.

¹⁰⁹ The fact that hesitations have been labelled at different phases of the transcription process made it possible to qualitatively assess how these elements are employed in film speech in a more confident way. As a matter of fact, the transcription process has been carried out over the span of various weeks. After the transcription phase, hesitations have been categorized in Excel spreadsheets. The time span in between these two activities — carried out by comparing the transcribed text with both the visual and the acoustic levels — made it possible

Drawing inspiration from Baños Piñero’s (2013, 2014, 2024) analyses of orality markers in both domestic prefabricated Spanish and dubbed Spanish, the quantitative and qualitative inquiry of the corpus has been carried out through Microsoft Excel. More specifically, the analytical framework developed for the analysis (which will be described in the following paragraphs) has been classified in Excel spreadsheets. An Excel spreadsheet was created for each film, including every turn containing a hesitation marker or a prolonged sound. As can be seen in Figure 3.2, Excel spreadsheets were organized into 13 columns: Character, Gender, English Transcription, Timecode, Context, Position, Hesitation Tag, Function, Prolongations, Notes on Prolongation, Italian Version, Translation Strategy, and Italian Function.

Character	Gender	English Transcription	Timecode	Context	Position	Hesitation Tag	Function	Prolongations	Notes on Prolongation	Italian Version	Translation Strategy	Italian Function
Miles	M	oh fuck me! I know I said that I would be there by noon but uh there's all of this work going on at my building and it's just a total, uh nightmare!*, and I had a bunch of stuff I had to deal with this morning but, I'm on my way, I'm out the door right now	00:02:06	PHONE CALL	INTRA-TURN	PLAN	LIE_EXCUSE WHY HE'S LATE			ah! sono uno stronzo! . si si lo-l-lo-lo so lo detto che sarei stato da te a-a mezzogiorno ma eh stanno facendo dei-dei-dei lavori al tetto della casa e-e ti assicuro è un vero incubo e-e-e ho-ho dovuto sbrigare un-un ca-un casino di cose stamattina ma, sto-sto arrivando, sto uscendo di casa in questo momento	CLOSE RENDERING 1) SUBSTITUTION WITH INTERJECTION 2) CLOSE RENDERING	PLAN_LIE_EXCUSE WHY HE'S LATE 1) LAUNCH ANSW QUESTION 2) PLAN_INSTANT DECISION WHAT TO ORDER
Miles	M	uh no actually I'm running late make it to go and give me a New York Times and a uh spinach croissant, thank you	00:02:40		1) START-TURN 2) INTRA-TURN	FIRST - 1) LAUNCH ANSW QUESTION 2) PLAN	INSTANT DECISION WHAT TO ORDER			ah no sono in ritardo lo porto via e il New York Times per favore e un uh croissant agli spinaci, grazie	CLOSE RENDERING	OMISSION OF HESITATION
Miles	M	thank you very much is Jack uh still here?	00:04:00		INTRA-TURN	PLAN WITHIN ASK QUESTION				grazie mille, Jack è ancora qui?	OMISSION OF HESITATION	prolongation interjection rendered = exclamative emphasis
Jack	M		00:04:17					herey somebody finally arrived, you finally got here mister Prompt	prolongation interjection = exclamative emphasis hello	echi: ce l'hai fatta finalmente	prolongation interjection rendered = exclamative emphasis	

Figure 3.2. Example of an Excel Spreadsheet from Sideways (2004)

The *Character* column includes the name of the speaker uttering the turn, while in the *Gender* column, information regarding the gender of the speaker is provided; this is particularly useful to disambiguate characters when general terms such as professions are used to refer to the speaker (see Extract 3 in §3.3)¹¹⁰. The *English Transcription* section comprises the source text, hence the transcription of Anglophone turns, followed by information concerning the timecode, which indicates the moment of the film in which the phenomenon appears. Timecode information is specifically useful for the analysis since it allows to quickly retrieve and compare dialogue with the corresponding video, on the one

to test whether the way the phenomenon was originally interpreted during the transcription process corresponded to the interpretation of the same item after the transcription phase. If not, further examination was necessary to qualitatively evaluate the hesitation marker.

¹¹⁰ Information regarding the gender of the speaker also allows to analyse whether, in film speech, hesitations are more common among men or women. Although this goes beyond the scope of this investigation, it represents one of the possible avenues for future developments.

hand, and the English soundtrack with the Italian one, on the other. The *Context* section encompasses contextual details concerning the turn; for example, if the turn is uttered during a phone call and, if so, whether the interlocutor can be heard by viewers or not, or if the turn is uttered in the form of off-screen narration. The *Position* column is used to investigate the location of hesitations within the turn; for instance, if they occur at the beginning (“start-turn” tag), in the middle (“intra-turn” tag) or at the end of the turn (“end-turn” tag), as will be further examined in §3.4.1. The *Hesitation Tag* and the *Function* sections are used to summarize information concerning the functional role played by hesitations within the turn¹¹¹, along with notes on said function; for example, the fact that a character is lying in the first string of the table in Figure 3.2.

Since lengthenings are also considered in the present study and since they frequently occur in spoken Italian (see §1.9 in Chapter One), two columns have been inserted with respect to this phenomenon: *Prolongations* and *Notes on Prolongations*, the former including the transcription of the lengthened item and the latter including comments on the way prolongations are used within the turn. Thus, these sections enclose prolonged items with the aim to verify whether they are used as hesitation markers or not. Next, the transcription of the corresponding dubbed Italian turn is reported in the *Italian Version* section. This column is followed by the *Translation Strategy* one, where information regarding the strategies used to render, or not to render, hesitations in the adapted version is reported¹¹². If the hesitation marker is rendered in the adapted version, the last column is used to summarize its function in Italian dubbing, i.e. whether the original function is reproduced or a change of function occurs.

As briefly mentioned above, the data schematized in these spreadsheets has been utilized to conduct both quantitative and qualitative analyses. Most notably, the quantitative analysis has been carried out by resorting to the filters inserted in the spreadsheet, which become visible by clicking on the arrow on the right of the title of each column; for example, filtering the text by selecting only those strings containing the label “start-turn”, i.e. turns where hesitations occur at the beginning of the turn. As for the qualitative analysis, hesitations have been examined with respect to their location in the English turn and to their function in both the source and the target text. In this case as well, filters have been used to search keywords related to the functional categories (see §3.6) and the translation strategies identified (see §3.7).

¹¹¹ The Functional Framework will be presented in more detail in §3.6.

¹¹² Further information on the translation strategies identified is provided in §3.7.

Having discussed the way hesitations and the information conveyed through them have been recognized, isolated and classified, the criteria established to qualitatively analyse these features should be examined, i.e. the rationale for allocating hesitations within the turn, the function they fulfil in on-screen interactions and the correlation between the position of hesitations and their functional role.

3.5. The Location of Hesitations in Film Speech

In order to investigate where hesitations occur in film speech, each character's turn has been examined. In particular, four different locations have been identified, namely: turn-initial, turn-medial, turn-final and hesitations occurring as free-standing items. These four locations have been assigned a label to use in the Excel classification presented above: start-turn, intra-turn, end-turn, and single-turn, respectively. More precisely, hesitation markers have been allocated to one of these four categories based on the peculiarities of the turn in which the phenomenon occurred — for example, the length of the turn — and according to the situational context that accompanies said turn.

The turn-initial location does not merely include those cases in which hesitations occur as the very first item uttered by speakers; on the contrary, this category generally includes those cases where hesitations appear among the first four linguistic items uttered by a character. When the turn includes a line of a few words, the start-turn label usually comprises the first two items uttered, as is the case, for instance, in “yeah *um* just Tiffany” (*Silver Linings Playbook*, 2012)¹¹³. Similarly, in a sentence like “that was nobody justuh-- just Clay” (*The Kids Are All Right*, 2010), made up of six items, the hesitation *uh* cliticized at the end of *just* (i.e. the fourth item) corresponds to a turn-medial rather than a turn-initial hesitation, even though it takes place at the end of the fourth word uttered by the character of Laser. In other words, it was important to consider the specific characteristics of each turn. Thereby, if the turn analysed included a very short sentence or a limited number of items, a hesitation occurring as the third or the fourth item was more likely to be classified as a turn-medial rather than a turn-initial element.

Moreover, when the talking character starts a new sentence after a silent pause that separates it from the previous one, the hesitation marker uttered at the beginning of said new sentence is categorized as a “start-turn” since it is used to launch a new phase of the

¹¹³ In a sentence like the one from *Silver Linings Playbook* (2012), if the hesitation had been delivered after *just* in “yeah just Tiffany”, it would have been categorized as a turn-medial hesitation marker.

exchange. Again, it was central to interpret the context so as to distinguish when the silence constitutes a pause within the turn from when it marks the boundary of a new utterance being launched; the latter case is exemplified in extract 23 from *Silence of the Lambs* (1991).

Extract 23 - *Silence of the Lambs* (1991)

Crawford	I don't expect him to talk to you .. but I have to be able to say we tried . so if he won't cooperate I want just straight reporting, how's he look how's his cell look, is he sketching? drawing? if he is, what's he sketching? ... here's a uh dossier on Lecter .. copy of our questionnaire and a special ID for you, have your memo on my desk by o eight hundred Wednesday

After the first two to four items (depending on the length of the turn), hesitations are classified as turn-medial items. The turn-final category, instead, includes all those instances where hesitations occur as the very last element uttered by the speaker before suspending his or her speech or before being interrupted by another character taking the floor.

In this case as well, each occurrence has been carefully observed in context since it may happen that the hesitation occurring at the end of a turn in the transcription actually corresponds to a turn-medial item. This can be observed, for instance, in the following scene from *Juno* (2007), when Bleeker hesitates to hold the turn before directing his speech to a specific interlocutor (i.e. the character of Juno) (Extract 24). A similar case occurs in the event of overlapping turns, as depicted in the following extract from *Manchester by the Sea* (2010): even though Nurse Irene joins the conversation, George does not interrupt his speech; he continues talking undisturbed. In transcribing the text, the sentences uttered by George had to be divided in order to represent the exact moment in which Nurse Irene takes the floor; nevertheless, since George does not stop talking, the hesitation marker *uh* corresponds to a turn-medial element (Extract 25).

Extract 24 - *Juno* (2007)

Bleeker	well I'm gonna set up the <i>u:h</i> apparatus um =
Bleeker to Juno	= Juno do you want to plug in the bunsen burner?

Extract 25 - *Manchester by the Sea* (2016)

George	no I mean uh --
Nurse Irene	+George brought him in*
George	+I brought him in* <i>uh</i> we were-- looking at the boat this morning, and-- he just-- I don't know he-he just-- like fell over I-- I thought he was kidding me at first, then I called the ambulance and <i>uh</i> -- .. that was it

One more aspect that needs to be taken into account concerns the fact that, in the context of film dialogue, conversational turns may be influenced by cinematographic mechanisms such as a change of shot or a change of scene, which may interrupt a character's turn. This can be observed in the following extract from *Manchester by the Sea* (2010), where Mister Martinez probably intends to continue the sentence; the turn is, however, interrupted by a change of scene (Extract 26).

Extract 26 - <i>Manchester by the Sea</i> (2016)	
Mr. Martinez	((sighs)) tomorrow, my sister in law's coming over with the nephews .. I gotta take my car in um--
	((change of scene))

In this regard, the following extract from *Michael Clayton* (2007) represents an interesting case. In this scene, a number of cutaways show the character of Karen, a corporate lawyer, preparing for an interview. The scene shows Karen rehearsing her answers to the interviewer's questions (in the past) and the lawyer doing the actual interview (in the present). Once again, it can be assumed that the hesitation at the end of the turn is used with the intention to continue the turn; nevertheless, the sentence is stopped by a change of shot, switching from the present to the past (Extract 27).

Extract 27 - <i>Michael Clayton</i> (2007)	
Interviewer	((offscreen)) so with all that pressure and workload, how do you keep a balance between work and life?
Karen during the interview	((offscreen)) balance? ((onscreen)) . ((laughter)) I think that's some-- that's something that-- you search for, your whole life isn't it? um--
	((change of shot))
Karen at home	it's a shifting balance really it's <i>uh</i> you know you <i>uh</i> ((chuckles)) you try to--

Finally, hesitations can also occur on their own, meaning as free-standing items in conversation (see Norrick 2014). When this occurs, hesitations have been tagged using the "single-turn" label. As will be shown in Chapter Four, this position may also include those instances where hesitations occur as single items isolated from the rest of the turn, i.e. preceded and followed by a medium-long silence that separates the hesitation from the previous and subsequent text.

As can be inferred from the considerations made in this section, contextual information retrieved from the overall scene depicted on-screen is of pivotal importance for the interpretation of the phenomenon and its qualitative categorization; for this reason, it was necessary to resort to a manual audio analysis and annotation of the corpus. This has been especially relevant when investigating the functional role hesitations fulfil in film speech. The Functional Framework adopted for the evaluation of hesitations in prefabricated dialogues is provided below.

3.6. A Functional Analysis of Hesitations in Film Speech

The Functional Framework adopted for the qualitative interpretation of the way hesitations are employed in film speech has been developed based on the functional categories identified by previous research with regard to hesitations in both spontaneous conversation and fictional dialogue (see Chapter One).

As discussed in Chapter One (see §1.10), in this study, hesitations are conceived as communicative tools speakers resort to for different pragmatic, discursive and interpersonal purposes. Since cinematic dialogues are produced with the intention to recreate spontaneous-sounding conversations, hesitations have been examined by drawing upon notions from the fields of Discourse and Conversational Analysis. Indeed, as pointed out by Brown & Yule (1983: 1), “[t]he analysis of discourse is, necessarily, the analysis of language in use. As such, it cannot be restricted to the description of linguistic forms independent of the purposes or functions which these forms are designed to serve in human affairs”. Based on these premises, hesitations have been examined and interpreted in context, reviewing the specific circumstances in which language is produced. More specifically, the pragmatic employment of hesitations in cinematic discourse is studied in relation to the story, the specific situation depicted on-screen, the relationship between characters and means of communication (i.e. face-to-face conversation or telephonic conversation). In so doing, the notions developed by Austin (1962) and Searle (1969) on speech acts have also been taken into account, as well as Grice’s (1975) conversational maxims. As pointed out by Schiffrin (1987: 10), “a complete analysis of the meaning of a text would specify both the propositional meanings displayed by cohesive ties, and the inferences and understandings derived through application of contextual and pragmatic principles”.

However, as previously pointed out in Chapter One, the study of hesitations poses a number of difficulties ensuing from their fuzzy nature, their unclear status within the

linguistic system and the various definitions used to categorize and describe the phenomena included among the “hesitation” label. This lack of clarity also concerns the functional categorization of hesitations. As a matter of fact, scholars have adopted various frameworks and have identified a number of functions that may be labelled differently across studies or may include different phenomena. In other words, inconsistency is found in the way hesitations are functionally classified. For example, what Tonetti Tübben & Landert (2022) identify as a “cognitive function” corresponds to what Tottie (2014a) simply defines as “planning”; in turn, this “planning function” is referred to as “hesitation proper” by Kjellmer (2003). In addition, Tonetti Tübben & Landert (2022) do not distinguish between retrieving and processing information and include among the general “cognitive function” hesitations fulfilling a planning function, hesitations uttered when retrieving lexical items and/or specialised vocabulary, as well as hesitations preceding specific information such as names, specialised vocabulary or idiomatic expressions (*ibidem*: 358-360). Thus, this cognitive function partially covers what Kjellmer (2003) identified as “highlighting function”, namely “focusing the listener’s attention on an important, semantically heavy element in the delivery that is about to follow” (*ibidem*: 187) or alerting the listener about subsequent important points, implications or innuendos (*ibidem*: 188). The situation is further complicated by the multifunctional nature of the phenomenon under study, i.e. hesitations may simultaneously fulfil multiple functions.

By relying upon previous research conducted on the employment of hesitations in spontaneous conversations — including face-to-face conversations, interviews and telephone conversations — storytelling and fiction (Stenström 1990; Smith & Clark 1993; Kjellmer 2003; Fox Tree 2007; Corley & Stewart 2008; Schegloff 2010; Tottie 2014a; Jucker 2015b; Boonsuk *et al.* 2019; Staley & Jucker 2021; Tonetti Tübben & Landert 2022), a Functional Framework for the investigation of hesitations in film speech has been developed (Table 3.3). More specifically, five Functional Domains have been recognized: the Structural Domain, the Planning Domain, the Interpersonal Domain, the Cognitive Domain and the Hesitation Proper Domain. Each domain includes a number of more specific functions, which have been identified based on the reference literature, on the one hand, and on the examination of hesitations in the analysed corpus of film speech, on the other. The latter case involves, for example, the “Filler pause while carrying out some other actions” function included within the Planning Domain. As briefly mentioned above, the functions hesitations fulfil in spoken interactions often tend to intersect and overlap. For this reason, as can be observed in Table 3.3, the Structural and the Planning Domain have been grouped

together under the macro-category of Discursive Functions, highlighting the fact that functions associated with the management of the turn also partly fulfil a planning function (i.e. planning how to elaborate the sentence launched or held). The five Functional Domains summarized in Table 3.3 are further discussed below.

Table 3.3. The Functional Framework

DISCURSIVE FUNCTIONS	1. Structural Domain
	Turn-managing; turn-taking, turn-holding, turn-yielding.
	Launching speech acts.
	Change of speech act; change of topic.
	To report speech and to quote someone.
	2. Planning Domain
	Planning and thinking about what to say and how to say it, i.e. plan continuation of the discourse; plan (within) storytelling; plan explanations and addition of information; plan (within) a list of elements.
	Looking for the right word and for the correct way to phrase the utterance, i.e. a thinking pause for the selection of specific lexical items and for structural choices.
	Thinking about instant decision: decision-making process.
	Plan self-corrections: to correct and redraft.
	Delay: to buy some time to plan.
	Filler pause while carrying out some other actions.
	Highlighting: to draw attention to important elements; to highlight specific lexical or highly informative items.
	3. Interpersonal Domain
	Grab attention: to attract the attention of the character(s) involved in the interaction and/or to whom the turn is directed.
	Politeness and Face-Saving: - mitigation of invitations, offers, suggestions and recommendations; - mitigation of dispreferred turns, responses, sequences or actions.
	Requests.
	Bad-news announcement: mitigate the announcement of something to some extent problematic.
	4. Cognitive Domain
	Thinking process to carry out a cognitive demanding task; cognitive effort, i.e. counting, making calculations, recalling specific information such as names, dates or details.
	5. Hesitation Proper Domain
	Uncertainty, difficulty, production trouble, vagueness, doubt, awkwardness.

3.6.1. The Structural Domain

The Structural Domain concerns those functions associated with the allocation and organization of turns — two processes that are central in carrying out conversational interactions (see Sacks *et al.* 1974).

As pointed out by Clayman (2013: 151), “Turns are incrementally built out of a succession of turn-conversational units”. Tottie (2014a), however, remarks that “coding is not always obvious or easy” (*ibidem*: 389), especially since “turn-taking is not a precise and rule-governed undertaking” (*ibidem*: 401); much depends on the situation and on the circumscribing context. Since “turn-change usually takes place without explicit verbal selection of a next speaker”, it is important to determine “what the cues for the turn-change are” (*ibidem*: 382). In this regard, hesitations constitute interactional and social devices employed by speakers to organize the sequential deployment of turns, on the one hand, and the speech acts and the topics tackled within a single turn, on the other. In the former case, hesitations are employed by speakers to take the floor and initiate the conversational turn (turn-taking devices), to hold the turn so as to continue saying something (turn-holding devices) and to yield the floor to someone else (turn-yielding devices) (Stenström 1990; Kjellmer 2003; Tottie 2014a; Boonsuk *et al.* 2019: 137). In the latter case, hesitations are used to structure and organize both information and speech acts taking place within the turn; for example, they are used to move from greetings to the “reason for the call” in phone calls (Schegloff 2010: 146) and, more generally, from greetings to the “reason-for-initiating an episode of interaction” (*ibidem*: 166). In this respect, these elements can be used to signal different stages within the turn and mark topic transition (see Stenström 1990: 247); as a result, change of topic and change of speech acts are included in this functional category, as is the employment of hesitations to signal a shift from direct speech to the reporting of someone else’s speech or to the quotation of a saying.

In addition to turn-yielding, hesitations may also be used to signal projection of continuation (Clayman 2013) or as an (unintentional) symptom of the speaker’s intention to keep the conversation going (Tottie 2014a: 402). “Turn-initiators” (Stenström 1990: 227) can be used to launch specific speech acts or specific sequences such as launching a question in the question-answer adjacency pair (Schegloff 2010: 141), launching pre-sequences such as pre-invitations, pre-tellings, etc., as well as to initiate a new conversational topic. Similarly, turn-holding hesitations can be used to “indicate that the speaker is preparing a new informational unit” (Kjellmer 2003: 184).

3.6.2. The Planning Domain

This functional category mainly concerns processes of speech encoding. In this respect, the Planning Domain includes hesitations reflecting the organizational effort necessary in

communicative interactions (Chafe 1980; Hieke 1981). As remarked by Fischer (2000), hesitation markers can be conceived as discourse particles with the invariant meaning of “I am thinking about what to say next” (*ibidem*: 126), that is, as planning signals used to indicate that the speaker needs some time to process his or her ideas (see Fox Tree 2007). This commonly occurs at points where it is necessary to make decisions on the selection of the next step concerning either the content of the utterance or how to structure said content (Goldman-Eisler 1958: 66-67; Maclay & Osgood 1959: 41); for example, how to extend the body of the discourse, how to tell the events of a story (i.e. storytelling) or which elements to include in a list of items. Therefore, the planning process may correspond to the speaker taking some time to look “for a good way to phrase the upcoming message” (Tottie 2014a: 397), hence selecting the right word to use in the specific context in which the interaction occurs.

At the same time, hesitations may be used to vocally explicate the decision-making process carried out by speakers. In this regard, the speaker may hesitate to buy some time, i.e. momentarily delaying the continuation of the utterance in order to plan how to move forward. This can result in the speaker having to stop his or her speech to self-correct, reformulate or redraft. As pointed out by Kjellmer (2003) when discussing the “correction” function, this allows speakers to indicate “that a more correct or more suitable word or phrase than the one(s) just said will follow” (*ibidem*: 188). When the self-correction is launched by speakers themselves, unless a meaning of awkwardness and insecurity emerges from the context on-screen (see §3.6.5 on the Hesitation Proper Domain), hesitations are conceived as tools that help the encoding process, i.e. help the speaker select the best way to continue.

Since all codes involved in the cinematic product need to be taken into account (i.e. verbal, acoustic and visual codes), it is necessary to relate the way hesitations are used by characters with what is visible on-screen. More specifically, the functional category examined here also includes those cases in which speakers utter vocalizations as pause fillers to avoid silence while carrying out some other actions.

Finally, even though the Planning Domain mainly concerns the encoding process, some functions included in this category also relate to the listeners’ decoding process. Most notably, speakers hesitate to highlight specific items to the attention of the listeners. This corresponds to what Kjellmer (2003) identified as the “highlighting” function, which focuses “the listener’s attention on an important, semantically heavy element in the delivery that is about to follow” (*ibidem*: 187). Likewise, in this investigation, hesitations are employed by speakers to draw attention toward specific elements or to highlight highly informative items.

Indeed, as pointed out by Goldman-Eisler (1957), hesitations tend to co-occur with an increase of information; the hesitation uttered by the character “highlights the following element, suggests that it is being chosen circumspectly and focuses the listener’s attention on it” (Kjellmer 2003: 187)¹¹⁴.

3.6.3. The Interpersonal Domain

In this functional category, hesitations represent an interaction device speakers resort to so as to establish contact with the interlocutor, on the one hand, and make sure that the conversation turns out well, on the other. In the former case, Kjellmer (2003: 172) talks about “attention-catching”, that is, hesitations being used “to attract attention to oneself and to establish contact” (*ibidem*: 186). In the present study, this is referred to using the “grab attention” label. In the latter case, the success of the exchange is linked to the speaker’s ability to be polite, i.e. showing respect and consideration toward both the interlocutor and what he or she says. This is achieved through linguistic politeness (see Watts 2003).

Linguistic politeness is commonly associated with Face Theory, i.e. a theory of social interaction linked to Goffman’s (1967) concept of “face”, which correlates with “notions of being embarrassed or humiliated, or ‘losing face’” (Brown & Levinson 1987: 61). As pointed out by Brown & Levinson (1987: 61):

In general, people cooperate (and assume each other’s cooperation) in maintaining face in interaction, such cooperation being based on the mutual vulnerability of face. That is, normally everyone’s face depends on everyone else’s being maintained, and since people can be expected to defend their faces if threatened, and in defending their own to threaten others’ faces, it is in general in every participant’s best interest to maintain each others’ face [...].

This complex process carried out by speakers in interactions is called “facework” (Watts 2003: 130).

In this regard, a distinction can be made between “positive politeness” (pos-politeness) and “negative politeness” (neg-politeness). Neg-politeness involves mitigation “to reduce or lessen possible causes of offense” (Leech 2014: 11) and it is often linked to non-imposition (Brown & Levinson 1987: 62). Pos-politeness, instead, “gives or assigns some positive value to the addressee” (Leech 2014: 12), for example, through offers,

¹¹⁴ Further remarks on the “highlighting” function in the Functional Framework developed for the analysis are reported in §3.6.6.

invitations, compliments and congratulations (*ibidem*), but also suggestions, advice, orders and requests (Brown & Levinson 1987: 66; see also Tonetti Tübben & Landert 2022: 365 in relation to pos-politeness in fiction). These kinds of speech acts intrinsically risk threatening “the face wants of the addressee and/or of the speaker” (Brown & Levinson 1987: 65). Nevertheless, in the context of interpersonal relationships, the risk of causing damage can be reduced by using politeness strategies, among others, formulaic utterances like *please*, modal verbs (*could*, *would*, *should*), discourse markers like *alright*, vagueness, hedges like *just* and *maybe* and specific tenses to increase tentativeness (the past tense) (Watts 2003: 70; Leech 2014; Ryan & Granville 2019: 8-9; Jucker & Landert 2021: 129-121)¹¹⁵. More specifically, particular attention is given to linguistic politeness in two contexts: dispreferred sequences and bad-news announcements.

The term “dispreferred sequences” has been used by Schegloff (2010) in the context of adjacency pair-based sequences, where adjacency pairs are described as the basic, minimal form of turn-constitutional units, composed of two consecutive turns (i.e. adjacently placed) uttered by different speakers, “the first implementing, on its occurrence, a recognizable initiating action and the second, a recognizably responding one, with the two being recognizably from the same sequence type: offer and acceptance or rejection, request and granting or rejection, invitation and acceptance or declining [...]” (*ibidem*: 141) and so on. According to the author, the second turn taking place within the adjacency sequence can be either the preferred or the dispreferred one: the former “advances the course of action” (*ibidem*) launched by the first part of the adjacency pair, while the latter “do not advance the course of action, but block it or in some other way interfere with its realization” (*ibidem*). For example, second pair parts such as avoiding answering a question and rejecting or declining an offer, a request or an invitation constitute face-threatening acts (FTAs). As a result, when the response is unexpected or dispreferred, elements like vocalizations are inserted before proceeding with the actual turn (*ibidem*: 142); for this reason, these elements have also been referred to by Levinson (2013: 108) as “warning particles”. With respect to the functional domain discussed here, it is noteworthy that, as remarked by Schegloff (2010: 143), requests are commonly treated as dispreferred actions within the first pair part (unlike invitations, for instance).

¹¹⁵ As for politeness in film dialogue, Jucker & Landert (2021) specifically focus on expressions with polite default values (i.e. to express non-imposition) and impolite default values such as swearwords; on the other hand, Ryan & Granville (2019) specifically focus on invitation sequences and on the opening and closing of phone calls.

Moreover, in the context of linguistic politeness, bad-news announcements also constitute dispreferred sequences. Indeed, as pointed out by Schegloff (*ibidem*: 168), “although “announcements” as a class of actions are not dispreferred”, announcements of bad news represent FTAs that the speaker tends to soften in order to make the news more acceptable (Brown & Levinson 1987: 160, 171-172; Leech 2014: 119).

To summarize, along with the attention-catching function, the Interpersonal Domain also includes hesitations being used for politeness purposes, hence connoting the speaker’s reluctance or unwillingness to be tactless and to possibly hurt or insult the interlocutor (Tottie 2017: 124)¹¹⁶.

3.6.4. The Cognitive Domain

Drawing on observations made by previous scholars (see Goodwin 1987; Bortfeld *et al.* 2001: 126; Fox Tree 2007; Staley & Jucker 2021), the Cognitive Domain has been introduced within the Functional Framework to distinguish the thinking process related to the continuation of an utterance (i.e. planning how to phrase the subsequent section of the sentence) from the thinking effort executed by speakers when accomplishing a cognitively demanding task.

Therefore, these tasks do not coincide with the effort made by speakers while planning their speech; this functional category rather concerns hesitations as tools used to verbalize the character’s reasoning, his or her endeavour to remember something (for example, to recall specific information) or to accomplish a task that requires some mental effort such as counting and calculations.

3.6.5. The Hesitation Proper Domain

The term “hesitation proper” was first used by Kjellmer (2003) to describe speakers’ tendency to resort to filled pauses to convey uncertainty; according to the author, filled pauses give speakers enough time to think about what to say when they are faced with difficult topics, topics they lack the knowledge of, or when they are at a loss of appropriate words (*ibidem*: 182-183). In line with Kjellmer’s description, in this study the Hesitation Proper Domain includes the meanings stereotypically associated with hesitation markers,

¹¹⁶ In this study Tottie (2017) is specifically referring to hesitations in writing; since film speech has been described as a hybrid between speech and writing (see Chapter Two on this), her considerations on written language are relevant to the present research.

which coincide with the definitions commonly found in dictionaries, i.e. “The action of hesitating; a pausing or delaying in deciding or acting, due to irresolution; the condition of doubt in relation to action” (1.a) or “Embarrassed halting in utterance; stammering” (2.a) (Oxford University Press: online)¹¹⁷.

In line with this view, this functional domain includes hesitation markers being used as disfluencies in film speech, meaning as a device that adds no propositional content but simply interferes with the fluency of the message (Lounsbury 1954; Maclay & Osgood 1959; Fox Tree 1995). More specifically, characters make use of these elements as the result of trouble and nervousness (Fox Tree 2007: 306) attributable, for example, to lies and awkward moments. Based on the situation and the interpersonal exchange depicted on-screen, these features can also be regarded as signs of the speaker being uncomfortable, ill-prepared and lacking in confidence (Christenfeld 1995: 173).

3.6.6. Further Remarks on the Functional Categorization of Hesitations

As pointed out by Kjellmer (2003), hesitations can simultaneously fulfil multiple functions. Within the framework described above, it is not always easy to distinguish how functions interrelate. For example, the turn-taking function (Structural Domain) partly intersects with the function of grabbing the interlocutor’s attention (Interpersonal Domain); indeed, when speakers announce themselves and establish their intention to take the floor, they also attract the attention of the other participant(s) to whom said turn is directed. Therefore, it is important to pay attention to the contextual factors that accompany a turn to interpret the phenomenon and establish which one of the multiple functions involved is the most dominant in the on-screen interaction. In carrying out the qualitative analysis, the distinction between the two abovementioned functions has been executed in the following way: while in the grabbing-attention function, the speaker wants to obtain the consideration of the participant to make sure that he or she is paying attention to what is being said — hence, oriented toward the interlocutor to whom the turn is directed — in the turn-taking function, the aim of the speaker is to attract the attention on the self — thus, oriented toward the speaker taking the floor.

Furthermore, it should be noted that the functional ambiguity of the features studied here has oftentimes resulted in blurred lines between the domains described above and with

¹¹⁷ Oxford University Press. (n.d.). Hesitation, n., 1.a, 2.a. In *Oxford English Dictionary*. Retrieved November 12, 2025, from <https://doi.org/10.1093/OED/3545649217>.

respect to the micro-functions included in each domain. Kjellmer's (2003) "highlighting" function, for instance, has been a subject of profound reflection. In particular, while Kjellmer (2003) identifies this function as a self-standing category, other scholars like Staley & Jucker (2021) include it among discourse structuring devices; Tonetti Tübben & Landert (2022), instead, consider this use of hesitations as pertaining to the cognitive function. Tonetti Tübben & Landert (2022), however, do not differentiate between the retrieving and the processing of information in cognitive functions, as it is instead done in this study. In the Functional Framework discussed here, the retrieving of information involves a cognitive effort, i.e. pertaining to the Cognitive Domain described in §3.6.4, which differs from the processing required to encode a message (Planning Domain).

The speaker making an effort to retrieve information was central to determining how to categorize the highlighting function. More specifically, since the highlighting function does not involve the same mental effort manifested by speakers when retrieving information, this function could not be included within the Cognitive Domain. Nevertheless, the question remained on whether to include said function within the Planning or the Structural Domain. Staley & Jucker (2021) considered the functions of "highlighting interactive aspects of turn taking and new or semantically heavy content" (*ibidem*: 29) as part of discourse structuring. Staley & Jucker's study, however, observes hesitations in one specific communicative situation: service encounters in restaurants. In this context, the highlighting function appears to be particularly relevant in relation to food terminology when servers present the items included in the menu to their guests. In the present study, instead, the Functional Framework outlined in Table 3.3 must apply to the interpretation of various communicative situations, i.e. depending on the story, the type of relationship existing between characters, the genre of the film, and so on. Hence, a more complex distinction of the specific ways in which hesitations are employed in multiple communicative situations is required.

In this investigation, as for the "highlighting" micro-function, the speaker hesitates drawing attention to a certain item while giving the interlocutor time to fully assimilate the information highlighted. Since the speaker plans the next section of the utterance with the intention of underlining specific elements to the interlocutor, this function is categorized as part of the encoding process included in the Planning Domain. On the other hand, as summarized in Table 3.3, when hesitations are used with a word-search function — i.e. the speaker needs some time to recall information that he or she cannot immediately remember — they are classified within the Cognitive Domain.

Having explained the rationale behind the functional analysis of hesitations in film speech, the next section will focus on the criteria for the investigation of the phenomenon from a translational point of view.

3.7. Investigating Hesitations in a Corpus of Dubbed Film Dialogue: Translation Strategies

As noted by Valentini (2008), in the investigation of audiovisual translation strategies, it is necessary to inspect and compare occurrences of categories that can relate to verbal items, on the one hand, and “to the whole meaning that can be inferred from the surrounding non-verbal elements co-occurring in the filmic sequence” (*ibidem*: 45), on the other. In the investigation, this is accomplished through an English-Italian comparative analysis to assess if the pragmatic features examined here have been transposed into the target text and, if so, how. Drawing on CBTS — i.e. Corpus Linguistics applied to the branch of Descriptive Translation Studies (Pavesi 2019b: 315), the latter being a discipline aimed at describing “the phenomena of translating and translation(s) as they manifest themselves in the world of our experience” (Holmes 1972: 176) — the dubbed portion of the corpus is studied with the purpose of describing and explaining the regularities found in the translation of hesitations in the film speech register. More specifically, by observing the way the original hesitations are treated in Italian dubbing, four translation strategies have been identified: Close Rendering, Omission, Addition and Substitution. In other words, these are the translation strategies consistently found in the dubbed section of the corpus. The abovementioned categories have been identified by drawing upon the taxonomies found in translation studies (Nida 1964; Newmark 1988; Delabastita 1993; Vinay & Darbelnet 1995; Toury 1995; Munday 2009) and, more specifically, in AVT research (Bucaria 2007, 2008; Pedersen 2008, 2011; Chaume 2012; Formentelli & Ghia 2021).

Drawing upon Bucaria’s (2007, 2008) and Formentelli & Ghia’s (2021) research, the Close Rendering category consists of cases where instances of original hesitations are maintained and reproduced in Italian, sometimes adapting the rendering to the target system; nonetheless, even when adapted, the hesitation in the dubbed version is performed in a way that remains close to the source text. In other words, this category includes instances where specific elements — i.e. vocalizations — are “rendered in a similar way to the ST” (Bucaria 2008: 231). Most notably, this corresponds to a faithful rendering in a functional and context-specific sense; that is, it represents a rendering that closely reproduces the hesitation of the

source text with minimal changes attributable to the dubbing process and to the constraints underlying said process (i.e. synchronization).

The use of Omission as a legitimate translation strategy was described by Toury (1995) as “an a priori, prescriptive attitude”, meaning that “writers intent on the rights of the source text refuse to grant ‘zero’ replacements legitimacy as translation solutions” (*ibidem*: 82); thereby, the original element is deleted, “leaving no trace in the target text” (*ibidem*). In this study, the category of Omission comprises those cases where hesitations are not maintained or reproduced in the dubbed text; this elimination, however, does not significantly modify the remaining part of the original text. Drawing upon Formentelli & Ghia (2021: 53), this strategy also encompasses the deletion of hesitations and concurrent insertion of words, phrases or other phenomena that are not present in the source text or that are completely unrelated to it as a form of compensation. In the context of AVT, it is necessary to inquire whether complete omission is an arbitrary decision made by one of the agents involved in the dubbing process (see §2.5.1 on the professionals involved) or whether it is attributable to the synchronization constraints imposed by the medium itself (see Chaume 2012: 75). When relevant, this will be addressed in the analysis in §5.3.2.

As opposed to Omission, Addition involves the inclusion in the Italian version of hesitation markers that are not present in the corresponding source text. This strategy has been identified as a form of adjustment — a translation technique suggested by Nida (1964: 226) as a means to achieve equivalence — where “information that is not specified in the ST is inserted into the TT” (Munday 2009: 167). In AVT, this strategy may be employed to complement the deletion of the original hesitation (i.e. cases of over-extension, where hesitations are used instead of the features present in the original English text), but also as a means of achieving synchronization (Chaume 2012: 75).

Substitution is described by Delabastita (1993) as a process through which “the relevant S.T. item is replaced by a T. code item which is characterized by a more or less equivalent relational value” (*ibidem*: 33-34). In the context of AVT and, more specifically, in subtitling extralinguistic culture-bound references, Pedersen (2008, 2011) identifies Substitution as an interventional strategy — as opposed to “minimal change strategies” (Pedersen 2011: 101) — where the original item is “replaced with something totally different that fits the situation” (Pedersen 2008: 104); in most cases, this would lead to target-orientation (Pedersen 2011: 101). In the present analysis, this strategy includes those cases where original hesitations are replaced with different elements in dubbing such as words, phrases, sounds and paralinguistic phenomena.

Thus far, the description of methodological aspects and frameworks has outlined the rationale followed to ‘manually’ recognize, interpret and categorize hesitations in original English film speech and in Italian dubbed dialogues. This process, based on the self-interpretation of the phenomenon in context, is a non-trivial operation that, nevertheless, required the employment of a reference corpus (*Movie Corpus*) to corroborate findings, as well as specific software. The latter have been used to assist the speech analysis process (Praat), inspect potential recurrent patterns (AntConc) and statistically investigate the corpus (R/RStudio). Further information regarding the tools used in the analysis is provided below.

3.8. Methodological Tools: Additional Software and the Reference Corpus

In addition to a reference corpus used for comparative purposes, the empirical analysis of hesitations carried out in this study is underpinned by various software. The specific tools employed to carry out the investigation will be described in more detail hereinafter.

3.8.1. Praat 6.2.23

Praat is an online freeware speech analysis program created by Paul Boersma and David Weenink of the Institute of Phonetic Sciences at the University of Amsterdam. This software allows to inspect the spectral content of an existing sound through a spectrogram, namely a spectro-temporal representation of the sound (Boersma & Weenink 2025: online)¹¹⁸. In particular, Praat was used to measure and verify the duration of silences, to closely examine the prosodic features of specific sounds or sound segments (i.e. for disambiguation purposes) and to distinguish a prolonged sound from a slightly withheld one due to intonation.

More specifically, the software has been employed in the following way: the soundtrack of the movie (English or Italian) has been extracted from the film and downloaded in the MP3 version. Then, from the MP3 audio track of the whole movie, the specific segment that needed more in-depth analysis was cut, using the audio editing program Audacity, and imported on Praat (*Open > Read from file*). Once uploaded, the specific sound segment appears in the Objects window. However, “Praat can’t reliably tell where one word starts and where another ends. Nor can it find the specific segment you’re looking for, nor identify the vowel in the word” (Styler 2023: 53). Therefore, further analysis has been carried out by creating TextGrid annotations, i.e. automated annotations that mark

¹¹⁸ Boersma P., Weenink D. 2025. *Praat: Doing Phonetics by Computer*. Online at <http://www.praat.org/>.

intervals or specific points within the audio file (*ibidem*). When analysing specific segments of the soundtrack, these sections were cut to the specific portion that needed to be inspected; hence, what was necessary to do through Praat was to distinguish the sounding portion from the silent one. From the Objects window, this was achieved through the following steps: *Annotate > To TextGrid (silences)*. In so doing, two tiers appear: the sound that has to be annotated and the associated TextGrid (*ibidem*: 54). By selecting both and clicking on *View & Edit*, a window like the one in Figure 3.3 is displayed. This window shows the speech waveform, the TextGrid and the spectrogram, which includes a series of elements providing prosodic information such as red dots representing formants, blue lines representing pitch and a yellow line representing intensity (see Figure 3.3).

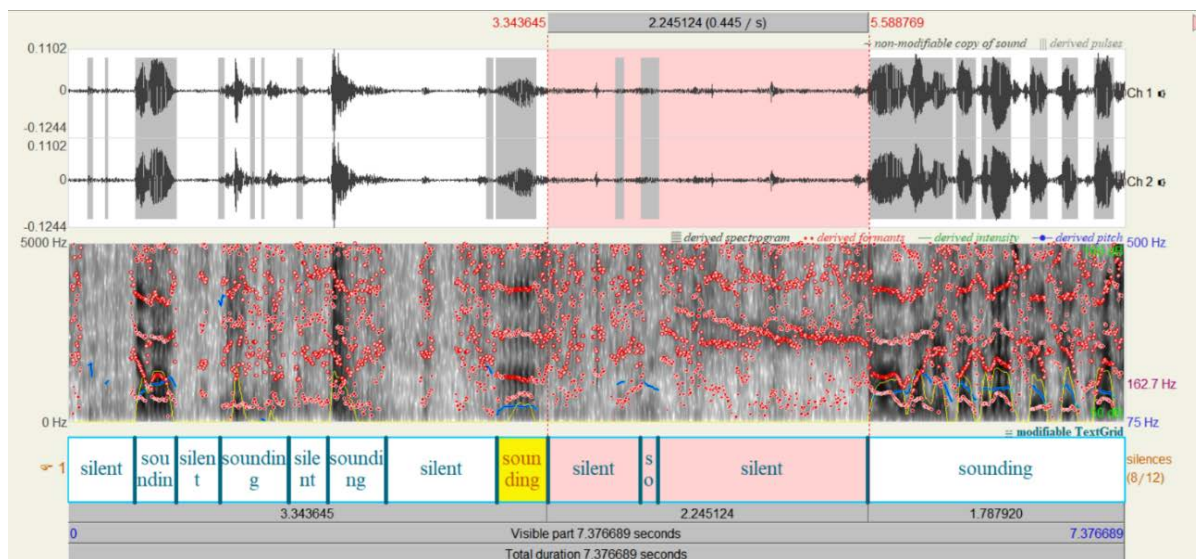


Figure 3.3. Example of a Praat Screenshot from Mississippi Burning (1988)

As briefly mentioned above, this software was used to measure sound segments and to disambiguate cases of doubt during the qualitative analysis and annotation process. As a matter of fact, this software makes it possible to listen to specific sections of a sound, split by the cursor or selection (for example, selecting one of the “sounding” segments visible in the spectrogram in Figure 3.3). When clicked, the selected section will play the corresponding part of the visible portions of the sound file. This allowed to closely examine the data and to carefully listen to the specific segment; for example, it was possible to verify whether what was heard was a background noise or a sound uttered by a character and if said sound corresponded to a hesitation. The information retrieved from the spectrogram was also used to verify the length of a sound, that is, whether a sound is prolonged (i.e. sound of an

unusual length) or it corresponds to an unmarked pronunciation influenced by the contextual outline (for example, a prepausal unit used to withhold a sentence).

In addition, Praat was used to measure the duration of silences. Most notably, by selecting with the cursor the portion of the audio that corresponds to silence, the duration (in seconds) of the corresponding section in red (see Figure 3.3 above) will be displayed either at the top of the selected segment or in the duration bar at the bottom, where the total duration of the segment under analysis is also reported. It is important to remark that Praat cannot distinguish dialogue from background noise or music; hence, the software cannot recognize segments where, even though background noises are present, characters are not speaking. Therefore, it was essential to manually identify the silent interval and select it with the mouse.

3.8.2. AntConc 4.0

As already discussed in Chapter Two, Corpus Linguistics has been recognized as a powerful methodology that relies upon corpus-linguistic software to research into the characteristic features, norms and patterns of audiovisual texts. In order to investigate whether regularities occur in the way hesitations are used in film speech, AntConc 4.0 (version 4.2.2) software was utilized, i.e. a freeware general-purpose offline corpus analysis toolkit for concordancing and text analysis (Anthony 2022: 107; Anthony 2025: online)¹¹⁹.

The analysis was carried out in the following way: first, transcriptions had to be modified and adapted to a machine-readable format. As pointed out by scholars (among others, Clark & Fox Tree 2002; Kjellmer 2003)¹²⁰, hesitations exhibit much variety in the way they are reported in the written form. As also mentioned above (see §3.3.1), in the corpus, hesitations have been transcribed using various conventional forms depending on the sound uttered by the (dubbing) actor and on the language being transcribed (Italian or English). In addition, labels such as *((indistinct sound))* or *((indistinct hesitating sound))* have also been used. Thereby, in order to conform these variants, the *hes* label has been employed. More specifically, new Excel spreadsheets were created; these files comprised dialogue transcriptions where each occurrence of hesitation was replaced with the annotation label *hes*. In addition, to make sure that all occurrences were recognized and detected by the software, those cases in which hesitations had been transcribed as cliticized to a previous

¹¹⁹ Anthony L. 2025. AntConc. In *Laurence Anthony's Website*. Retrieved online at <https://www.laurenceanthony.net/software/antconc/>.

¹²⁰ See Chapter One for a more detailed discussion on the ways hesitations are transcribed.

word had to be modified: the *hes* label had to be preceded and followed by a blank space. These Excel files (.xlsx format) were then converted into Text files (.txt format) and imported into AntConc 4.0 as a “Quick Corpus”.

Once all 28 files were imported, a first query was carried out to verify whether the total hits of *hes* corresponded to the total number of hesitations registered through the manual analysis of the corpus (i.e. 2001 tokens). Having confirmed that all tokens had been correctly annotated using the *hes* label, the corpus was examined for collocates and concordances counting. This was achieved by resorting to the Collocate Tool and the KWIC (Key-Word-In-Context) Tool. The former shows “words that appear frequently within a certain distance of the search term” (i.e. collocates), while the latter “allows you to see how words and phrases are commonly used in a corpus of texts” (Anthony 2023, AntConc version 4.2.2 help page). Further details on how these tools have been employed to verify and extend the findings obtained from the qualitative analysis of hesitation are provided in Chapter Four (§4.4).

As remarked by Anthony (2022: 121), “general-purpose corpus toolkits like AntConc provide many of the basic features for analysing corpora, but they do not provide advanced statistical measures”; for this reason, RStudio software was used for statistical analyses.

3.8.3. RStudio 25.09.01+401

In this investigation, statistical analyses have been computed to verify whether the frequencies of hesitations registered in film speech over the years are significantly influenced by factors such as the production year and film genre. Statistical tests have been carried out through RStudio software (Posit Team 2025)¹²¹, an integrated development environment for R (R Core Team 2025)¹²², which is a free software environment for statistical computing and graphics. More specifically, RStudio was used to carry out two statistical tests: Conditional Inference Trees (CTrees) and the Analysis of Variance (ANOVA).

The CTree method “is a non-parametric class of regression trees embedding tree-structured regression models into a well defined theory of conditional inference procedures. It is applicable to [...] multivariate response variables and arbitrary measurement scales of

¹²¹ Posit team. 2025. *RStudio: Integrated Development Environment for R*. Posit Software. PBC, Boston, MA. Online: <http://www.posit.co/>.

¹²² R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. Online: <https://www.R-project.org>.

covarieties” (Hothorn *et al.* 2015: 1). In other words, it allows the researcher to assess the relationship between a response variable and various predictors (Levshina 2021: 611). Most notably, this method “is based on testing the null hypothesis that the distribution of the response variable [...] is equal to the conditional distribution of the response variable given some predictor [...]” (*ibidem*: 615). The null hypothesis states that “there is no relation between the independent variable(s) and the dependent variable(s)” (Gries 2009:13) and it is tested by computing the probability value p (i.e. p -value). More specifically, as pointed out by Gries (2009: 30):

When that probability p is equal to or larger than 5%, then you stick to the null hypothesis because, so to speak, the result is still too compatible with the null hypothesis to reject it and accept the alternative hypothesis. If, however, that probability p is smaller than 5%, then you can reject the null hypothesis and adopt the alternative hypothesis.

In this study, the alternative hypothesis corresponds to: “the genre and the decade in which a film was produced/dubbed influence the frequency of occurrence of hesitations in original/dubbed film speech”, while the null hypothesis corresponds to: “the genre and the decade in which a film was produced/dubbed do not influence the frequency of occurrence of hesitations in original/dubbed film speech”. If the p -value is smaller than 0.05 (i.e. 5%), it is possible to reject the null hypothesis that the predictor variables (i.e. decade and genre) do not affect the frequency of the phenomenon; if the p -value is equal to or larger than 0.05, the null hypothesis cannot be rejected; hence, no significant difference occurs.

In testing the null hypothesis, the CTree method was used because of its simplicity and ease of interpretation since it is possible to inspect the results of the analysis by easily looking at a graphical representation of the tree. Most notably, having established the predictor variables, the algorithm determines “the best binary split in one selected covariate” (Hothorn *et al.* 2006: 655); in particular, it uses statistical independence tests to select the predictor variable with the strongest link to the response variable (*ibidem*), then finds “the optimal binary split in one selected covariate” (*ibidem*: 658). A split is made only if the test is statistically significant, that is, only if the global null hypothesis can be rejected at a certain level of statistical significance (i.e. p -value < 0.05) (see Levshina 2021: 616). If the hypothesis cannot be rejected — the significance level is not reached — the splitting process stops (see Hothorn *et al.* 2006: 655). Hence, if no predictor is significant, the node is not split.

Based on these premises, the analysis was carried out in the following way: data retrieved from both English and Italian dialogues were organized in Excel spreadsheets; most notably, the variables studied here were organized into four columns: Film (i.e. the 28 titles), Decade (i.e. the decade in which each film was produced), Genre (i.e. comedy or non-comedy) and FreqHes (i.e. normalized frequencies of hesitations) (see Figure 3.4).

1	Film	Decade	Genre	FreqHes
2	Father of the Bride	1950s	Comedy	4.88515598
3	Mister Roberts	1950s	Comedy	2.39310785
4	Twelve Angry Men	1950s	Non-Comedy	3.43160193
5	Separate Tables	1950s	Non-Comedy	2.66936672
6	The Apartment	1960s	Comedy	7.61962816
7	The Hustler	1960s	Non-Comedy	2.69163253
8	Guess Who's Coming to Dinner	1960s	Comedy	5.31837389
9	The Graduate	1960s	Non-Comedy	2.74690973

Figure 3.4. Example of Excel Spreadsheet for Analyses on RStudio (dataITA file)

The Excel files (.xlsx format) created for both English and Italian data were then converted into comma-separated files (.csv format). In order to access said data on RStudio, the comma-separated file (called “dataENG” and “dataITA”) must be saved locally in a directory on the researcher’s computer. Having done that, materials were opened on RStudio as data frames called “data” or “dati”, the former referring to the English data and the latter to the Italian ones. Namely, files were opened through the following codes:

- English film speech: `data <- read.csv2 ("dataENG.csv", stringsAsFactors = TRUE)`
- Dubbed film speech: `dati <- read.csv2 ("dataITA.csv", stringsAsFactors = TRUE)`

For the sake of brevity, the other codes written on RStudio and reported hereafter for exemplifying purposes will simply include the data frame called “data”; the statistical analysis for Italian dubbing was executed through the exact same codes by simply replacing *data* with the data frame corresponding to the Italian versions, i.e. *dati*.

Having uploaded the comma-separated file on the software, the next step was to make sure that the *FreqHes* column (i.e. normalized frequencies) was read by the program as a numerical value. This was achieved by resorting to the following code: “`data$FreqHes <- as.numeric (as.character (data$FreqHes))`”. At the moment of writing the code for the CTree, the necessary add-on packages were opened in RStudio, i.e. *party* and *partykit* (see Hothorn *et al.* 2006; Hothorn & Zeileis 2015). Subsequently, the *ctree()* function was employed to fit the Conditional Inference Tree; most notably, the line “`(FreqHes ~ Decade + Genre, data = data)`” was used to relate the dependent variable (i.e. *FreqHes*) to the predictable variables (i.e. *Decade* and *Genre*) within the selected data frame (i.e. *data*). Finally, the graph of the

CTree was plotted by using the *plot()* function. The steps explained here are summarized in the codes below, where *tree1* refers to the Conditional Inference Tree obtained for the English dialogues; for the analysis of Italian data, *tree2* has been written within the following codes:

- `tree1 <- ctree (FreqHes ~ Decade + Genre, data = data)`
- `plot (tree1)`

If the effects of the individual variables on the response are not statistically significant, no split occurs in the graph and the p-value is not reported. Nevertheless, even if not statistically significant, it is still important to visualize the p-value registered in order to have a better look at the relationship between the predictors and the response variable, as well as the relation between the predictor variables themselves (i.e. decade and genre). Exact values were obtained through the ANOVA test.

ANOVA “can be regarded as a special case of regression analysis” (Lavshina 2015: 171) that allows comparing more than two means or groups, assessing the effects of multiple independent variables at the same time (Rietveld & Hout 2005: 1). In particular, “you look both for main effects (i.e., each independent variable’s effect in isolation) and an interaction (i.e., each independent variable’s effects in the context of the other independent variable [...])” (Gries 2009: 284). In order to achieve this, categorical conversion is essential in ANOVA: a vector object must be converted to a factor. This was executed through the following lines, used to convert both predictor variables:

- `dati$Decade <- as.factor (dati$Decade)`
- `dati$Genre <- as.factor (dati$Genre)`

To compute ANOVA, the *aov()* function was utilized with the following formula for a two-way factorial design: “`anova_mod <- aov (FreqHes ~ Decade + Genre + Decade:Genre, data = data)`”. Then, the command “`summary(anova_mod)`” was used to obtain a statistical summary of the analysis, showing key outputs through a table organized in 5 columns — Df, Sum Sq, Mean Sq, F value and Pr(>F) — like the one reported in Figure 3.4 extracted from Gries (2009: 278).

```

Analysis of Variance Table
Response: DICE
.....Df...Sum Sq...Mean Sq...F value...Pr(>F)
PROCESS.....2..0.225784..0.112892...332.06<.2.2e-16***
Residuals..117..0.039777..0.000340
[...]

```

Figure 3.5. Example of ANOVA Table Retrieved from Gries (2009: 278)

The *Df* column reports the degrees of freedom for the independent variables and for the residual variance that remains unexplained. The *Sum Sq* column (i.e. sum of squares) includes the deviation sum of squares within and between groups (Rietveld & Hout 2005: 54; Gries 2009: 278); each sum of squares is divided by its associated degree of freedom (*Sum Sq*:*Df*) so as to obtain a variance estimate. The result of this operation is reported in the *Mean Sq* column (i.e. mean square) (Rietveld & Hout 2005: 56). The *F value* contains the statistical ratio used to compare two groups, while the *Pr(>F)* column lists the p-value (Gries 2009: 278). This last column is the one that provides the most relevant finding for the current analysis because, even if not statistically significant, it allows to verify how much the computed value differs from the significance threshold.

3.8.4. Using the *Movie Corpus* as a Reference Corpus

Traditionally, the term “reference corpus” has been employed to describe a corpus that “is not a sample of any particular language variety, domain or text type, but is instead an attempt to represent the general nature of the language through a wide-sampling corpus design” (Baker *et al.* 2006: 138). In translation studies, reference corpora have been described as a “tertium comparationis” (Baños Piñero *et al.* 2013: 484), that is, a corpus to which the focus corpus is compared. In order to carry out this comparison process, a suitable dataset has to be selected as the reference corpus, i.e. the two corpora must present similar parameters in terms of the type of texts and the languages involved; moreover, in the case of diachronic studies, the time span covered must also coincide.

In corpus-based research on film dialogue, reference corpora are employed as “representative corpora of the source or target language, functioning as baselines for comparison” (Pavesi 2022: 549); these databases are normally larger than the focus corpus since the aim is to verify whether the data obtained from the analysis are unusually frequent compared with the reference corpus (Anthony 2022: 105). This can also be established by using a subsection of these large databases as the reference corpus; for example, in the case

of diachronic studies, it is possible to focus on data concerning a specific decade (Mautner 2022: 256).

As described in Chapter One, it is difficult to compare corpora by focusing on pragmatic features like hesitations for various reasons such as the different methodologies and transcription principles adopted (Borfeld *et al.* 2001; Giannini 2003a; Jucker 2021). Nevertheless, some considerations made with regard to the data registered in this investigation required further corroboration by confronting results with the data obtained from similar and comparable databases. More specifically, when necessary, the English component of the corpus described here has been compared with the *Movie Corpus* (Davies 2019)¹²³.

The *Movie Corpus*, created by Mark Davies, contains 200 million words of data in more than 25,000 movies from 1930 to 2018 from different English-speaking countries. In particular, the corpus includes texts extracted from the Open Subtitles collection. This aspect is particularly relevant for the present investigation. Indeed, the use of the *Movie Corpus* as a reference corpus poses some limits related to the nature of the texts included in the database: while the focus corpus includes verbatim transcriptions of movie soundtracks, the *Movie Corpus* includes subtitle files, i.e. a written text that endeavours to convey the information contained in the soundtrack (Díaz Cintas & Remael 2007: 8). As noted by scholars, “the subtitler is limited to two lines of approximately 37 characters each for 5 to 6 seconds of speech. [...] This is why subtitling has developed a style of its own that has an impact on grammar and register, as well as on the interactional and other oral features of dialogue” (*ibidem*: 63). As a result, subtitles have traditionally corresponded to a reduced form of the original spoken text, i.e. a text that goes through condensation and the deletion of elements due to limitations of space and time (*ibidem*: 46). Among the elements commonly omitted from the written text, hesitations are also included (Bruti 2019: 197). Therefore, the information retrieved from the investigation of subtitle files does not necessarily correspond to the way certain phenomena are employed in the source spoken text. Nevertheless, as will be explained in §4.1.3, the data extracted from the *Movie Corpus* can yield useful information for the interpretation of the results; in particular, if some of the tendencies found in the analysis are mirrored in the *Movie Corpus*, they may indicate a trend specific to the film speech register. In the opposite case, the results obtained from the

¹²³ The *Movie Corpus* is available online at <https://www.english-corpora.org/movies/>.

analysis may be attributable to stylistic and/or narrative specificities of the films included in the investigation, rather than being ascribable to film speech as a whole.

It must be noted that, for the aims of the present study, the ideal reference corpus would include verbatim transcriptions of dialogues from movies produced (among other countries) in the United States of America in the same time span considered for this study (i.e. 1950-2019). However, general filmic corpora built on word-for-word transcriptions of dialogues are either not freely available online or not representative of the period selected for the diachronic analysis; that is, the films included in these databases do not cover all seven decades comprised in the analysed corpus¹²⁴. As pointed out by Zago (2021: 170), the *Movie Corpus* is “[t]he largest — and most recent — filmic corpus to date”, representing “an invaluable tool that is highly likely to become a standard reference point in the study of English language film discourse”. Thus, despite the differences between the types of texts included in the abovementioned databases, the *Movie Corpus* is a free, open-access resource that, unlike other corpora available online, covers the same time span considered in the current investigation and allows for a comparative analysis based on the same country of production¹²⁵.

Having described the criteria and methodology used to build and examine the parallel corpus developed for the diachronic analysis of film speech, results obtained from the investigation of hesitations in original Anglophone dialogue and in Italian dubbed dialogue will be reported, explained and discussed in the following chapters.

¹²⁴ The *American Movie Corpus* (Forchini & Palmero Aprosio 2025), for instance, includes orthographic transcriptions of dialogues from movies produced in the USA between 1959 and 2019; hence, unlike the present corpus, it starts from the end of the 1950s. Moreover, the corpus includes only one film from the 1950s (*Some Like it Hot*, 1959) and no movie from the 1960s.

¹²⁵ Since the corpus includes movies produced in various Anglophone countries, it is possible to limit the search to specific sections of the corpus, for example, by selecting the American component of the database.

Chapter Four

Hesitations in Anglo-American Film Dialogue: A Diachronic Analysis

The aim of this chapter is to describe the results obtained from the analysis carried out on hesitations in the English component of the corpus. Firstly, the chapter will focus on the discussion of the data obtained from the diachronic quantitative analysis; subsequently, the functions hesitations fulfil in original Anglophone film dialogue are reviewed through a qualitative examination. The data obtained from the functional description have also been cross-checked through corpus analysis using the AntConc software.

4.1. Hesitations in Film Speech over Time

4.1.1. Hesitation in Film Speech over Time: A Quantitative Analysis

The aim of the present study is to investigate the use of hesitations in film speech, as well as to ascertain whether the way these elements have been employed in cinematic dialogue over the years may provide some insights regarding the evolution of film speech and how orality is codified in the cinematic context. In order to do so, a first quantitative analysis has been carried out. The examination of hesitations has been accomplished by observing the phenomenon over the span of seventy years, considering movies from 1950 to 2019, for a total of 28 movies, namely four films per decade. The corpus, composed of the simplified prosodic transcription of the original soundtrack of each title, comprises a total of 347,396 words. Within the overall Anglophone corpus, 2001 tokens of hesitation markers have been detected, corresponding to 0.58% of the overall corpus.

Table 4.1. Tokens of Hesitations in the Anglophone Section of the Corpus		
Year of Release	Original Title	Tokens in the English Transcription
1950	<i>Father of the Bride</i>	100
1955	<i>Mister Roberts</i>	52
1957	<i>Twelve Angry Men</i>	94

1958	<i>Separate Tables</i>	34
1960	<i>The Apartment</i>	60
1961	<i>The Hustler</i>	41
1967	<i>Guess Who's Coming to Dinner</i>	89
1967	<i>The Graduate</i>	31
1974	<i>The Conversation</i>	70
1977	<i>Annie Hall</i>	132
1978	<i>Heaven Can Wait</i>	131
1979	<i>Kramer vs. Kramer</i>	104
1982	<i>Tootsie</i>	91
1982	<i>The Verdict</i>	129
1988	<i>Mississippi Burning</i>	27
1989	<i>Driving Miss Daisy</i>	20
1991	<i>The Silence of the Lambs</i>	56
1994	<i>Forrest Gump</i>	11
1996	<i>Jerry Maguire</i>	60
1999	<i>American Beauty</i>	19
2001	<i>In the Bedroom</i>	95
2004	<i>Sideways</i>	132
2007	<i>Michael Clayton</i>	53
2007	<i>Juno</i>	87
2010	<i>The Kids Are All Right</i>	130
2012	<i>Silver Linings Playbook</i>	30
2016	<i>Manchester by the Sea</i>	105
2019	<i>Joker</i>	18

As can be observed in Table 4.1, every film included in the analysis presents at least ten tokens of hesitation markers, with *Forrest Gump* (1994) being the title with fewer occurrences (11 tokens) and *Sideways* (2004) and *Annie Hall* (1977) the titles with the most occurrences (132 tokens). It is interesting to notice that in the titles where hesitations appear with a higher frequency, the phenomenon commonly approximates to 130 tokens; this is the case for the film *Annie Hall* (132 tokens), *The Verdict* (129 tokens), *Heaven Can Wait* (131), as well as *The Kids Are All Right* (130 tokens). This tendency seems to confirm that a certain amount of orality features is tolerated in film speech; however, if exceeded, they risk compromising the comprehensibility and the overall coherence of the dialogues, or they would simply take too much space within the narrative (Valdeón 2009: 206, 2011: 230; Bublitz 2017: 244; Jucker 2021: 356).

Since the corpus spans seven decades (from the 1950s to the 2010s) and each decade comprises four titles, grouping together the tokens of hesitations found in each of the four films released within a decade provides a first glimpse into the presence of the phenomenon in film speech over time. An initial observation of hesitations in the original English version of the 28 titles included in the analysis yields an important preliminary finding: within the time span observed, hesitations appear to have always been quite present in film speech.

Indeed, as reported in Table 4.2, hesitations are sufficiently represented in earlier movies as they are in more recent ones.

Table 4.2. Raw Frequencies of Hesitations per Decade						
1950s	1960s	1970s	1980s	1990s	2000s	2010s
280	221	437	267	146	367	283

Raw frequencies, however, do not allow for a straightforward comparison of data in order to describe the actual evolution of film dialogue over time, as the occurrences found in each movie must be considered with respect to the quantity of dialogue present in the film and/or in a specific decade. Since, as shown in Chapter Three (see Table 3.2), the various sub-sections of the corpus (i.e. the seven decades included in the study) differ in size, a more detailed quantitative analysis has been carried out by observing normalized frequencies, making it possible to compare results across the different components of the corpus¹²⁶. As shown in Table 4.3 below, raw frequencies have been normalized per 1,000 words¹²⁷.

Table 4.3. Normalized Frequencies of Hesitations per Decade (1,000 words)			
Decade	Total Number of Words	Raw Frequencies	Normalized Frequencies
1950s	55,071	280	5.084
1960s	48,631	221	4.544
1970s	50,770	437	8.607
1980s	49,875	267	5.353
1990s	46,832	146	3.118
2000s	45,274	367	8.106
2010s	50,943	283	5.555

First, it can be noticed that the use of hesitations in film speech fluctuates over time. These fluctuations, however, are not statistically significant. This has been attested using the Conditional Inference Tree analysis (CTree), a statistical test conducted through the open-source software RStudio. The analysis has allowed to test whether the predictor variable (i.e. decade) significantly contributes to the way normalized frequencies (i.e. the dependent variable) fluctuate over the years. When plotting the graph, the tree yielded a single node; hence, no splits emerged regarding time-related factors that significantly influence

¹²⁶ Normalized frequency is used to obtain the frequency of occurrence of a given item per a given number of words, for instance, per 100,000 words (Bednarek 2018: 96). The normalized frequency is calculated by dividing the raw frequencies of a linguistic item by the total number of tokens in a corpus, then multiplying the result of the division by the standardized value established, for instance, per 100,000 (Baker 2006: 86).

¹²⁷ For the sake of readability, the current chapter provides frequencies normalized to three digits after the decimal point. The exact occurrences normalized per 1,000 words can be found in the Appendices.

frequencies¹²⁸. Further support for this claim is offered by the statistical test ANOVA (Analysis of Variance), used to compute the p-value of the time variable examined here. Results show that the time variable presents a p-value of 0.268 (the statistical relevance threshold is $p < 0.05$). Thereby, the increases and decreases registered over the years could be interpreted as weak rising and falling trends that are not significantly and systematically influenced by the time variable (i.e. decade).

The findings show a slight decrease in occurrences of approximately 11% in the films from the 1960s, followed by a significant increase in the 1970s, where hesitations appear twice as much as in the previous decade. This seems to suggest that in older movies the presence of the phenomenon remains more or less stable for twenty years (with a slight diminution in the 1960s), until the sharp increase in the 1970s, which corresponds to the highest peak in the corpus. A decline occurs in the following decades, with a lessening of approximately 38% in the 1980s, reaching the lowest value in the 1990s (falling by 64% compared to the 1970s and 42% compared to the 1980s)¹²⁹. The situation changes in the films produced at the beginning of the new millennium: in the 2000s, the frequency of hesitation markers augments, reaching the second highest value after the 1970s. As can be observed in Figure 4.1, during the decade spanning 2000 to 2010, the presence of hesitations is roughly equivalent to what it was in the 1970s¹³⁰, with a reduction of just 6%. The phenomenon then finds a new balance in more recent films.

¹²⁸ The graph can be found in the Appendices.

¹²⁹ The percentage of reduction between decades has been obtained using the formula: $\frac{\text{previous value} - \text{current value}}{\text{previous value}} \times 100$, where current value refers to the more frequent decade taken into account as against a previous decade. The percentage value of reduction has been obtained by using the normalized frequencies (1,000 words).

¹³⁰ This is particularly significant considering the fact that the transcriptions of the movies included in the 2000s sub-section present the lowest number of words compared to the other decades taken into account (45,274 words). Consequently, this allows to re-establish the fact that the incidence of the phenomenon in a specific moment in time is not directly related to how much dialogue is contained in the text.

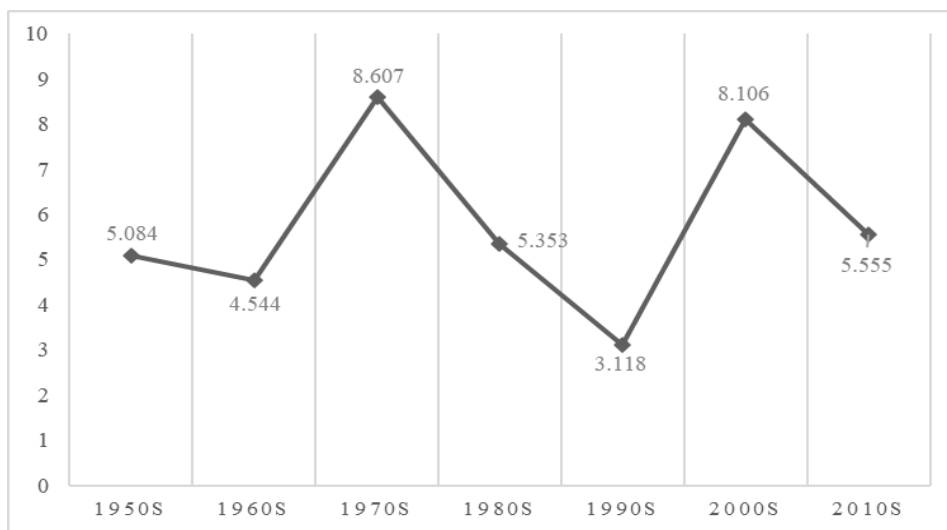


Figure 4.1. Normalized Frequencies of Hesitations in Film Speech over the Years (1,000 words)

As can be seen from the graph above, there is no clear evidence of an increase of the phenomenon over time. Specifically, the analysis of normalized frequencies between 1950 and 2019 does not suggest a linear increase. On the contrary, the presence of the phenomenon seems to oscillate over the years, with a positive peak in the 1970s (8.607 occurrences every 1,000 words) and then again in the 2000s (8.106 occurrences every 1,000 words), after a progressive decline that reached its lowest point in the 1990s (3.118 occurrences every 1,000 words). With the exception of these positive and negative instances — which, as highlighted by the tests described above, are not statistically significant — the level of usage of hesitations in film dialogues over the years seems to have remained relatively stable (ranging from 4.544 to 5.555), with a slight increase in the latest titles compared to the previous decades. Figures show that, from the 1950s to the 2010s, there is an increase of approximately 9% (from 5.084 to 5.555 per 1,000 words)¹³¹.

Despite these fluctuations, hesitations appear to be an integral and lasting aspect of film dialogue, a phenomenon that was already present in older productions.

The lack of a consistent trajectory of the phenomenon over the years may be a reflection of the dynamic manner in which hesitations are employed in film speech as a result of the influence of stylistic factors. One aspect that may impact on the way dialogues are written and performed is film genre. In this respect, it is important to consider whether a correlation exists between the incidence of occurrence of hesitations in film speech and the genre variable.

¹³¹ The percentage of increase has been obtained using the following formula: $\frac{\text{current value} - \text{previous value}}{\text{previous value}} \times 100$.

4.1.2. Hesitations and the Genre Variable: A Diachronic Analysis

In this section, the genre variable is studied in order to ascertain whether film genre impacts on the occurrence of hesitations in the dialogue and, if so, whether changes can be observed over the years. Various scholars have demonstrated the impact genre may have on the colloquiality of a film (Zago 2016) and how movie genre influences the degree of naturalness and spontaneity of the dialogue (Forchini 2012; Veirano Pinto 2014). In the current study, the aim is to verify whether a correlation exists between the frequency of occurrence of hesitations and the genre of the film (comedies vs non-comedies). This is achieved by calculating the normalized frequency (per 1,000 words) of occurrence of hesitations in each title (see Table 4.4). Subsequently, by grouping together the normalized frequencies of hesitations in both comedy and non-comedy films, the mean score between the two genre categories is calculated. In so doing, movie length (i.e. the number of words contained in each title) will not impact on the results. Having obtained the mean score of both genre categories for each decade, the phenomenon trend has been studied by looking at the percentage variation between one decade and the next one¹³².

Table 4.4. Normalized Frequencies of Hesitations in the Corpus (1,000 words)					
Year	Film	Genre	Number of Words	Hesitation Tokens	Normalized Frequencies
1950	<i>Father of the Bride</i>	Comedy	13,409	100	7.458
1955	<i>Mister Roberts</i>	Comedy	14,186	52	3.666
1957	<i>Twelve Angry Men</i>	Non-Comedy	15,189	94	6.189
1958	<i>Separate Tables</i>	Non-Comedy	12,287	34	2.767
1960	<i>The Apartment</i>	Comedy	14,634	60	4.100
1961	<i>The Hustler</i>	Non-Comedy	10,040	41	4.084
1967	<i>Guess Who's Coming to Dinner</i>	Comedy	14,826	89	6.003
1967	<i>The Graduate</i>	Non-Comedy	9,131	31	3.395
1974	<i>The Conversation</i>	Non-Comedy	8,835	70	7.923
1977	<i>Annie Hall</i>	Comedy	16,210	132	8.143
1978	<i>Heaven Can Wait</i>	Comedy	12,731	131	10.290
1979	<i>Kramer vs. Kramer</i>	Non-Comedy	12,994	104	8.004
1982	<i>Tootsie</i>	Comedy	17,387	91	5.234
1982	<i>The Verdict</i>	Non-Comedy	12,357	129	10.439
1988	<i>Mississippi Burning</i>	Non-Comedy	10,433	27	2.588
1989	<i>Driving Miss Daisy</i>	Comedy	9,698	20	2.060
1991	<i>The Silence of the Lambs</i>	Non-Comedy	9,445	56	5.929
1994	<i>Forrest Gump</i>	Comedy	12,014	11	0.916

¹³² The percentage variation is obtained by using the following equation: $\frac{\text{more recent mean score} - \text{previous mean score}}{\text{previous mean score}} \times 100$.

1996	<i>Jerry Maguire</i>	Comedy	16,008	60	3.748
1999	<i>American Beauty</i>	Non-Comedy	9,365	19	2.029
2001	<i>In the Bedroom</i>	Non-Comedy	9,328	95	10.184
2004	<i>Sideways</i>	Comedy	12,600	132	10.476
2007	<i>Michael Clayton</i>	Non-Comedy	13,060	53	4.058
2007	<i>Juno</i>	Comedy	10,286	87	8.458
2010	<i>The Kids Are All Right</i>	Comedy	11,035	130	11.781
2012	<i>Silver Linings Playbook</i>	Comedy	17,370	30	1.727
2016	<i>Manchester by the Sea</i>	Non-Comedy	13,992	105	7.504
2019	<i>Joker</i>	Non-Comedy	8,546	18	2.106

The mean score obtained from the analysis of the standardized frequencies of hesitations based on genre (see Figure 4.2)¹³³ confirms the expected result; that is to say, hesitations are systematically more present in comedies than they are in non-comedy films, probably as a consequence of the stylistic properties typical of comedies and the pragmatic features of their dialogues, generally characterized by higher degrees of attitudinal discourse (Veirano Pinto 2014: 129)¹³⁴ and a colloquial tone (Zago 2016: 163-164), which makes this genre slightly more similar to spontaneous conversation than non-comedies (Forchini 2012: 96). At the same time, the impact the genre variable has on the rate of occurrence of hesitation markers is not as important as it may be thought. Based on the mean score obtained for both genre categories (occurrences normalized to 1,000 words), the frequency of occurrence of hesitations in comedies compared to those in non-comedies rarely exceeds 40%. More specifically, in the 1950s, comedies present, on average, approximately 24% more hesitations than those featured in non-comedy films, in the 1960s the percentage reaches 35%, 16% in the 1970s, up to 33% in the 2000s and almost 41% in the 2010s¹³⁵. All things considered, even though a preference to employ hesitation markers in comedies can be noticed, genre does not seem to significantly influence the way these features are used in film speech over the years. This is further corroborated by the CTree and the ANOVA statistical analyses. These tests have been computed encompassing both predictor variables simultaneously, i.e. genre and decade. As can be observed from the graph in the Appendices, the CTree test yielded a single node (i.e. no splits), while the ANOVA test registered a p-value of 0.640 for the genre variable and a p-value of 0.8762 for the ratio between genre and

¹³³ The mean score for each film genre has been obtained by summing the normalized frequencies of the two films included for each genre and dividing the result by two.

¹³⁴ According to Veirano Pinto (2014), comedies present more marks of attitudinal discourse as a consequence of the fact that, in this genre, films tend to “rely on clash between situations and attitudes to create misunderstandings and confusion that, in turn, create humour” (*ibidem*: 129; see also Kozloff 2000: 198).

¹³⁵ To find out how much higher the mean score of one genre category is compared to the other, the following formula has been used: $\frac{\text{higher mean score} - \text{lower mean score}}{\text{lower mean score}} \times 100$.

decade. These results seem to suggest that normalized frequencies are statistically stable over time; also, they are not affected by genre in a statistically meaningful way.

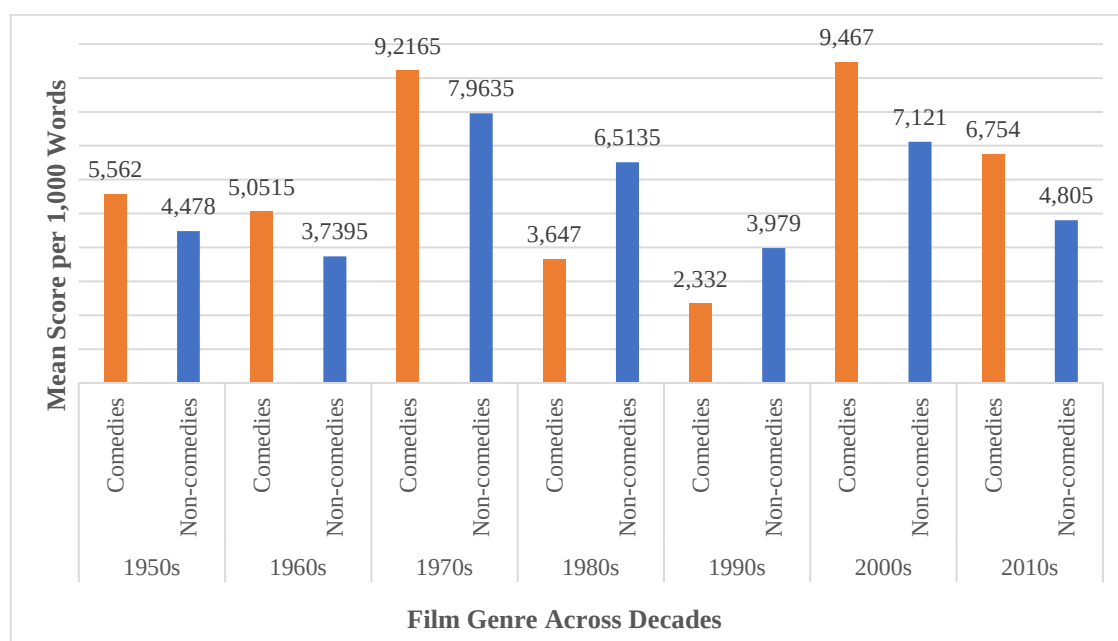


Figure 4.2. *Hesitations in Comedies and Non-Comedies over the Years*

As expected, hesitations are particularly present in the films produced in the 1970s and in the 2000s. Interestingly, in these two decades not only comedies but also drama films present a higher mean score compared to the value obtained for the same genre in other moments in time. As a consequence, even though hesitations are more present in comedies than they are in non-comedy movies, the increase of hesitations in the 1970s and in the 2000s occurs regardless of the genre of the film. To phrase it differently, in both genre categories taken into account, orality features like the ones observed in the current study seem to be encouraged, approximating film speech to a more spontaneous sounding style. Non-comedy films reach their highest peak in the 1970s (in line with the overall trend established in the previous section §4.1.1), where hesitations saw a more marked increase compared to comedy films (+113% as against +82%) (see Table 4.5). This may indicate that the receptiveness to spontaneity and orality has concerned primarily “non-narrative” (Forchini 2012: 98) film genres. On the contrary, comedy films reach the highest value in the 2000s, exhibiting a stronger increase (+305%) compared to non-comedies, where the phenomenon rises in a more moderate way (+79%). This may reflect an extensive use of hesitations in comedies in more recent titles, where perhaps they work as a more dominant stylistic element not only to convey naturalness and improvisation but also as an expressive and comical resource, i.e.

to emphasize humour. This positive peak established for comedy films is followed by a slight downsizing in the 2010s.

Table 4.5. Percentage Variation of Hesitations in Comedies and Non-Comedies		
Decades	Percentage Variation in Comedies	Percentage Variation in Non-Comedies
1950 > 1960	-9%	-17%
1960 > 1970	+82%	+113%
1970 > 1980	-60%	-18%
1980 > 1990	-36%	-39%
1990 > 2000	+305%	+79%
2000 > 2010	-29%	-33%

As already pointed out, hesitations are, generally speaking, more frequent in comedies than they are in non-comedies. The 1980s and the 1990s, however, exhibit the opposite trend. As a matter of fact, even though both genre categories show a decline in the phenomenon, comedy films exhibit a much sharper diminution. To be more precise, in the 1980s and in the 1990s, non-comedy films feature a normalized frequency that is almost 80% higher and 71% higher than comedies, respectively. In order to explain this, a more accurate inquiry of the titles selected for the study has been carried out.

As for the 1980s, the normalized frequencies of hesitations in relation to the genre variable indicate that, among the four titles taken into account (*Tootsie*, 1982; *The Verdict*, 1982; *Mississippi Burning*, 1988; *Driving Miss Daisy*, 1989), the main differences are observable in the comedy and the drama film produced at the beginning of the decade, namely *Tootsie* (1982) and *The Verdict* (1982). The two other titles selected for the analysis — i.e. *Mississippi Burning*, 1988 (drama) and *Driving Miss Daisy*, 1989 (comedy) — present similar normalized rates (2.588 and 2.060, respectively). As for the two titles produced in 1982, it is interesting to notice that in both films hesitations are quite present (with 129 tokens in the drama film as against 91 in the comedy); nevertheless, *Tootsie* presents a mean score that is nearly half the mean score of *The Verdict*. In this case it may be possible to hypothesize that these results occur as a consequence of the actors' style of performance and/or directing choices; in other words, linguistic and stylistic choices made by the scriptwriter, the film director or the actor might have influenced the ratio of employment of hesitation markers as pragmatic tools in the narration. In the case of the two titles analysed here, it is possible to hypothesize that, while in *Tootsie* hesitations mainly contribute to the comical effect of the film — a talented actor with an impossible reputation has to face the consequences of adopting a female alter ego, Dorothy Michaels, to get a job in a TV soap

opera — in *The Verdict*, hesitations are used as an expressive tool for the characterization of the film protagonist, the lawyer Frank Galvin (Paul Newman), involved in a medical malpractice case. As also pointed out in an interview by the director of the film, Sidney Lumet, this is a film about personal salvation where “the case only serves as the instrument by which a man saves himself” (Lumet 2009: online)¹³⁶. Indeed, “Galvin’s journey of self-interrogation, discovery, and affirmation, as a lawyer, is complex. It is punctuated throughout by pauses, by silences, some brief, others lengthy, each a gnawing and revelatory insistence that Galvin confront and engage, not just his circumstance, but his very self” (DeCoste 2002: online)¹³⁷.

Similar considerations apply to the 1990s. In this case, the lower mean score obtained for the comedy genre category is influenced by the very few tokens of hesitation markers found in the film *Forrest Gump* (1994), i.e. 0.916 as opposed to 3.748 in *Jerry Maguire* (1996), 5.929 in *The Silence of the Lambs* (1991) and 2.029 in *American Beauty* (1999). Further discussion of the stylistic and narrative tools employed in this film (for instance, the voice-over technique) will follow in the next section.

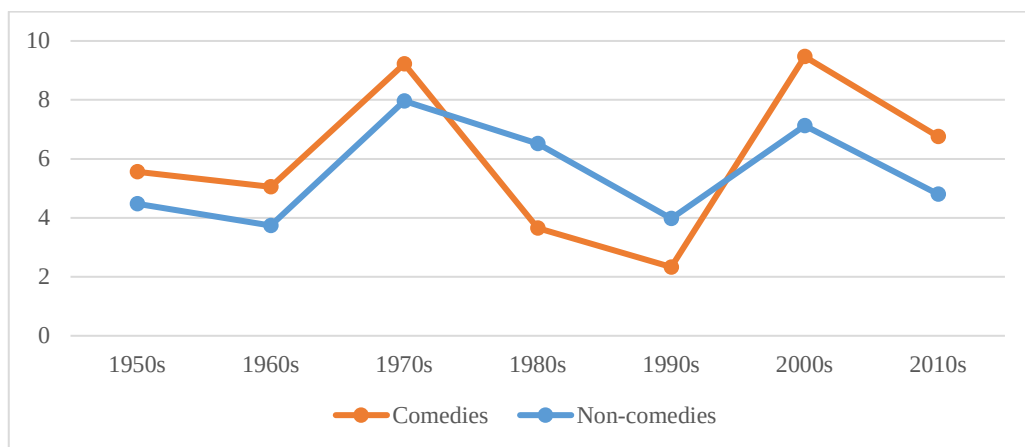


Figure 4.3. Variation of Hesitations in Comedies and in Non-Comedies (1950s-2010s)

Although hesitations occur slightly more often in comedies than in non-comedy films, the collected data seem to suggest that movie genre does not significantly influence the amount of hesitation markers featured in cinematic dialogue. On the contrary, both comedies and non-comedies exhibit similar trends over the years. This can be observed in the chart above (Figure 4.3), where the variation of hesitations in comedies and non-comedies over the years

¹³⁶ Sidney Lumet’s interview, available online at <https://www.youtube.com/watch?v=QaPsVydfWvY>.

¹³⁷ Review available online at https://cap-press.com/sites/pj/verdict_decoste.htm?srsId=AfmBOorQ-4X1Acx4VgUnAkuwfd1YR42xSvQmvvO5yfmXNATDEl-hElxh.

performs in the same way. The only difference occurs in the transition from the 1970s > 1980s to the 1980s > 1990s, where the negative percentage change decreases slightly in comedies (+24 percentage points) and increases in drama (-21 percentage points)¹³⁸. To phrase it differently, in the transition from the 1970s to the 1980s, the decrease is more dramatic in comedy films compared to non-comedies, whereas the opposite occurs in the transition from the 1980s to the 1990s, i.e. the rate of employment of the phenomenon lessens more markedly in the non-comedy category. Therefore, while in comedies the decline is particularly pronounced in the 1980s, marked by an abrupt diminution of the phenomenon, in non-comedy films the decrease occurs more gradually over time. In the transition to the following decade, the opposite trend can be observed.

All things considered, in non-comedy films the trend is more regular with less extreme variations compared to comedies, where changes are more abrupt. This could be interpreted as a sign that in comedy films there is a tendency to experiment more. The data reported in the current section highlight a non-linear pattern in the use of hesitations in film speech with respect to genre, which may reflect the fact that the presence of this phenomenon in film speech seems to be influenced by linguistic fads, stylistic preferences and changes due to the socio-cultural context, rather than by a structural transformation of film language depending on genre. More specifically, albeit with different positive and negative variations, the two film genre categories showcase a similar trend over the years. That is to say, the evolution displayed regarding the use of hesitations in film speech over time is not limited to a single genre; on the contrary, the same trend occurs in both contexts (comedies and non-comedies) simultaneously. This may indicate that the incidence of hesitations in film speech is not attributable to factors contingent to genre or specifically related to the genre of a film. In other words, there must be other factors common to both genre categories, and therefore attributable to changes in cinema as a whole, that may affect the use of these orality markers — factors such as the scriptwriting style, the acting style or the influence of contemporary speech. A more detailed discussion on this is provided hereafter.

¹³⁸ The difference between percentage points is obtained by making the difference between the percentage of the most recent value (-36% for comedies and -39% for drama films) and the previous value (-60% for comedies and -18% for drama films) (see data in Table 4.5).

4.1.3. The Evolution of Hesitations in Film Speech: A Discussion of the Quantitative Data

As already pointed out, not only stylistic, but also cultural and generational changes in the movie production system inevitably affect the way scripts are written and performed, also impacting on the linguistic preferences specific to each period. Indeed, as noted by Kozloff (2000), film dialogue is not an independent entity: “[i]ndustrial, technological, and social changes have all affected [and will continue to affect] the ways in which films have their characters speak” (*ibidem*: 19).

The findings presented in the current study seem to be in line with the linguistic changes highlighted by previous scholars as a result of the evolution of norms in the cinematic industry over time (Kozloff 2000; Veirano Pinto 2014). In fact, while in the 1930s, the introduction of the Production Code as a form of control and censorship concerning all aspects of film production (plot development, language, moral attitudes, etc.) resulted in the standardization of film language — generating dialogues characterized by a planned speech style — the progressive opposition to the Code in the late 1950s triggered a change in style, with films favouring a more colloquial language. This tendency became even more evident in the 1960s, when the lifting of the Production Code gave new impulse to movie language, with less rhythmic and more colloquial dialogues, as well as the use of more explicit language, also more marked by improvisation (Kozloff 2000: 22-23; Veirano Pinto 2014: 123). In Kozloff’s (2000: 23) words, “[t]he accompanying acting style was less declamatory, faster, and more throwaway”, resulting in more realistic and informal language, thereby contributing to the “loosening of the careful precision of mainstream films’ dialogue” (*ibidem*: 24).

The state of affairs of the film production industry presented above would explain the significant increase of hesitations found in the films from the 1970s, as well as the progressive decrease in the following decades. In fact, as pointed out by Veirano Pinto (2014), the drive for innovation that characterized the 1970s found some new balance in the 1980s, with the consolidation of those “trends in moviemaking that had been developing since the end of the studio system in the early 1950s” (*ibidem*: 123). According to the author, these trends comprised “the use of more fragmented language, which is more typical of natural speech” (*ibidem*). The author’s investigation into the spontaneity of film speech compared to unplanned conversation thus offers valuable insights for interpreting the data obtained from the corpus built for the present study.

Furthermore, she points out that movies produced between the 1990s and the 2000s present a lower mean score of spontaneity features compared to the previous years. According to the author, this drop could be explained by the increase in the number of films longer than two hours in those twenty years (Veirano Pinto 2014: 123). She argues that longer movies present more marks of planned speech compared to movies of an average length (up to two hours), which present more features of spontaneous speech (*ibidem*: 124). The present study likewise reveals a noticeable increase in film length beginning in the 1990s. The films included in this sub-section of the corpus range from 118 to 142 minutes, i.e. from two to almost two and a half hours. This observation is further supported by the calculation of the mean duration of films in the seven decades (obtained from the total length of films per each decennium divided by the number of titles taken into account for each ten-year span). As it can be evinced from Table 4.6, the negative peak noticed in the 1990s could be explained by an increase in the duration of films, also offering support to the statement that longer movies are characterized by features of planned speech, making the language used in dialogues less spontaneous (Veirano Pinto 2014: 123).

Table 4.6. Average Length of Films per Decades		
Decade	Total Running Time (minutes)	Mean Running Time (minutes)
1950s	410	102.5
1960s	474	118.5
1970s	411	102.75
1980s	469	117.25
1990s	520	130
2000s	469	117.25
2010s	485	121.25

Moreover, it is noteworthy that three of the films from the 1990s resort to voice-over to convey both the protagonists feelings and thoughts and to advance the narrative. Specifically, the use of the voice-over is particularly prominent in the film *Forrest Gump*, where the lowest occurrence of hesitations has been registered; de facto, in the transcription of the film, the “voice-over” tag or “v.o.” has been used more than 100 times in relation to lines uttered by the protagonist of the film. Thus, it may be inferred that the lower rate of hesitations found in the films from this decade could be attributed to a reduced presence of interpersonal verbal exchanges between characters. In fact, the language of voice-over can be seen as a form of narrative language; as suggested by Zago (2016: 135), “[t]he language of voiceover can be situated in the area of the language of distance in that it has a descriptive rather than an interpersonal focus”.

The impact of the specific stylistic and narrative techniques employed in the movies from the 1990s on the rate of employment of hesitation markers in their dialogues has been further corroborated through a comparison of the analysed corpus with the *Movie Corpus* (Davies 2019)¹³⁹, which contains 200 million words of data in more than 25,000 movies from 1930 to 2018. A comparison of the frequency of *uh* and *um* has been carried out using the List tool. More specifically, the analysis was accomplished by limiting the search to specific sections of the corpus by selecting the American component of the database (the US/CA section includes 153,308,236 out of the total 199,479,302 words) and only the decades included in the present investigation (1950s-2010s), looking at both raw frequencies (raw f.) and normalized frequencies (per mil).

Table 4.7. Hesitations in the <i>Movie Corpus</i> over Time									
Investigation of <i>uh</i> in the <i>Movie Corpus</i>									
	All	1950s	1960s	1970s	1980s	1990s	2000s	2010s	US/CA
raw f.	65,872	4,341	4,737	6,144	8,645	17,547	31,198	51,335	131,745
per mil	429.70	408.91	520.39	608.04	612.32	712.90	608.97	810.74	859.41
Investigation of <i>um</i> in the <i>Movie Corpus</i>									
	All	1950s	1960s	1970s	1980s	1990s	2000s	2010s	US/CA
raw f.	25,437	373	602	785	2,012	5,443	13,041	27,959	50,875
per mil	165.93	35.14	66.13	77.69	142.51	221.14	254.56	441.56	331.87

The *Movie Corpus* highlights a linear growing tendency of hesitations in film language (see Table 4.7). Nevertheless, it must be pointed out that because the data of the reference corpus were extracted from subtitle files, this linear growing tendency does not necessarily imply that hesitation markers are more frequently employed in more recent movies compared to older ones; this could reflect a growing attention toward these pragmatic features in subtitled texts over time¹⁴⁰. In other words, it could be hypothesized that in more recent subtitled files, adapters tend to be more mindful of the meanings conveyed through these elements, which are more likely to be reproduced in the written text. On the other hand, the data obtained from the analysis of the *Movie Corpus* seem to suggest that the decline observed in connection with the 1990s in the diachronic corpus compiled for this study may be attributable to the stylistic and narrative preferences characterizing the movies selected for the analysis (longer running time and voice-over). To phrase it differently, since the negative

¹³⁹ Davies M. 2019. *The Movie Corpus*. Available online at <https://www.english-corpora.org/movies/>.

¹⁴⁰ On the limits of using the *Movie Corpus* as a reference corpus, see §3.8.4 in Chapter Three.

peak found in the 1990s is not mirrored in the reference corpus, the limited amount of hesitation markers observed in the database created for the present study may be ascribed to the specificities of the current corpus rather than indicating a trend attributable to film speech as a whole.

Having established the incidence of the phenomenon in relation to both the year of production and the genre of the film, as well as some of the stylistic and contextual factors related to the evolution of the moviemaking industry, a more in-depth analysis of the way in which hesitations are employed in cinematic dialogue has been carried out. The qualitative analysis presented in the following sections aims to observe the functions these elements fulfil in the narrative context displayed on the screen.

4.2. Hesitations in Film Speech over Time: A Functional Analysis

In the following sections, the phenomenon of hesitations will be presented and observed in relation to their location in the conversational turn. An analysis of the specific functions hesitations fulfil in these locations will be presented, with a view to ascertaining whether there is any kind of correlation between the location of hesitations and the way they are employed in film speech. Results will be compared with the findings of previous research on spontaneous speech (see Chapter One). The insights gained from the functional analysis will allow for further reflection on the status of the category of hesitations as a whole.

4.2.1. The Location of Hesitations in Film Speech: A Turn-Based Analysis

In order to investigate where hesitations occur in film speech, attention has been paid to their position within the speaker's turn. In particular, four different locations have been identified, namely: turn-initial, turn-medial, turn-final position, as well as hesitations occurring as single-turns, meaning free-standing items in conversation. In the analysis, each token has been allocated to one of these four categories based on the peculiarities of each specific turn (see §3.5 for a detailed discussion of the criteria for the allocation of hesitations to the four categories presented above).

Results show that hesitations mainly occur in two positions, namely at the beginning of a turn or in middle position. On the contrary, hesitations are quite rare at the end of a turn, and even rarer as free-standing items (see Table 4.8). More specifically, hesitations are most commonly employed in middle position within the turn; as a matter of fact, the turn-medial location exhibits the highest frequency of occurrence of the phenomenon in 19 out of the 28

analysed films. Among these, there are various cases (titles followed by the asterisk symbol in Table 4.8) in which, even though the turn-medial location is the most frequent one, the difference between the number of hesitations in turn-initial and turn-medial position is not particularly significant. For instance, in *Separate Tables* (1958) or in *Silver Linings Playbook* (2012), the difference is only a matter of a few occurrences. Moreover, there are eight cases in which hesitations at the beginning of the turn prevail over those occurring in turn-medial position. It is interesting to notice that this commonly happens in movies where the phenomenon of hesitations is not particularly present — such as *Driving Miss Daisy* (1989), *American Beauty* (1999) or *Joker* (2019), which are among the films with the lowest number of tokens (20, 19 and 18, respectively) — with the exception of *In the Bedroom* (2001), *Juno* (2007) and *Manchester by the Sea* (2016), which present a higher number of occurrences (95, 87 and 105, respectively).

Table 4.8. Number of Hesitations in Different Locations				
Film Title	Turn-Initial	Turn-Medial	Turn-Final	Single-Turn
<i>Father of the Bride*</i>	43	47	7	3
<i>Mister Roberts</i>	20	28	4	0
<i>Twelve Angry Men</i>	31	59	3	1
<i>Separate Tables*</i>	16	17	1	0
<i>The Apartment</i>	23	35	2	0
<i>The Hustler</i>	20	18	3	0
<i>Guess Who's Coming to Dinner</i>	37	50	2	0
<i>The Graduate</i>	18	11	1	1
<i>The Conversation</i>	27	40	3	0
<i>Annie Hall</i>	28	98	6	0
<i>Heaven Can Wait*</i>	57	60	9	5
<i>Kramer vs. Kramer</i>	34	65	5	0
<i>Tootsie</i>	33	52	5	1
<i>The Verdict</i>	42	86	1	0
<i>Mississippi Burning</i>	7	19	0	1
<i>Driving Miss Daisy</i>	16	4	0	0
<i>The Silence of the Lambs</i>	23	32	0	1
<i>Forrest Gump*</i>	4	6	0	1
<i>Jerry Maguire</i>	20	32	5	3
<i>American Beauty</i>	11	8	0	0
<i>In the Bedroom</i>	48	44	2	1
<i>Sideways</i>	47	80	3	2
<i>Michael Clayton</i>	13	37	3	0
<i>Juno</i>	43	41	3	0
<i>The Kids Are All Right*</i>	59	63	4	4
<i>Silver Linings Playbook*</i>	12	16	2	0
<i>Manchester by the Sea</i>	53	35	10	7
<i>Joker</i>	8	7	2	1
Total	793 (40%)	1090 (54%)	86 (4%)	32 (1.6%)

These results are in line with previous investigations conducted on hesitations in spontaneous speech (see §1.4.3). The results reported in Table 4.8 are in accordance with Tottie's (2014a) findings obtained from the investigation of the *Santa Barbara Corpus of Spoken American English*¹⁴¹, where hesitations mainly occur in turn-medial position, followed by turn-initial and turn-final position (*ibidem*: 393). These findings are consistent with those reported by Boonsuk *et al.*'s (2019) study of filled pauses in informal spontaneous conversation, according to which the phenomenon more frequently occurs in the middle of the speaker's utterance. Hence, based on the current data, it can be argued that the use of hesitations in film speech seems to reflect the way speakers employ these features in spontaneous conversation. This claim, however, needs to be further examined by looking into the specific ways in which hesitations are used in film dialogues, i.e. their functions within the interactions on the screen and the correlation between their function and their location in the sentence.

4.2.2. The Location of Hesitations in the Turn: A Functional Analysis

In order to investigate the functional role of hesitation markers in film speech, the phenomenon has been observed in the four identified locations (turn-initial, turn-medial, turn-final, and single-turn). By doing so, the correlation between the pragmatic and communicative function carried out by these elements in cinematic dialogue and their location within the turn has been examined. For each location, tokens have been classified according to the five Functional Domains identified as relevant for the analysis, i.e. Structural Domain, Planning Domain, Interpersonal Domain, Cognitive Domain, and Hesitation Proper Domain (see § 3.6 for a detailed discussion).

The correlation between the position of hesitations within the turn and the function they fulfil in context reveals that turn-initial hesitations most commonly fulfil functions within the Structural Domain; turn-medial hesitations typically fulfil functions grouped within the Planning Domain, as is also the case for hesitations occurring at the end of the turn. However, as it will be explained hereafter, the end-of-turn category is more elusive, as it is often difficult to discriminate how a talking character intends to avail of turn-final hesitation markers; for example, whether the speaker is planning how to continue the

¹⁴¹ Du Bois J.W., Chafe W.L., Meyer C., Thompson S.A., Englebretson R., Martey N. 2000-2005. *Santa Barbara Corpus of Spoken American English, Parts 1-4*. Philadelphia: Linguistic Data Consortium. Available online at <https://www.linguistics.ucsb.edu/research/santa-barbara-corpus>.

sentence but gets interrupted by another character or if the speaker is intentionally ceding the floor. Finally, the single-turn category exhibits a balanced functional use of hesitations with regard to the Planning and the Structural Domain; moreover, in this section, the Hesitation Proper category also assumes a more relevant role.

A more detailed discussion of the five Functional Domains in each of the four positions established within the turn is provided hereafter.

4.2.2.1. Hesitations in Film Speech: The Occurrence of Hesitations at the Beginning of the Turn

As pointed out in Table 4.8, hesitations are commonly used by speakers at the beginning of the turn. It is interesting to notice that, in most cases, hesitations do not occur as the very first element of the turn, but they appear within the first few items uttered by the speaker. As it has been pointed out in Chapter Three (see §3.5), in the current study the turn-initial category tendentially includes the first four items uttered by a character. In the analysed data, hesitations occur as the very first item of the turn 357 times; that is to say, they represent 45% of the total occurrences of the phenomenon at the beginning of the turn. These results are in line with the evidence reported in the pioneering studies conducted on hesitations in spontaneous speech reviewed in Chapter One (see §1.2.2) such as Boomer (1965) and Hawkins (1971). In compliance with Boomer's (1965) data, hesitations in film speech usually occur after the first word of the clause or at clause boundaries, i.e. after a pause so as to start a new utterance or a new turn (see also Hawkins 1971). Hence, as noted for spontaneous conversation, film speech likewise exhibits a tendency to employ hesitations following a preliminary selection of words, which reflects the initial decision fictitiously made by the speaker regarding the structure of the utterance¹⁴²; more specifically, the speaker starts verbalizing the sentence (i.e. makes a first decision concerning how to begin the sentence), then stops and thinks about "the next word, or the next series of words" (Goldman-Eisler 1958: 67) to continue such sentence. Moreover, this seems to support the claim that, when remembering and performing out loud the line they learned in advance, actors reproduce what commonly happens in spontaneous conversation, notably the fact that speakers often start talking before they have actually selected every word or phrase they are going to utter (Clark 2004: 373).

¹⁴² These decisions are 'fictitiously' made because, as already discussed in Chapter Two, film dialogues are pre-planned, that is, scripted and rehearsed in advance; nevertheless, they are performed in such a way that they should sound spontaneous or improvised.

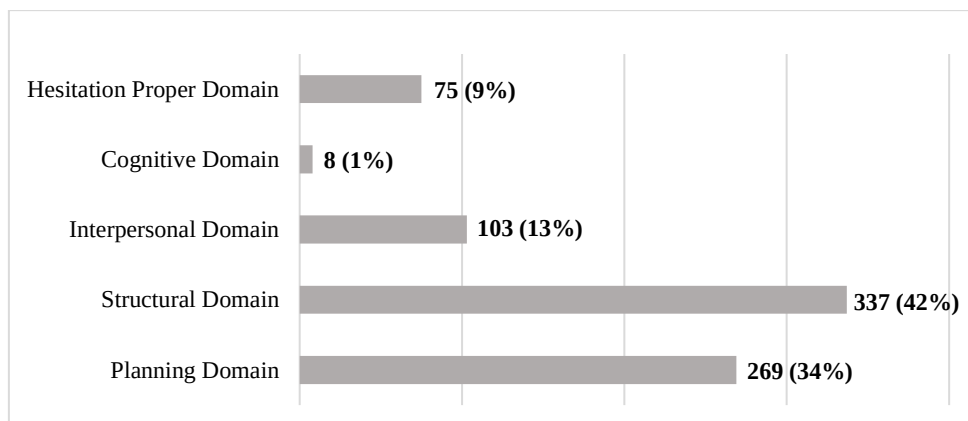


Figure 4.4. Functional Domains of Hesitation in the Start-Turn Category (Raw Frequencies and Percentages)

The functional analysis of hesitations in the turn-initial category highlights that, in this location, the phenomenon is predominantly employed to accomplish discursive tasks (Figure 4.4). It is no surprise that turn-initial hesitations mainly perform functions within the Structural Domain (approximately 42%), particularly given that this category includes tasks associated with turn-taking and turn-launching; that is, reflecting the speaker's intention to take the floor, join the conversation and actively contribute to the communicative exchange. Indeed, these represent the most frequent uses of hesitations at the beginning of the turn. The Structural Domain is followed by the Planning Domain, reproducing the fact that speakers launch their turn before having fully decided what they are going to say, thus replicating speakers' need to buy some time to arrange their utterances and decide how to carry on (34%). The Planning Domain is followed by the Interpersonal Domain (approximately 13%); in particular, turn-initial hesitations are often employed to grab the attention of the interlocutor when taking the turn, as well as to plan and launch particular types of speech acts that could potentially threaten the face of the interlocutor such as delivering and/or responding to requests, offers and invitations. This category is immediately followed by the Hesitation Proper Domain, most notably transmitting characters' tentativeness in answering questions and in launching turns (approximately 9%). Finally, the Cognitive Domain is the one where fewer hesitation markers occur at the beginning of the turn, reproducing a cognitive effort made by the speaker to recall information or to think about details, especially when having to answer a question (1%).

4.2.2.2. Hesitations in Film Speech: The Occurrence of Hesitations in the Middle of the Turn

Turn-medial hesitations are the most frequent in the corpus. In accordance with findings obtained from the investigation of these features in impromptu conversation, speakers most

commonly appeal to these elements in the middle of the turn as holding devices (Structural Domain) or as planning tools within their conversational turn (Planning Domain) (Tottie 2014a; Boonsuk *et al.* 2019).

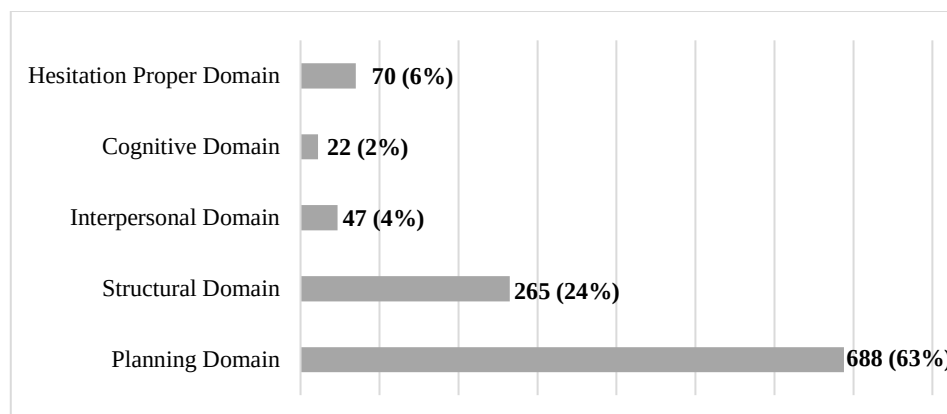


Figure 4.5. Functional Domains of Hesitations in the Intra-Turn Category (Raw Frequencies and Percentages)

As can be seen from the graph above (Figure 4.5), turn-medial hesitations most frequently fulfil planning functions (63%), followed by structural functions (24%), functions related to the Hesitation Proper Domain (6%), the Interpersonal Domain (4%) and, finally, the Cognitive Domain (2%). It is unsurprising that turn-medial hesitations most commonly perform functions within the Planning Domain. In fact, as also pointed out by Biber *et al.* (1999: 1054), “a filled pause is most useful at a point of grammatical incompleteness which is nevertheless a major planning point, such as the beginning of a dependent or coordinate clause”. Consequently, it can be expected that hesitation markers are introduced in the middle of a turn to plan and organize how to continue constructing verbal meanings through lexical and structural choices (Clark 2004). As for the Structural Domain, turn-medial hesitations are typically employed as holding devices, i.e. tools that speakers utilize to maintain and organize their turn, to signal that they still have something to add before concluding the turn, as well as to smoothly navigate between different types of speech acts. The Hesitation Proper Domain mainly regards hesitations as conveying a sense of nervousness, confusion, awkwardness and discomfort in certain situations, for instance, when one of the protagonists is lying or to highlight moments of misunderstanding between characters. The Interpersonal Domain in turn-medial positions is mainly used to mitigate the content of the turn or to mitigate specific speech acts such as requests, invitations and/or dispreferred turns. Finally, even in this case, hesitations categorized within the Cognitive Domain are mainly used to render the mental process carried out by speakers when retrieving information.

4.2.2.3. *Hesitations in Film Speech: The Occurrence of Hesitations at the End of the Turn*

Analysing hesitations in turn-final position can be challenging. Scholars have highlighted how tricky it can be to interpret hesitations at the end of a turn in spontaneous conversation — distinguishing, for instance, between a turn-holding and a turn-yielding function (Stenström 1990; Kjellmer 2003; Tottie 2014a). The same occurs in film speech. Even though in this context dialogues are planned, it often remains difficult to confidently establish whether the hesitation at the end of a turn occurs because the speaker is being interrupted by another character or because the speaker is ceding the floor to someone else; alternatively, the speaker may even decide to simply interrupt the turn and leave it unfinished, in a similar manner to what Tottie (2014a: 402) defines as “unintentional symptoms of planning”.

Given the nature of the corpus under study, it is essential to consider the occurrence of a hesitation within the actual scene of the movie. However, it must be pointed out that the conversational interaction depicted on-screen is not always straightforward and the role played by the hesitation marker cannot always be interpreted with certainty. For these reasons, a sixth category has been incorporated into the present study in addition to the five functional categories originally established, i.e. the Uncertain Domain. This encompasses the less transparent functions associated with the discursive planning and management of the utterance, without incurring the risk of misinterpreting the speakers’ intentions and, consequently, erroneously categorizing the phenomenon into one of the two more narrowly defined groups, namely the Planning and the Structural Domain. This ambiguity, however, persists only in 6 out of the 86 instances of turn-final hesitations found in the corpus (roughly 7% of cases) (see Figure 4.6).

These preliminary considerations have made it possible to uncover some of the ways in which hesitations are employed in cinematic dialogues at the end of a turn. More precisely, turn-final hesitations are mainly used to carry out a planning process or a reasoning process, either left unfinished by speakers themselves (leaving the sentence hanging) or interrupted by another character taking the turn (Planning Domain, 69%)¹⁴³. Within the Structural Domain (8%), they are commonly employed as holding devices interrupted by another

¹⁴³ If not interrupted, the character would presumably continue with the sentence and finish his or her turn; hence, if not interrupted, these turn-final hesitations would actually correspond to turn-medial hesitations. In line with the considerations made regarding the most frequent functional employment of turn-medial hesitations, the present remark explains the high rate registered for the Planning Domain within the turn-final category.

character inserting in the conversation (or by cinematic mechanisms such as a change of shot or a change of scene) and as turn-yielding devices, implicitly asking another character to step in and contribute to the exchange or simply leaving some space for someone else to take the turn. Interestingly, in this position the Hesitation Proper Domain appears to be more balanced with respect to the other functional categories (9%), as is the case with the Cognitive Domain (5%), which becomes slightly more present than the Interpersonal one (2%).

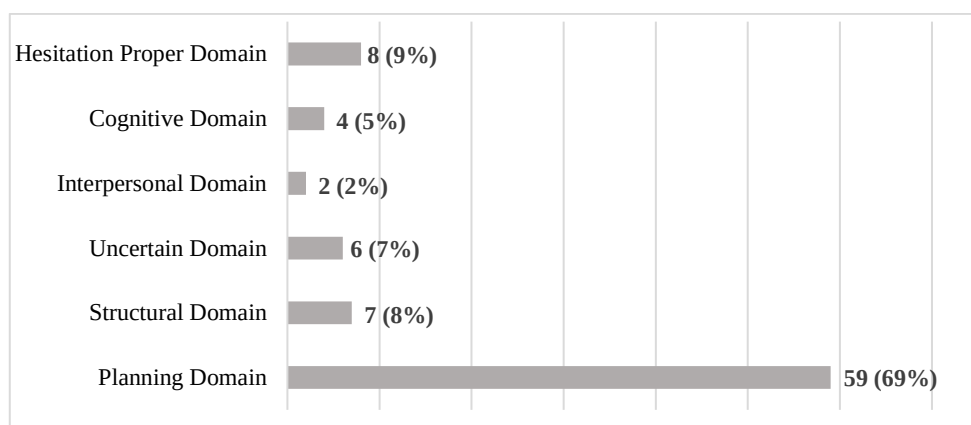


Figure 4.6. Functional Domains of Hesitations in the End-Turn Category (Raw Frequencies and Percentages)

4.2.2.4. Hesitations in Film Speech: The Occurrence of Hesitations as Free-Standing Items

As pointed out by Norrick (2014), features like *uh* or *um* can occur as free-standing items in conversation, meaning that hesitations alone can constitute a turn in verbal exchanges. This occurs quite uncommonly in the corpus, with only 31 occurrences, spread across 13 out of the 28 films analysed. These cases include not only those instances in which hesitations occur as the only component of a communicative turn, but also those cases in which vocalizations occur as a single item preceded and followed by long pauses of silence that separate the hesitation marker from the rest of the turn.

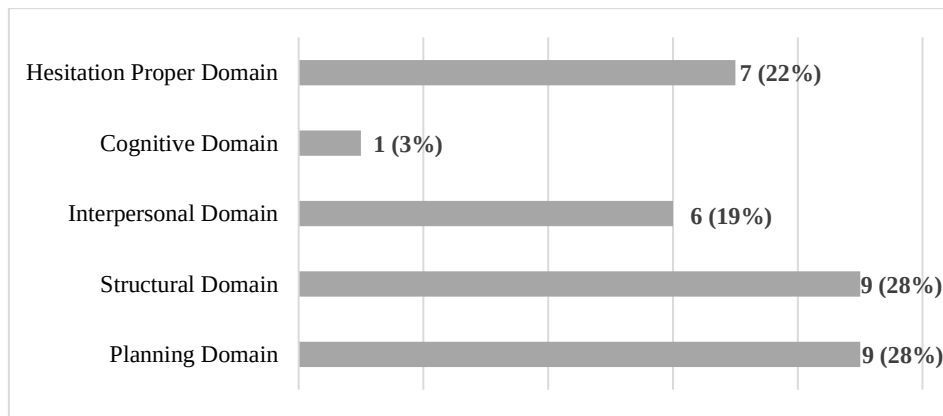


Figure 4.7. Functional Domains of Hesitations in the Single-Turn Category (Raw Frequencies and Percentages)

As for the Planning and Structural Domains (both 28%), hesitations are used as free-standing items to fulfil three main functions: to fill a moment of silence while carrying out some other action visible on the screen, as well as to launch such action; as tools to plan or launch a speech act which is immediately blocked either by the speaker who self-interrupts or, more commonly, by another character taking the turn. As regards the Hesitation Proper Domain (22%), it is interesting to notice that the number of occurrences of this functional category is almost balanced with the discursive functions of planning and structuring the turn (7 tokens for the Hesitation Proper Domain as against 9 for both the Structural and the Planning Domain). When occurring as single-turn items, hesitations mainly represent a tool to express the character's difficulty in coming up with something to say or as a backchannel, i.e. a reaction to a moment of difficulty or trouble provoked by the situation depicted on the screen. As for the Interpersonal Domain (19%), hesitations are commonly used as free-standing items to grab the attention of another character, hence maintaining the connection and preventing the interlocutor from interrupting the verbal exchange. Finally, within the Cognitive Domain (3%), there is only one instance of a hesitation marker specifically used as a free-standing item to express some sort of cognitive effort made by a character to recall precise information.

4.3. The Multifunctional Use of Hesitations in Film Speech: Examples from the Corpus

Having examined the functional ways in which hesitations are employed in film speech according to their location within the turn, the next section will further investigate the employment of hesitation markers across the five Functional Domains with the aim to

highlight some of the prototypical ways in which these features are used in the cinematic context.

4.3.1. Hesitations in Film Speech: The Structural Domain

As already pointed out, the Structural Domain includes those functions related to the management of the conversational turn (turn-taking, turn-holding and turn-yielding), as well as to the organization of the content of the turn such as, for instance, a change of speech act.

As previously noted, hesitations mainly fulfil structural functions when occurring at the beginning of the turn, where they are used to manifest the speaker's intention to take the floor and contribute to the communicative exchange (Extract 1), to directly launch the turn, as well as to retake the turn after a moment of silence (Extract 2). In extract 1, for instance, Natalie is being questioned about what happened the night her boyfriend, Frank, died. While telling the facts, Mrs. Keyes, the defendant's attorney, utters the vocalization *uh* to manifest her intention to take the floor to ask Natalie to clarify one aspect of her deposition. In extract 2, instead, the protagonist Juno retakes the floor, starting a new verbal exchange with Mark, after a long silence (2.976085 seconds), which signals the end of the previous conversational interaction.

Extract 1 - *In the Bedroom* (2001)

Natalie	I-- .. I closed the door ... I've-- .. I walked through my room .. I heard some-- yelling I heard yelling, heard yelling .. I got to the stairs ... ((sighs)) . and then I heard-- .. I heard the shot .. and-and Richard-- ((cries))
Mrs. Keyes	uh I'm sorry
Natalie	((crying)) Richard--
Ms. Keyes	could we go back just a minute? . mrs Strout you said you HEARD the shot

Extract 2 - *Juno* (2007)

Juno	you should feel happy homes! I'm giving you and Vanessa the gift of life sweet screaming pooping life and you don't even have to be there when it comes out all covered in <i>um</i> --
Mark	viscera
Juno	blood and guts ... uh whoa? . is that a Les Paul?
Mark	yes it is, Vanessa gave me my own room for all my stuff

Of course, when used to initiate the communicative exchange, hesitations also serve, at least in part, the function of grabbing the interlocutor's attention. This is exemplified in extract 3

from *The Verdict* (1982); in this scene, the funeral director takes the floor, attracting Mrs. Dee's attention, in order to introduce both her and the audience to the protagonist of the film: the lawyer Frank Galvin.

Extract 3 - *The Verdict* (1982)

Funeral director	uh mrs Dee, this is Frank Galvin, he's a good friend of ours and a VERY fine! attorney

Hesitations, however, are also used to actually launch the turn. This is the case in extract 4 below. In this scene from *Sideways* (2004), the protagonist Miles calls Maya, his love interest, and leaves her a voicemail message apologizing for lying to her. Since no one is answering the phone, no interactive verbal exchange is taking place; hence, there is no need to compete with the interlocutor for the "allocation of speakership" (Sacks, Schegloff & Jefferson 1974: 699). Consequently, in this case the hesitation uttered at the beginning of the turn has been interpreted as a pragmatic marker employed by the speaker to open and initiate the telephone soliloquy. Likewise, in extract 5, the hesitation marker *uh* is used to launch a parenthetical aside. In this scene from *The Verdict* (1982), the attorney Concannon is preparing his client, Dr. Towler, to answer some of the questions he will probably be asked in court. When Concannon corrects the *when she* to *when Debby*, Dr. Towler acknowledges the correction with the hesitation marker *uh* and accepts it, validating Concannon's interruption by thanking him before continuing with the utterance. In this extract, Dr. Towler's initial *uh thank you* can be interpreted as what Clark (2004) defines as "side exchanges", namely "projective pairs interposed within or between utterances to deal with local, ongoing performance" (*ibidem*: 374).

Extract 4 - *Sideways* (2004)

Miles on the phone	((offscreen)) uh it's Miles . ((onscreen)) I had to call and <i>uh</i> tell you again . how much I enjoyed out time together and <i>uh</i> .. how sorry I am that the things turned out the way they did .. I think you're great, Maya .. <i>uh</i> always have from the first time that you <i>uh</i> waited on me ... <i>uh</i> and while I'm at it . <i>uh</i> I guess maybe you should know that, my book is not getting published . I thought this one <i>uh</i> had a chance but I was wrong again .. once again .. so you see <i>uh</i> I'm not really much a writer .. ((offscreen)) I am not much of anything, really

Extract 5 - *The Verdict* (1982)

Concannon	anything special about the case?
Dr. Towler	well-- w-when she had been on--
Concannon	when Debby, thank you
Dr. Towler	uh thank you, when Debbie had been under medical--

When a simultaneous face-to-face interaction between two conversing characters occurs, hesitations are commonly used as turn-launching devices to direct what is about to be said in the turn to a new interlocutor. This can be more clearly observed in extract 6 and in extract 7. In the former, retrieved from the film *The Apartment* (1960), after attempting suicide, Fran Kubelik spends two days recovering in Baxter's apartment until her brother-in-law (Karl Matuschka) shows up looking for her. In extract 6, *uh* is used by Miss Kubelik (or Fran) to launch the introduction of his brother-in-law to the male protagonist of the film (Baxter). As a matter of fact, Miss Kubelik ignores Matuschka's questions asking her what happened and why she has not been home for days, and proceeds to address her turn to Baxter. In extract 7, instead, the hesitation marker is used to organize the conversational turns, allowing the speaker to shift interlocutors; specifically, Lester moves from talking to Ricky to talking to Ricky's dad. This transition is structured through the hesitation marker *uh*, which is used to launch a new turn directed to the new interlocutor.

Extract 6 - *The Apartment* (1960)

Matuschka	what's with you Fran? did you forget where you live?
Miss Kubelik to Baxter	uh this is my brother in law Karl Matuschka
Baxter to Matuschka	oh! how do you do <i>uh</i> mister Matusch--

Extract 7 - *American Beauty* (1999)

Lester to Ricky	hey yo Ricky ((panting)) my entire life is passing before my eyes and those two have barely broken a sweat ((panting)) =
Lester to Ricky's dad	= uh sorry hi ((panting)) Lester Burnham I live next door we haven't met

Sometimes, it may happen for a turn to be immediately blocked either by the speaker who self-interrupts — as in the case below (Extract 8) from *The Kids Are All Right* (2010), where Paul launches the answer to Tanya's question but immediately interrupts to greet Brook — or, more commonly, by another character taking the turn.

Extract 8 - *The Kids Are All Right* (2010)

Tanya	wow, so how'd you leave it? I mean are you gonna see them again?
Paul	uh-- =
Paul to Brooke	= oh hi Brooke
Brooke	hey Paul check these out

At the same time, hesitation markers may also work as a verbal launcher of an action visible on the screen; for instance, handing a gift to another character (Extract 9a) or starting a farewell handshake (Extract 9b). As can be inferred from the extracts below, this commonly occurs with hesitations as free-standing turns. In particular, in these cases the hesitation is categorized as a single free-standing item, as it is separated from the previous sentence by a silent pause, which makes the action independent from the sentence preceding the hesitation marker. In the scene extracted from *Forrest Gump* (1994), for instance, Forrest is invited by Jenny, his childhood best friend, to visit her house. In this case, a long pause (3.341473 seconds) separates Forrest's comments on the house and the air conditioning from the next turn, which is launched by the hesitation marker *uh*. This is used to fill the silence while handing the gift Forrest brought for Jenny, i.e. a box full of chocolate. In the case reported from *The Silence of the Lambs* (1991), instead, the hesitation marker is used after a long silent pause (4.024281 seconds) to launch the action of a handshake, which is used to conclude the communicative exchange between Clarice and Mr. Crawford.

Extracts 9 - Hesitations to Launch an Action Visible on the Screen

9a	<i>Forrest Gump</i> (1994)	it's nice . you got air conditioning ... uh--
9b	<i>The Silence of the Lambs</i> (1991)	okay, sure <i>um</i> . thank you, mister Crawford ... um--

Along with a turn-taking function, hesitations often occur at the beginning of the turn in the question/answer adjacency pair (see Schegloff and Sacks 1973)¹⁴⁴. Even though they may occur in asking questions, they are more frequently employed in answering sequences. In the former case, hesitations are employed as tools to launch the question (Extracts 10a-d), to delay asking the actual question (usually followed by a pause, as in extract 10e, where the hesitation marker is followed by a silence of 2.465598 seconds), as well as to ask for confirmation or echo questions before actually launching the response (10f). In the latter

¹⁴⁴ As reported by Schegloff & Sacks (1973: 295), "adjacency pairs consist of sequences which properly have the following features: (1) two utterance length, (2) adjacent positioning of component utterances, (3) different speakers producing each utterance".

case, hesitations are employed in various contexts: to answer a polar question (with both positive and negative answers) (11a), to launch explanatory answers (11b), to provide a vague, uncertain or approximate response (11c) as well as to reply with a precise referent, precise information or facts (11d); hesitations may also be used by a speaker to buy some time to think about how to answer (11e) or to delay the actual answer (usually followed by a pause, as in extract 11f, where the hesitation marker is followed by a silence of 2.000000 seconds).

Extracts 10 - Turn-Initial Hesitations in Asking Questions

10a	<i>Manchester by the Sea</i> (2016)	uh you know we're Christian too right?
10b	<i>The Kids Are All Right</i> (2010)	ehm . do you want to keep playing?
10c	<i>Heaven Can Wait</i> (1978)	uh wh-wh-what's the matter with him? he can't keep his eyes open
10d	<i>Michael Clayton</i> (2007)	uh did you ((offscreen)) think about what we said?
10e	<i>Heaven Can Wait</i> (1978)	uh .. you want to tell me why you're calling me Joe?
10f	<i>Sideways</i> (2004)	uh me? no, no no I'd just fuck 'em up, yeah yeah, no that was the one, unpolluted part of my divorce, no kids

Extracts 11 - Turn-Initial Hesitations in Answering Questions

11a	<i>Tootsie</i> (1982)	uh no he's no longer with the show mister Bishop is rehearsing the Iceman Cometh for Broadway
	<i>The Kids Are All Right</i> (2010)	uh yeah ((chuckles)) I always liked food
11b	<i>Driving Miss Daisy</i> (1989)	uh I'm just dusting your bulbs miss Daisy
11c	<i>Sideways</i> (2004)	((offscreen)) u:h oh seventy something
11d	<i>In the Bedroom</i> (2001)	uh October
11e	<i>Michael Clayton</i> (2007)	u:h yeah yeah I'll take one to go
11f	<i>Juno</i> (2007)	um .. it's-- it's Paulie Bleeker

In addition, within the Structural Domain, hesitations are often employed to retake the conversational turn. This frequently occurs at the beginning of the turn to continue a previous sentence momentarily blocked by the insertion of another character and, consequently, put on hold by the speaker. In these cases, the talking character resorts to hesitation markers to return to the sentence temporarily interrupted and to continue with what he or she was saying. This is the case in both extract 12 and extract 13, where *uh* is used to restart the turn, continuing the previous sentence interrupted by other characters' backchannels.

Extract 12 - Sideways (2004)

Stephanie	oh Miles <i>um</i> --
Miles	yeah?
Stephanie	uh Maya's in the back, back's just past the kitchen
Miles	okay, thanks

Extract 13 - The Graduate (1967)

Mr. Braddock	this boy, I'm sorry, this young man is soon to continue his education as a Frank Halpingham Award scholar
	((verses of people being impressed))
Mr. Braddock	uh ((chuckles)) but before he does, before he does! . =
Mr. Braddock to Ben	= you're disappointing them Ben, you're disappointing them

Nevertheless, a character may decide to add something to what has already been said even with no interruption; hence, maintaining the conversational turn and moving forward with what he or she wants to say. Turn-holding hesitations mainly occur in turn-medial position; nevertheless, they are also employed at the beginning of a turn to signal that after a first answer, backchannel or response, the speaker still has something to add (Extracts 14a), sometimes giving him/her time to think about how to continue (14b). Accordingly, in extract 14c, the hesitation marker allows the speaker to move forward with the storytelling of what happened. In addition, hesitations are also employed to rephrase the content of the turn in a clearer way, namely adding *close by* to explicate the meaning of the word *adjoining* in extract 14d. In this sense, these features allow characters to provide further clarification of the speaker's intended meaning or to offer an explicit elaboration of the information (Extract 14e).

Extracts 14 - Hesitations as Turn-Holding Devices

14a	<i>Kramer vs. Kramer</i> (1979)	he's terrific uh he <i>uh</i> had a little accident a couple weeks ago on a playground and cut himself [...]
	<i>Sideways</i> (2004)	oh Christ uh good thing you have a voice over career
14b	<i>Silence of the Lambs</i> (1991)	anger um . social . acceptance, and <i>um</i> sexual frustration sir
14c	<i>Manchester by the Sea</i> (2016)	no I mean <i>uh</i> -- I brought him in uh we were-- looking at the boat this morning, and-- he just-- I don't know he-he just-- like fell over I-- I thought he was kidding me at first, then I called the ambulance and <i>uh</i> -- .. that was it
14d	<i>The Conversation</i> (1974)	ah ... do you have a room that would be-- <i>uh</i> adjoining uh close by?
14e	<i>Separate Tables</i> (1958)	well don't . it makes me embarrassed ... you really think I haven't changed much? . uh to look at I mean

When fulfilling a turn-holding function, hesitations also allow the speaker to organize the distinct actions carried out within the same turn; for instance, moving from one speech act to a different one. This is the case in extract 15a, where the speaker, Victoria, holds the turn shifting from the polite formula *thanks* (thanking the character of Miles for the compliment he gave to her: “wow you look beautiful”) to the introduction of her husband. Similarly, in extract 15b, the hesitation marker *uh* is used to signal a shift from the “howareyou” sequence typical of phone calls to what has been defined as “reason for the call” (Schegloff 2010: 146). This organizing function can be further exemplified by the scene from *Twelve Angry Men* (1957), where the turn-medial hesitation marker allows the shift from talking about Juror no. 8 (using the third singular person) to actually talking to him, using the second singular person pronoun *you*.

Extracts 15 - Hesitations to Internally Organize the Turn

15a	<i>Sideways</i> (2004)	thanks ((chuckles)) um this is Ken Cortland my husband
15b	<i>The Kids Are All Right</i> (2010)	((offscreen)) I’m well thanks uh ((onscreen)) Wendy at the:-- the Cryobank <i>uh</i> was talking about maybe--
15c	<i>Twelve Angry Men</i> (1957)	okay, <i>um</i> now perhaps if the <i>uh</i> gentleman down there, who is disagreeing with us uh well perhaps you could tell us why you know, let us what you're thinking and <i>uh</i> we might be able to show you where you're mixed up

In this regard, a hesitation marker may be employed to signal a transition from direct speech to reading something — such as reading the discourse prepared for a toast in *Father of the Bride* (1950) — or from direct speech to the reporting of someone else’s words (Extract 16b) or the quotation of a saying (Extract 16c). It may also mark the transition from a declarative sentence to a question (Extract 16d).

Extracts 16 - Turn-Holding Hesitations to Perform Different Speech Acts

16a	<i>Father of the Bride</i> (1950)	now-now what do you think of this? I thought I’d start out . uh friends when I was a very young lawyer I made a speech--
16b	<i>Annie Hall</i> (1977)	((offscreen)) wha--? how am I a par-- wha--? I pick up on those kind of things! you know I was having lunch with some guys from NBC so I said u:h did you eat yet? or what? and Tom Christie said no! jew! not did you? d’you eat? d’you eat? not did you eat? but jew eat? jew, you get it? d’you eat?
16c	<i>In the Bedroom</i> (2001)	well look he lost an arm . you see the trap has nylon nets, called heads, two side-- side heads to let the lobster crawl in, and inside, what's called a bedroom head holds the bait and keep him from escaping now, you know the old saying uh two's company three's a crowd?, well it's like that, you uh-- get more than two of these in a bedroom and chances are something like

		that's happen, that's why Frank can't leave these traps for more than a day
16d	<i>Separate Tables</i> (1958)	that's quite all right major uh when you were at Sandhurst?

To conclude, hesitations also occur as turn-yielding tools (turn-final position), implicitly asking another character to step in and contribute to the exchange or simply leaving the space for someone else to take the turn. This mainly occurs in film speech when the speaking character tacitly asks someone to provide the material of which the speaker (momentarily) lacks. This is the case, for instance, in extract 17, where Katie is looking for the exact term (*acronyms*) to complete the sentence. The right expression, however, does not come to her mind; hence, the speaker turns to Ruth and indirectly asks her to get involved in the exchange so as to help her with the missing word.

Extract 17 - In the Bedroom (2001)	
Katie	it's a little jeep!, S U V? and A T V and K F C, what's with all these <i>uh</i> these uh --?
Ruth	acronyms
Katie	well yeah! I guess it's just too much trouble to--, say what something is anymore

4.3.2. Hesitations in Film Speech: The Planning Domain

The Planning Domain represents the second most frequent functional category with which hesitations are employed within the turn. Hesitations are used at the beginning and in the middle of the turn to plan what to say and how to say it, as a thinking pause in which the speaker takes the time to accomplish some choices regarding the structure of the sentence, the selection of the right word or specific lexical items, as well as the right way to phrase the utterance. In extract 18a, for instance, the whole turn is planned through the insertion of a prolonged hesitation marker *u:h*, which signals the speaker's thinking process about what to include in the list of elements, namely a list of adjectives that could be used to describe the wine during a wine tasting. In other cases, instead, the pause filled by the hesitation marker follows a preliminary selection — hence, the speaker is making an initial decision on how to start the utterance (Extract 18a) — allowing the character to plan how to continue (Extract 18c) or to buy some time while looking for the right word to formulate the utterance (*terrific* and *empty* in extracts 18d and 18e, respectively).

Extracts 18 - Hesitations to Plan the Turn and to Continue the Utterance

18a	<i>Sideways</i> (2004)	u:h thick? thin? watery? syrupy? . okay?
18b	<i>In the Bedroom</i> (2001)	I uh hope I'm not disturbing you
18c	<i>The Hustler</i> (1961)	okay okay, what do I do now lie down on the floor and u:h bow from the ankles? what do I do? go home?
18d	<i>Kramer vs. Kramer</i> (1979)	that's uh terrific, how do you feel?
18e	<i>Silver Linings Playbook</i> (2012)	I know! because I do this time after time after time after time I do all this shit for other people, and then I wake up and-and I'm uh empty! I have nothing!

This planning process carried out by speakers more typically occurs in interactive exchanges between characters, but it can also be noticed in non-interactive sequences. In the former case, the speaker may use hesitation markers as a tool to direct the attention of the interlocutor toward specific information (high-informative items) such as, for instance, proper names or specific referents — as in extracts 19a-b, where a hesitation is used to highlight the name of the person introduced to the interlocutors or as in extract 19c, where *um* precedes the term *sternum*. Alternatively, these elements may be used to highlight central elements in the turn such as, for example, the items around which the question revolves in extracts 19d-e.

Extracts 19 - Hesitations to Direct the Listener's Attention

19a	<i>Juno</i> (2007)	that is uh Herschell Gordon Lewis he's the ultimate master of horror
19b	<i>Annie Hall</i> (1977)	this is uh Shawn and then <i>uh</i> Bobby, Bob and Petronia
19c	<i>Silence of the Lambs</i> (1991)	Bill ... star shaped contact entrance wound over the uh sternum .. <i>um</i> muzzle stamp at the top
19d	<i>Michel Clayton</i> (2007)	they killed them Michael . those small farms the family farms . did you-did you- did you meet uh Anna?
19e	<i>In the Bedroom</i> (2001)	you got any uh tabasco?

In the case of non-interactive scenes, hesitations are employed in sequences of monologues — as in the initial scene from *Annie Hall* (1977) (Extract 20a), where the protagonist plans the recital of two jokes — or in sequences of voice-over narrations, where a non-diegetic (inaudible to characters) voice is added to the soundtrack to provide narrative information, a character's thoughts or commentary on the on-screen action. In extract 20b, for instance, the protagonist's line provides some information regarding the unfolding of his life events. The protagonist, Forrest Gump, plans this sort of storytelling, explaining to the audience what happened in his life using the voice-over technique. More precisely, the character

launches the turn using the particle *so* and, before proceeding with the narration of what happened, he momentarily stops to plan how to continue.

Extracts 20 - Hesitations in Non-Interactive Scenes

Extract 20a - *Annie Hall* (1977)

Alvy Singer	there's an old joke uhm two elderly women are at a catskill mountain resort ((clears throat)) and one of them says, wow the food at this place is +really terrible*! the other says +yeah I know*! and such +small portions! well that's essentially how I feel about life full of loneliness and misery and suffering and unhappiness and, it's all over much too quickly, the-the other important joke for me is one that's uh is usually attributed to Groucho Marx but I think it appears originally in Freud's Wit and its Relation to the Unconscious and it goes, like this I'm paraphrasing, <i>uhm</i> I would never want to belong to any club that would have someone like me for a member, that's the +key!* joke of my adult life in terms of my relationships with women ((inhales))

Extract 20b - *Forrest Gump* (1994)

Forrest (v.o.)	so uh I never went back to work for Lieutenant Dan, though he did take care of my Bubba-Gump money . he got me invested in some kind of, fruit company . and so then I got a call from him saying we don't have to worry about money no more =

As pointed out by Biber *et al.* (1999: 1074), “[c]ommonly an utterance opens with a single prefatory word, which has the function of orienting the listener to the following utterance, especially in relation to what has preceded”. Hesitations generally come after this prefatory word to build interpersonal and descriptive dynamics, allowing speakers to plan their turn and move forward with it. Hence, similarly to what occurs in spontaneous conversation, in film speech, hesitations are employed at points of transition across turns, that is, in conjunction with an utterance launcher (Biber *et al.* 1999: 1054). In the following extract from the film *Sideways* (2004), for instance, the hesitation marker follows the interjection *oh*, which is used to initiate the turn; the turn, however, has not been fully planned by the speaker, who hesitates so as to signal that the character is still thinking about what to say (Extract 21a). Similarly, hesitations often follow other utterance-launching items such as discourse markers (*you know, now, look* in 21b-d), greetings formulae (*hi, good afternoon* in 21e-f) and attention signals such as *hey* (Extract 21g), response forms like *okay* (Extract 21h) or linking adverbs such as *then* (Extract 21i).

Extracts 21 - Hesitations Occurring with Utterance Launchers

21a	<i>Sideways</i> (2004)	oh uh well I like them both, but <i>um</i> if pressed I would have to say that I prefer the dark

21b	<i>The Hustler</i> (1961)	you know uh they say Minnesota Fats is the best in the country out where I come from
21c	<i>Guess Who's Coming to Dinner</i> (1967)	now uh you'll-you'll excuse us please
21d	<i>Kramer vs. Kramer</i> (1979)	look uh Ted I don't honestly think she'd go to the trouble of suing for custody of the child if she's going to kidnap him
21e	<i>Mississippi Burning</i> (1988)	hi uh there was a-- couple of things I needed to check with you
21f	<i>Silence of the Lambs</i> (1991)	good afternoon um sorry to bother you <i>um uh</i> I'm looking for mrs Lippman's family
21g	<i>The Hustler</i> (1961)	well hey uh I'll tell you what I'll play you jack up pool just keep one hand in my pocket
21h	<i>The Kids Are All Right</i> (2010)	right, okay um this is our plan, I don't think you should worry, and I think you should just call me back if it continues
21i	<i>Twelve Angry Men</i> (1957)	okay then uh just remember that this has to be twelve to nothing either way <i>um</i> that's the law, okay are we ready? now all those voting guilty, please raise your hands ((others raise their hands)) one two three four five six seven . eight nine ten, eleven, okay that's eleven, guilty, those voting not guilty?

When preceded by these elements, hesitations are used by speakers to plan the content of the turn. In this context, these pragmatic features are employed in a manner similar to the function of “deliberation signal” attributed by Bibet *et al.* (1999: 1086) to the discourse marker *well*, “indicating the speaker’s need to give (brief) thought or consideration to the point at issue” before continuing with his or her utterance. Similarly, when carrying out a planning function, hesitations are commonly uttered in combination with adverbial elements, i.e. *actually*, *really* and *just* in extracts 22a-c; function words such as conjunctions (Extracts 23a-b), articles (23c), demonstrative pronouns (Extract 23d) and prepositions (Extract 23e), as well as with verbal elements, like the past participle of the verb “to be” used to form the present perfect continuous tense and the future form “we’ll go” (Extracts 24). In this planning context, hesitations also tend to appear in between elements that are usually not disjointed, as in the case of a noun from its preceding article (23c) or a verb from its subject pronoun (24c-d). Moreover, the following examples showcase how hesitations do not always occur as independent items, but are often cliticized at the end of “monosyllabic words whose last consonant is lengthened with vocalizations” (Tottie 2017: 125).

Extracts 22 - Hesitations with Adverbial Elements

22a	<i>Manchester by the Sea</i> (2016)	actually um I-- oh I'm sorry I didn't mean to cut you off I just gotta go I'm--
22b	<i>Annie Hall</i> (1977)	I'm really uh looking forward to tomorrow I mean you know I think it'll be really nice you'll meet Mother and Father
22c	<i>Silver Linings Playbook</i> (2012)	((sighs)) well I'm just um doing this project with Tiffany I just wanna tell her I'm not gonna be on time

Extracts 23 - Hesitations with Function Words

23a	<i>The Silence of the Lambs</i> (1991)	no he was-- very strong he lasted more than a month . my mother died when I was very young so-- ... my father had become the whole world to me and uh when he left me I had nothing I was ten years old
23b	<i>Separate Tables</i> (1958)	well yes but uh well the fellow saved my life once just outside El Alamein, pinned down short of petrol he was a good man
23c	<i>Jerry Maguire</i> (1996)	oh good, good all right . here's Arizona Cardinals, this is <u>the uh team</u> that, Rod plays for
23d	<i>The Hustler</i> (1961)	there's that uh that big brown one's mine it goes in drawing room a huh? thank you
23e	<i>Mississippi Burning</i> (1988)	two . well I was, as I remember, it didn't last very long I was never home, guess she got fed up <u>with uh</u> phone calls from Miami and postcards from Des Moines and . there was always a guy around . any guy that could spare the time for a movie or a beer or-- . a quarter for the jukebox ((chuckles)) ... she left... ((chuckles)) how about you?

Extracts 24 - Hesitations Occurring with Verbal Elements

24a	<i>The Apartment</i> (1960)	been uh hearing some nice things about you, I have a report here from mister Dobisch says you're loyal cooperative resourceful--
24b	<i>Kramer vs. Kramer</i> (1979)	((offscreen)) and after that, if you're really certain you wanna retain custody of your child, then <u>we'll uh go</u> in there and beat the pants off them, okay?
24c	<i>Father of the Bride</i> (1950)	well in the first place Buckley <u>you uh look</u> around at this house and everything and you probably think well, he's pretty well off but let me tell you my boy ((smoking his pie)) that doesn't mean a thing ((smokes)) the day I married Ellie my position with the law firm of Bartham, Henderson and Peck ((smokes)) was as nebulous as--
24d	<i>Tootsie</i> (1982)	Michael I-I'm trying to stay calm here <u>you uh are</u> a wonderful actor

The planning function of hesitations also emerges from the fact that these pragmatic markers often occur in combination with the repetition of grammatical elements, as reported in extracts 25, where hesitation markers are used in between the reiteration of the personal pronouns *we* (25a) and *I* (25b); in the middle of the repetition of the subject pronoun and the verbal element *he's* in extract 25c and *I'll* in extract 25d (where two hesitations are uttered

within the repetition, probably as a tool to relieve the planning pressure at the beginning of a new turn in which an instant decision is being taken, i.e. which chocolates bar to buy from a young scout girl); in between the repetition of demonstrative pronouns (25e), as well as within the reiteration of the quantifier *a little* used as a determiner (25f).

Extracts 25 - Hesitations and Repetitions

25a	<i>Michael Clayton</i> (2007)	we uh we tried calling you six times but uh your machine must be busted or something . we-we gave it our best shot u:h you know we ran out of red covers we only had two thousand in stock so uh we did the other thousand in blue . =
25b	<i>Mister Roberts</i> (1955)	well I uh sure as heck haven't seen much action this past year ((chuckles)) you mean a battle action well you know I uh I got a little shrapnel in my left leg but it's-- it gives me a twinge every now and then that's all
25c	<i>Kramer vs. Kramer</i> (1979)	he's ok he's uh he's sleeping . it's okay I can do that
25d	<i>In the Bedroom</i> (2001)	okay I'll uh-I'll uh-I'll do that .. uh sh-should I hold something?
25e	<i>Annie Hall</i> (1977)	that's uh that's-- pretty serious stuff there
25f	<i>Jerry Maguire</i> (1996)	yeah . okay, okay look uh a little uh a little speech before I go

The repetition within a turn may also coincide with disfluency phenomena such as false starts, namely those cases in which “the speaker says something, then goes back to the beginning of the piece and repeats what has already been said” (Biber *et al.* 1999: 1062). In this context, however, hesitations are not conceived as disfluencies; these features are rather intended as tools used by the speaker as a strategy to gain time to begin and then re-begin the same piece of speech (*ibidem*: 1055). This occurs, for instance, in the line reported in extract 26a, where the false start includes the repetition of *it's* discontinued by a hesitation marker which gives the character enough time to plan his explanation; similarly, in extract 26b, the hesitation within the repetition of the modal verb *wouldn't* provides the speaker some extra time to plan the hypothetical situation.

Extracts 26 - Hesitations and False Starts

26a	<i>In the Bedroom</i> (2001)	see it's uh-- . it's the trial we can't go through that my wife and I, that's why you're leaving I've got you a ticket ... my wife, keeps seeing you I can't have that anymore
26b	<i>The Verdict</i> (1982)	look I swear to you I wouldn't uh-- . I wouldn't turn down the offer if I thought I couldn't win the case

The manifestation of hesitations in combination with repetitions may also occur in questions. As a matter of fact, a pattern which often occurs in the corpus is that of a question being planned through the following sequence: ‘one element uttered to initiate the question + hesitation marker + continuation of the question’. As can be inferred from the following extracts (Extracts 27), the hesitation marker may either be followed by a direct continuation of the sentence or by the repetition of the element initiating the question, i.e. the repetition of the element preceding the hesitation marker (Extracts 28):

Extracts 27 - Hesitations to Plan Asking Questions

27a	<i>Father of the Bride</i> (1950)	[...] now <i>uh</i> a-about the food let's see, <u>what about</u> <i>uh</i> a large salmon at either end of the table? [...]
27b	<i>Guess Who's Coming to Dinner</i> (1967)	= are you <i>uh</i> Catholics mrs Drayton?
27c	<i>Kramer vs. Kramer</i> (1979)	would you <i>uh</i> describe your husband as an alcoholic?
27d	<i>The Verdict</i> (1982)	when can we <i>uh</i> get together again and get a deposition?
27e	<i>Jerry Maguire</i> (1996)	you <i>uh</i> didn't have the talk did you? . hm?
27f	<i>Juno</i> (2007)	how about you Mark? <u>are you</u> <i>uh</i> looking forward to being a dad?
27g	<i>Manchester by the Sea</i> (2016)	which-- . <i>uh</i> part are you having trouble with?
27h	<i>Joker</i> (2019)	did you <i>uh</i> get a new gig?

Extracts 28 - Hesitations within Repetitions to Plan Asking Questions

28a	<i>The Hustler</i> (1961)	you're <i>uh</i> you're Minnesota Fats aren't you?
28b	<i>The Verdict</i> (1982)	<u>is she</u> <i>uh</i> <u>is she</u> getting good treatment over there?
28c	<i>Jerry Maguire</i> (1996)	<u>where's</u> <i>uh--</i> . where's the little guy?
28d	<i>In the Bedroom</i> (2001)	did-did he ever <i>uh--</i> <u>did he ever</u> talk to you about what happened? on that day?
28d	<i>Michael Clayton</i> (2007)	((offscreen)) hello <u>is</u> <i>uh</i> <u>is</u> Anna there?
28e	<i>The Kids Are All Right</i> (2010)	<u>did you</u> <i>uh</i> <u>did you</u> ever play any sports in school or--

Sometimes the planning does not go through. In these cases, the sentence is left hanging unfinished, as in the scenes reported in extracts 29. In the first case, the character of Kirkeby seems to want to keep saying something but self-interrupts to pick up some object from the ground, while in the other cases, characters leave the previous sentence incomplete to start a new one.

Extracts 29 - Hesitations and Sentences Left Unfinished

29a	<i>The Apartment</i> (1960)	uh sure sure uhm -- oh! ((takes something from the ground)) uh look kid I-I put in a good word for you with Sheldrake in personnel
29b	<i>Father of the Bride</i> (1950)	oh imagine I said Edith where did I-- well anyway she just uh she just took like a duck to water, just like a uh -- when she was about six I started taking her out to the raft with the other kids and throwing her in
29c	<i>Michael Clayton</i> (2007)	Don said that I could uh -- .. I'm not exactly sure how this works

In this context, hesitations commonly occur at the end of a turn, as a symptom of a planning function that is sometimes interrupted by the speakers themselves. As pointed out by Clayman (2013: 153), turn-final hesitations may be employed by speakers to project continuation since the speaker continues to vocalize even if he or she is not adding anything to the substance of the turn. This allows speakers to conclude the turn without actually concluding, as is the case, for instance, in the scene extracted from the movie *Twelve Angry Men* (1957), where Juror no. 1 tries to bring back the order after an argument between Juror no. 5 and Juror no. 10. As can be seen from the transcription of the scene, Juror no. 1 concludes the turn with a hesitation marker and then starts a new turn directed to a specific interlocutor. The shift to a new interlocutor, however, does not occur immediately; the character hesitates to conclude the previous sentence directed to the whole group of Jurors gathered in the room and launches a new turn specifically directed to Juror no. 8 after a silent pause of 1.714703 seconds.

Extract 30 - *Twelve Angry Men* (1957)

Juror no.1	okay look let's stop the arguing we're only wasting time uh .. =
Juror no. 1 to Juror no. 8	= uh your turn down there let's go
Juror no.8	well I didn't expect a turn I thought you were all gonna to try to convince ME, wasn't that the idea?

More commonly, however, the speaker happens to be interrupted by another character taking the turn. This is the case in the following scene from *The Kids Are All Right* (2010). At dinner, Paul is telling the other characters about his trip to Mexico; the hesitation at the end of the first turn would be fulfilling a planning function, i.e. the character plans how to continue the storytelling of this vacation. The continuation of the storytelling, however, is interrupted by another character (Joni) taking the turn to ask a question. Consequently, Paul's turn remains unfinished as he begins a new turn to answer Joni's question.

Extract 31 - *The Kids Are All Right* (2010).

Paul	a big rattlesnake when I was in <i>uh</i> I was in the desert in <i>uh</i> in Baja Mexico we should-- I'll take you to Baja . I was in the-I was in the desert and <i>uhm--</i>
Joni	((offscreen)) what's it like there?
Paul	((offscreen)) oh it's beautiful and-- I was <i>uh--</i> I was surfing but then there was no waves so we climbed up into the-into the mountains [...]

This is not always the case; it may also happen that a character momentarily suspends the turn, interrupted by another speaker, and then continues the interrupted sentence in a different turn. In this regard, in the scene presented in extract 32, Juno confesses to her father (Mac) and her stepmother (Bren) that she is pregnant. As can be seen in the example, while discussing about the father of the baby (Paulie Bleeker), Juno momentarily stops to think about the correct way to phrase the sentence. This moment of pause is interpreted by Mac as a Transition Relevance Place (Sacks, Schegloff & Jefferson 1974), i.e. a point where he can insert himself into the conversation with the intention to stop Juno from finishing the sentence. Despite Mac's insertion, Juno completes the utterance previously left unfinished (*in chair*). This also commonly occurs when a listening character provides some sort of backchannel to the speaker such as *mm* or *mm-mh*.

Extract 32 - *Juno* (2007)

Mac	Paulie Bleeker?
Juno	what?
Mac	I didn't think he had it in him
Leah	((laughing)) I know right?
Mac	all right this is no laughing matter
Juno	no it's not! . and you know Paulie is actually great in <i>uh--</i>
Mac	okay
Juno	<i>in chair</i>
Mac	you're thinking about adoption?

In other cases, the sentence left unfinished is a symptom of what can be defined as “retrace-and-repair sequence” (Biber *et al.* 1999: 1062), where the character stops, plans and redrafts the whole sentence (Extracts 33):

Extracts 33 - Hesitations in Retrace-and-Repair Sequences

33a	<i>Annie Hall</i> (1977)	we-- <i>uh</i> you should have told me it was Tony Lacey's party
33b	<i>Heaven Can Wait</i> (1978)	I'll just <i>uh</i> I'm going to sit here for a minute
33c	<i>Kramer vs. Kramer</i> (1979)	well-- .. I <i>uh--</i> my French professor finally asked me out

33d	<i>In the Bedroom</i> (2001)	look you uh -- . can you just go now? I-I don't want you here when they get back
33e	<i>The Verdict</i> (1982)	well I-- it is a shame what happened to your husband mrs Dee is uh -- . I knew him vaguely through the Lodge he was a wonderful man

The planning process may also involve some delay in the continuation of the actual sentence; in this case, the hesitation marker is usually followed by a medium-long pause during which the speaker buys some time to decide how to continue (Extract 34a-b), to postpone the actual perpetuation of the sentence in emotionally demanding turns such as characters declaring their love to another character (Extract 34c), a character providing a dispreferred opinion (Extract 34d) or an apologizing turn. The latter case is reported in extract 34f, where the hesitation marker is used to plan and delay the follow-up phase in which the character of Ben justifies his behaviour (*I was kind of upset there*) before moving to the actual apologies.

Extracts 34 - Hesitations to Delay the Continuation of the Sentence

34a	<i>In the Bedroom</i> (2001)	I'll uh .. (2.063636 seconds of silence) I'll see you later then
34b	<i>Manchester by the Sea</i> (2016)	I guess you don't know this but I uh -- .. (1.972089 seconds of silence) I really kept in touch with Joe, +it's been kind of weird* for me not seeing Patrick--
34c	<i>Juno</i> (2007)	and also um ... (3.694110 seconds of silence) I think I'm in love with you?
34d	<i>Kramer vs. Kramer</i> (1979)	all right may sound a little rough . but uh .. (1.752000 seconds of silence) I think you ought to send Billy away to stay with relatives for a while
34f	<i>The Graduate</i> (1967)	well look uh ... (2.530136 seconds of silence) ((sighs)) I was kind of upset there .. I'm sorry I said those things

Sometimes, the delay may occur as a consequence of the character's need to take some time to think about how to continue or as a consequence of a character's decision-making process. An example of the former case is provided in extract 35, where Lieutenant Doug Roberts and the ship's doctor (Doc) help Ensign Frank Thurlowe Pulver recreate scotch with the ingredients available on the ship. Doc's hesitation at the end of the turn is a symptom of an ongoing thinking process; more precisely, Doc is thinking about the right referent for the comparison to answer the question posed by Mr. Roberts. An example of the latter case, instead, is provided in extract 36 from *Tootsie* (1982), where the turn-final hesitation marks the character's reasoning about whether to accept the invitation or not, interrupted by Julie's new turn.

Extract 35 - *Mister Roberts* (1955)

Mr. Roberts	Doc? . what does scotch taste like?
Doc	oh well it tastes like u:h-- . u:h--
Mr. Roberts	you know what it's always tasted a little like to me? . iodine

Extract 36 - *Tootsie* (1982)

Julie	what are you gonna do for the holidays?
Dorothy	why?
Julie	well, the baby and I are gonna go up to my dad's farm, upstate and it is not exactly the fast lane but it's-it's kind of fun, maybe you'd like to come along?
Dorothy	oh uh--
Julie	you know since my dad met you, he's your biggest fan

These moments of silence must also be interpreted within the entertaining context of cinematic products, where pathos is created to convey the intensity of the moment and the emotional involvement of the character on the screen, hence generating suspense and stimulating curiosity among viewers (see Kozloff 2000: 49-51; Bednarek 2018: 55). For this reason, the combination 'hesitations + pauses of silence' is often used at the beginning of longer turns in which characters usually prepare themselves for a long stretch of dialogue in which they get to speak without being interrupted (Kozloff 2000: 60), as in extract 37. Once again, these performances usually involve some kind of revelation or turning-point in the narration, as well as a character opening up about a personal situation or preparing to narrate something emotionally challenging, like the end of a marriage (Extract 37) or the accident that caused a house fire that killed the protagonist's children (Extract 38).

Extract 37 - *Kramer vs. Kramer* (1979)

Joanna	yes ... um .. (1.888000 seconds of silence) look .. during the last five years of our marriage ... I was becoming more and more unhappy, more and more troubled, and I really needed-- . somebody to help me, but when I turned to Ted he just-- .. wasn't there for me . so we became more and more isolated from one another more and more separate .. he was very involved in his career . and because of, his attitude towards my, fears and his inability to deal with my feelings .. I had come to have almost no self esteem . I was scared and I was very unhappy . and in my mind I had no other choice but to leave .. at the time I left I felt, that there was something terribly wrong with me . and that my son would be better off without me ... and it was only after I got to California that I realized . after getting into therapy, that I wasn't such a terrible person . and-- ((sighs)) . just because I needed some kind of, creative or em-emotional outlet other than my child that didn't make me unfit to be a mother

Extract 38 - Manchester by the Sea (2016)

Lee	<i>u:h</i> anyway . our bedroom is in the downstairs and-- the kids sleep upstairs so Randi . makes everyone leave around two o'clock . or three am . and then .. she goes back to bed ... so I go upstairs to check on the kids and it's fucking freezing upstairs . but I can't use the central heat cause it dries out Randi's sinuses and gives her really bad headaches . so I go downstairs and put a fire in the fireplace and sit down to watch TV but there's no more beer .. I'm still jumping around like a-- .. jackrabbit, so-- I put a couple fire logs in the fire and I-- .. just to warm up the house while I was gone ... and then, I-- am going to the mini mart but I'm too wasted and I don't want to drive . so I walk it's about twenty minutes each way, about halfway there, and I can't remember if I put the screen on the fireplace ... I figure it's okay ... so I just keep going to the store ... a:nd <i>u:h</i> that's it .. a log must've rolled out on to the floor . and-and the firemen said they pulled Randi out, she was passed out downstairs and uh ... (3.677714 seconds of silence) then the furnace blew, and they couldn't go back in again . and-- .. that's all I remember

Finally, before moving on to the next functional domain, hesitations can also be used as pause fillers employed to occupy a pause before continuing with the utterance or while some other action visible on the screen is carried out. In extract 39a, for instance, Juror no. 6 hesitates after initiating the answer to the Baliff's question "What is your name sir?", inserting a planning pause while looking for his name on a sheet of paper. Likewise, in extract 39b, the hesitation marker uttered by Arthur is used to fill a momentary pause in which the character is looking for notes about a joke in his notebook; or again, in extract 39c, a hesitation marker is used as a pause filler, verbalizing the fact that the hotel receptionist is looking for information regarding a reservation.

Extracts 39 - Hesitations as Pause Fillers

39a	<i>Twelve Angry Men</i> (1957)	oh it's u:h that one yeah
39b	<i>Joker</i> (2019)	my mother would say, ((mimicking his mum's voice)) you should enjoy it one day you'll have to work for a living . ((back to his voice)) no I won't ma I'm gonna be a comedian ((laughter)) ... and um .. here's one .. you know, I was just thinking the other day, why are the rich people so:--
39c	<i>The Hustler</i> (1961)	I'll <i>uh</i> see what I can do uh , <i>uh</i> you were right mister Gordon I mislaid your wire uh two adjoining suites?

In this context, hesitations are often used as single-turn items in between longer turns. This is the case, for instance, in the following extract from the film *Mississippi Burning* (1988), where Detective Word visits the local pastor and his son with the intention of gaining some information that could help him in the investigation of three missing civil rights workers. In

the following scene, after having established contact with the other character using the greeting *hello*, a hesitation marker is used to interrupt the silence (2.232923 seconds), filling the pause while getting closer to the interlocutors before asking them a question (preceded by a silence of 2.45125 seconds).

Extract 40 - *Mississippi Burning* (1988)

Word	hello .. uh .. I wonder if I could ask you a few questions

4.3.3. Hesitations in Film Speech: The Interpersonal Domain

The functions associated with the Interpersonal Domain mainly concern establishing and modulating characters' interactions on the screen. Thus, this functional domain commonly interests one character establishing contact with another character, as well as guaranteeing, or trying to guarantee, that the verbal exchange is carried out successfully, minimizing friction between participants.

One of the basic functions identified within the Interpersonal Domain regards the speaker wanting to attract the attention of the interlocutor(s) to whom he or she is directing the message. This function is commonly carried out by turn-initial hesitation markers used with a turn-opening function while also serving as an attention getter (for instance, extracts 41a-h). This also happens when the talking character launches a new turn directed to a new participant, that is, when a change of interlocutor occurs (41f). In this context, the pragmatic features observed here are commonly used in the pattern 'hesitation + vocative and/or familiarizer'.

Extracts 41 - Turn-Initial Hesitations to Grab Attention

41a	<i>Twelve Angry Men</i> (1957)	okay uh gentlemen <i>uh</i> I can have your attention? [...]
41b	<i>Separate Tables</i> (1958)	((offscreen)) uh miss Cooper? ((onscreen)) I've been waiting for you . it's a matter of vital importance I can't go into it here I merely ask you to read this
41c	<i>The Hustler</i> (1961)	hey uh mister? .. is it okay if I grab a cue?
41d	<i>Kramer vs. Kramer</i> (1979)	oh hey uh Ted come on in, have a drink!
41e	<i>The Verdict</i> (1982)	you honour ... u:h ladies and gentlemen of the jury [...]
41f	<i>Tootsie</i> (1982)	= uh Julie there's been an alert at your station =
41g	<i>The Kids Are All Right</i> (2010)	oh I forgot uh Nic I got a bitching nineteen ninety eight Alma Rosa that I think you're just gonna love
41h	<i>Silver Linings Playbook</i> (2012)	((offscreen)) uh Jake? you better get your brother

Even though this pattern typically occurs at the beginning of the turn, speakers also use it in turn-medial position so as to draw the listener's attention either to what has been said until then or to what the speaker is about to add. In other words, the combination 'hesitation + vocative' can be used with both an anaphoric (Extracts 42) and a cataphoric function (Extracts 41b and 41g). In the former case, this pattern is used to direct the listener's attention toward what the speaker has just said; in the latter case, this pattern is used by the speaker to direct the listener's attention toward what is about to follow.

Extracts 42 - Turn-Medial Hesitations to Grab Attention

42a	<i>The Apartment</i> (1960)	oh! how do you do uh mister Matusch--
42b	<i>Kramer vs. Kramer</i> (1979)	I won't hurt him please uh mister Kramer? <i>uh</i> can I talk to you for a second outside?
42c	<i>In the Bedroom</i> (2001)	= <i>uh</i> rephrase your question uh ms Keyes

In some cases, the attention-getting function is made explicit by the action on the screen; for instance, in extract 41i, the protagonist Arthur hesitates, grabbing the attention of Murray, who is walking away. The hesitation marker is therefore employed to re-establish contact and prevent the interlocutor from leaving the conversational exchange. The same occurs in the following scenes from *Twelve Angry Men* (1955) and from *Father of the Bride* (1950). In the former case, twelve jurors are gathered to discuss the fate of a young man accused of murder. Eleven jurors believe he is guilty, while one person (Juror no. 8) has a reasonable doubt. In the example below, Juror no. 6 confronts Juror no. 8 regarding his decision to vote not guilty. When Juror no. 8 starts walking away, Juror no. 6 employs a hesitation marker (preceded by a silence of 4.288000 seconds and followed by a silence of 2.928000 seconds) to attract the attention of the other character in order to keep the communicative channel open, i.e. with the intention to signal to the interlocutor that the interaction is not yet concluded, as there is still something that he wants to add.

In the latter case, during the wedding rehearsal, Stanley (the father of the bride) feels like he is the only one who cares about the success of the ceremony; in fact, rehearsals are hastily carried out, with no groom, missing guests and no Priest. When the Priest finally arrives at the church, he immediately runs away with the excuse of a meeting. Hence, when everyone starts leaving, Stanley tries to stop them, capturing everyone's attention by means of a turn consisting of a cluster of three hesitation markers.

Extract 43 - *Twelve Angry Men* (1955)

Juror no.6	ah he's guilty for sure, not a doubt in the whole world we should've- - we should've been done already ((sighs)) oh I don't mind you know ((chuckles)) beats working ... uh ... you think he's not guilty huh?
Juror no.8	I don't know, it's possible

Extract 44 - *Father of the Bride* (1950)

Priest to Stanley	and now I must run I hate to but I must and don't worry because everything will go beautifully tomorrow you see I know I've done it before
Stanley	uh uh uh--
Priest to woman	good luck my dear such a beautiful bride I'm sure you'll be very happy goodbye

These examples show that hesitation markers can be used to fulfil what Jakobson (1960) identified as the phatic function of the language, that is to say, “the endeavor to start and sustain communication” (*ibidem*: 356). More specifically, in the context of film dialogue, hesitations are also used to maintain the communicative channel open, ensuring that the interlocutor remains engaged in the conversational act, as in the case reported in extract 45 where, while talking on the phone, Jerry attracts Dorothy's attention to check whether she is still on the line.

Extract 45 - *Jerry Maguire* (1996)

Jerry Maguire (on the phone)	((offscreen)) Dorothy? uh hello?

Nevertheless, the Interpersonal Domain does not only include the phatic functions presented above; it also encompasses tasks related to linguistic politeness, i.e. “the language a person uses to avoid being too direct” or the “language which displays respect towards or consideration for others” (Watts 2003: 1). In this context, hesitations are used by speakers to plan and modulate certain types of speech activities such as an apologizing turn (Extracts 46a-c) or condolences (Extract 46d), but also requests, offers, invitations, suggestions and so on (Brown & Levinson 1987: 65-68; Leech 2014).

Extracts 46 - Apologies and Condolences

46a	<i>Heaven Can Wait</i> (1978)	hey uh I'm sorry, I didn't think Corkle
46b	<i>Tootsie</i> (1982)	I gotta go, uh listen uh forgive me for rushing off like this it was a wonderful dinner
46c	<i>The Kids Are All Right</i> (201)	um it's all right I just wanted to say um I'm-- I'm sorry about last night

46d	<i>Manchester by the Sea</i> (2016)	yeah I'm really uh I'm sorry to hear about Joe

Indeed, the speech activities listed above are usually carried out in an indirect manner by resorting to formulaic formulas, hedges and discourse markers, to name a few (Watts 2003: 70). Amidst these elements, hesitations also play an important function. More specifically, turn-initial hesitations are used to plan and launch indirect speech acts such as requests (in extract 47a, Tony invites another character, Betty, to explain the reason for her visit and in extract 47b, Galvin requests to talk with Mr. Alito over the phone), offers (Extract 47c, offering another cup of coffee), invitations (Extract 47d, an indirect invitation to go to Frank's cabin) and suggestions (Extract 47e, where the hesitation mitigates the suggestion to meet in person).

Extracts 47 - Hesitations to Mitigate Specific Speech Acts

47a	<i>Heaven Can Wait</i> (1978)	uh perhaps you would like to tell me what it's about
47b	<i>The Verdict</i> (1982)	oh u:h mr Alito please
47c	<i>The Hustler</i> (1961)	= would you uh like another cup?
47d	<i>Mister Roberts</i> (1955)	cause uh it just so happens I have a bottle of Red Label in my cabin
47e	<i>Jerry Maguire</i> (1996)	hey maybe uh the best thing is that we get together in person ((offscreen)) Rod uh I'm sorry would you mind holding +I just* got someone--

This function also occurs when hesitations are used in the middle of a turn; in this case, the mitigating function also partially overlaps with the planning function, as illustrated by the following extracts. In the examples below, the hesitation marker co-occurs with the redraft of the offer by switching from the modal verb *could* to *would* (48a), with the repetition of the preposition *to* in the request reported in extract 48b, as well as with the character looking for the right words to phrase the request, as in extracts 48c and 48d. In the latter case, the talking character is indirectly asking the protagonist, Ted Kramer, to leave the room so that he and his colleague can talk privately about Ted's job application.

Extracts 48 - Hesitations to Mitigate Requests

48a	<i>Father of the Bride</i> (1950)	I don't know how you feel about it but could you uh would you like a little snifter before dinner?
48b	<i>Separate Tables</i> (1958)	((clears throat)) it's just if you don't mind to uh to get my bill ready

48c	<i>Heaven Can Wait</i> (1978)	Joe if you could just um not so much lower but broaden your standards, we've seen a great many bodies and the choices are becoming limited
48d	<i>Kramer vs. Kramer</i> (1979)	<i>uh</i> mister Kramer could we talk uh privately for a moment?

In this respect, a turn may also be left unfinished as a form of linguistic politeness, thereby mitigating a turn that could potentially threaten the face of the interlocutor. This is the case in the following scene from *The Conversation* (1974): during a party, the surveillance expert Harry Caul is talking with some friends and colleagues about how he managed to succeed in his most recent job, i.e. intercepting the conversation between two lovers. After having revealed that the recordings were obtained through three special microphones specifically designed by Harry himself, the colleague William P. Moran asks Harry to have a closer look at the equipment. The use of the hesitation marker at the end of the turn allows the character to carry out the request in a more indirect manner, hence mitigating the speech act.

Extract 49 - *The Conversation* (1974)

William P. Moran	yeah, it sounds very pretty, I'd like to take a look at that mic too uh--

As can be evinced from the examples above, these elements are often employed in question/answer sequences. On the one hand, hesitation markers are used to modulate the interrogative turn when accomplishing offers and/or invitations; on the other, they are used to answer the received offer/invitation. In the latter case, hesitations are usually employed to mitigate a dispreferred answer that could potentially threaten the face of the interlocutor such as refusing someone's invite (Brown & Levinson 1987: 38; Leech 2014: 32). This is the case in extract 50a from *The Conversation* (1974), where the protagonist Harry Caul is invited for drinks by some colleagues while attending a convention. Harry mitigates the refusal of their invitation because he wants to talk to another character first. Similarly, in extract 50b, the indistinct hesitating sound is employed by Juno to hedge the refusal of Bleeker's invite ("Do you wanna come?") to hang out with him and some other friends after school, as can also be seen in extract 50c, where Lee refuses the offer of something to drink ("Could I get you another beer Lee?"). In this respect, while turn-initial hesitations are employed to launch the turn used to decline the offer and/or invitation, turn-medial hesitations are usually utilized to plan the explanation used to justify the non-acceptance. This is displayed, for instance, in extract 50d, where Major Pollock, having decided to leave Miss Cooper's hotel,

refuses her proposition to stay, and in extract 50e, where Clarice Starling refuses Doctor Chilton’s invitation to hang out with him in the evening. In both cases, the hesitation marker follows the coordinating conjunction *but*, which connects a first sentence addressing the request/invitation and the second sentence, which contrasts with the first one by executing the refusal.

Extracts 50 - Hesitations to Mitigate the Refusal of Invitations

50a	<i>The Conversation</i> (1974)	uh you guys go on without me I will have a talk to Stan
50b	<i>Juno</i> (2007)	oh ((indistinct sound)) . it sounds awesome but I got-I got my ultrasound
50c	<i>Manchester by the Sea</i> (2016)	uh I'm good . thank you
50d	<i>Separate Tables</i> (1958)	well that’s very good of you but uh of course I have to go
50e	<i>The Silence of the Lambs</i> (1991)	((chuckles)) well I’m sure this is a great town doctor Chilton but um my instructions are to-- talk to doctor Lecter and report back this afternoon

Likewise, in extract 51, Ted is invited to lunch by his boss; however, he is uncertain whether to accept or not, opting for a dispreferred response in the adjacency pair invitation/acceptance of invitation (Clayman 2002: 231). Consequently, even in this case, the speaker concludes the turn without actually concluding the explanation of the reason why he may not be available for lunch, hence mitigating his reluctance to accept the invite.

Extract 51 - Kramer vs. Kramer (1979)

Jim O’ Connor	Ted you got lunch?
Ted	oh hi Jim gee uh I don’t know I was trying to work this idea out here uh

Hesitations are also employed in film speech to carry out another potential face-threatening act, i.e. correcting the interlocutor regarding specific information (for instance, Laser correcting Paul for calling him Laze in extract 52a), a particular term (as in extract 52b, where Jerry corrects Dorothy who uses the term “memo” to describe the document about changing the sports management business) or the state of affairs, namely correcting the information provided by the interlocutor regarding a specific situation as it unfolds, for instance, in extract 52c from *Sideways* (2004), where Miles corrects the other characters’ belief regarding his book being published. As a matter of fact, correcting someone may constitute a threatening act toward the hearer’s positive face, defined as the “individual’s

desire that her/his wants be appreciated and approved of in social interactions” (Watts 2003: 86).

Extracts 52 - Hesitations to Mitigate the Correction of the Interlocutor

52a	<i>The Kids Are All Right</i> (2010)	((softly)) um it’s-it’s Laser ((mimics)) +yeah*
52b	<i>Jerry Maguire</i> (1996)	thanks, uh actually it was just a mission, statement
52c	<i>Sideways</i> (2004)	yes well uh it’s not exactly finalized yet so-- but there has been some interest, and--

Hesitations, together with some other linguistic features, can hedge the speaker’s utterance, mitigating the strength of the face threat. This is the case in the following line uttered by Juno in the homonymous film of 2007: “quite frankly *uh* this looks a little stupid”, where the hesitation marker *uh*, together with the adverbial phrase *quite frankly* and the hedge *little* are used to soften the face-threat. Likewise, hesitation markers are employed to downplay the connotative strength of a word, as in the case of the lexical item *gay* in the sentence “Madison? wait hold on w-- . isn’t that? like a little-- *uh--* gay?” from the same film.

Another way in which the speaker mitigates a potential threat to the interlocutor’s face is by refraining from imposing, thereby allowing to the interlocutor the freedom to choose, as in the line “we *uh* we don’t have to go” extracted from the film *In the Bedroom* (2001), where Matt asks his wife about spending the weekend at some friends’ house; or as in the line “and then I want you to *uh* to go out to the home and see the girl” extracted from the film *The Verdict* (1982), where the hesitation *uh*, in conjunction with the repetition of the particle *to*, serves to avoid imposing on the interlocutor.

More common, however, is the employment of these features to mitigate a momentary interruption in the conversation; for example, by suspending the ongoing interaction (Extract 53a). This most generally involves telephonic exchanges in which the character implicitly apologizes to the interlocutor for having to momentarily truncate the call (Extract 53b), sometimes asking the listener to wait on the line while talking with someone else. In the latter case, the hesitation marker commonly accompanies the formula *just a minute*, as reported in extracts 53c-d.

Extracts 53 - Hesitations to Mitigate the Interruption of the Interaction

53a	<i>Guess Who’s Coming to Dinner</i> (1967)	now uh you’ll-you’ll excuse us please
53b	<i>In the Bedroom</i> (2001)	I’ll uh I’ll call you back later okay? yeah, yeah, okay . =

53c	<i>The Apartment</i> (1960)	Vanderhof public relations, oh yes Baxter uh just a minute =
53d	<i>Guess Who's Coming to Dinner</i> (1967)	yes that's her ((laughter)) no uh just a minute =

Finally, in film speech, hesitations are also used to plan and mitigate the delivery of dispreferred news; that is, delicate announcements that the speaker must communicate to others, such as losing an important client (Extract 54a), not being selected for a job (Extract 54b) or having to get stitches after a bad fall (Extract 54c).

Extracts 54 - Mitigating Bad-News Announcements

54a	<i>Kramer vs. Kramer</i> (1979)	the uh Mid-Atlantic people have invited them to pitch the account
54b	<i>Tootsie</i> (1982)	uh we're looking for someone a little younger
54c	<i>Kramer vs. Kramer</i> (1979)	no he's not gonna lose it but I'm going to have to uh take some stitches

4.3.4. Hesitations in Film Speech: The Hesitation Proper Domain

As pointed out by Jucker (2015b), hesitation markers are employed in fictional texts to portray characters in a specific way and to delineate their communicative behaviours. In this regard, hesitations are stereotypically employed in film speech to convey nervousness, confusion, doubt and awkwardness, i.e. they are the manifestation of a state of uncertainty and insecurity in which characters are unsure of what to say. This is the case with the scene reported in extract 55, where Patrick asks his uncle Lee if his girlfriend can spend the night at their house. When confronted about the need to talk about safe sex, Patrick hesitates as a consequence of the momentary awkwardness due to the topic of the conversation.

Extract 55 - *Manchester by the Sea* (2016)

Patrick	so-- .. do you mind saying she stayed downstairs like if-- they-- call or something?
Lee	okay ... am I supposed to tell you to use a condom?
Patrick	((chuckles)) no I mean-- uh--
Lee	is that what Joe would say?
Patrick	no, well I mean yeah yeah we've had the discussion and everything . so--

In this functional domain, hesitations can be a reflection of the character's emotional state, expressing some sort of tentativeness regarding what to say and how to say it. In extract 56, for instance, hesitations are used to reflect Nic's uncertainty about how to conclude her turn

directed to Paul (the sperm donor through whom Nic and Jules gave birth to their children); that is, how to express what she wants to say in a nice way. Similarly, in extract 57, the character of Mrs. Drayton finds out that her daughter is engaged to a man of a different race. In this context, the hesitation marker reflects the character's state of shock and the difficulty of finding something to say.

Extract 56 - *The Kids Are All Right* (2010)

Nic	((chuckles)) oh yeah, no I was asking because I remember when I was reading your file back when we were looking for-- uh you know um the sperm

Extract 57 - *Guess Who's Coming to Dinner* (1967)

Mrs. Drayton	((chuckles)) well I-I um ((chuckles)) I-I suppose it would be all right if I said, my goodness! wouldn't it? ((chuckles)) . well my goodness

In this regard, hesitations can be used to avoid silence even though the character is uncertain about the topic of the conversation or how to actually contribute to the verbal exchange. Sometimes, these markers are also used to avoid providing an actual response, as in the case of extract 58 from *Heaven Can Wait* (1978), where Joe, recently reincarnated in the body of Leo Farnsworth (a multimillionaire industrialist), is talking with Betty Logan, an environmental activist critical of Leo's corporate policies. In this extract, Joe hesitates, uncertain about what the interlocutor is referring to and, as a result, unsure about how to reply. In addition, hesitation markers may also be employed as a form of reaction to a moment of difficulty or trouble provoked by the situation depicted on screen, as in extract 59 from *Joker* (2019), where Arthur accidentally fires a gunshot inside the house.

Extract 58 - *Heaven Can Wait* (1978)

Betty	I want to know what you're going to do about Pagglesham
Joe	uh--
Betty	PAGGLESHAM! PAGGLESHAM! it's what we're talking about!, [...]

Extract 59 - *Joker* (2019)

Arthur's mother	((offscreen)) Happy? what was that? are you okay?
Arthur	MOM! WHAT?
Arthur's mother	THAT NOISE! DID YOU HEAR THAT NOISE?
Arthur	I'M WA-I'M WATCHING AN OLD WAR MOVIE!
Arthur's mother	TURN IT DOWN!
Arthur	um

The character's difficulty is sometimes also expressed through the co-occurrence of hesitations with disfluency phenomena such as the stuttering false start in the question "why-uh what's the problem officer?" from *The Silence of The Lambs* (1991), or in the retrace-and-repair sequence in "what-d-- uh I mean-- . you guys are supposed to take care of this you know?" from *Juno* (2007). These features may serve to verbalize the self-interruption and self-correction processes carried out by characters with the intention of redrafting the sentence by reformulating the intended reply. In the narrative of a cinematic product, these elements may also be uttered to deliver some emotionally loaded turn, as exemplified in the speaker's admission of being heartbroken in extract 60.

Extract 60 - Juno (2007)

Bleeker	you're being really immature .. you've no reason to be mad at me I mean-- you know you broke my heart I uh -- I should be royally ticked off at you you know? I should be really cheesed off I shouldn't wanna-- talk to you anymore

In the context of cinematic dialogue, these elements are often used in sequences where the character makes a mistake, enhancing the comical effect of the overall scene. For instance, in extract 61, the hesitation marks the moment in which Stanley inadvertently repeats Ellie's words (*there's a lake*), then he immediately self-corrects and reformulates.

Extract 61 - Father of the Bride (1950)

Stanley to Ellie	= hey it's Aunt Hattie she's been at the station all morning wants to know how to get to the church
Ellie	we'll take the first car you'll follow with Kay
Stanley	Ellie! Ellie Ellie, Ellie! what am I gonna tell her now this is your relative
Ellie	tell her to go jump in a lake
Stanley on the phone with aunt Hattie	aunt Hattie? aunt Hattie go outside the station there's a lake-- uh there's-there's some taxis out there take one and go to the church I'll pay for it later, all right aunt Hattie ((hanging up the phone)) all right

Similarly, in the scene below from *Heaven Can Wait* (1978), Joe has just reincarnated in the body of Leo Farnsworth, who has been killed by his wife, Julia Farnsworth, and Leo's personal assistant, Tony Abbott. In the extract below, while Joe is talking with Betty, Leo's wife enters the room and, surprised upon discovering that her husband is still alive, screams in a state of shock. The opposition between the words *I don't frighten anybody* and Julia's yell creates a comical moment, which is even more emphasised by Joe having to explain to

Betty that the person showcasing this strong negative reaction when seeing him is his own wife. The explanation is provided through the hesitation *um* in the midst of stuttering repetitions, highlighting the character's tentativeness and unreadiness to face the situation.

Extract 62 - *Heaven Can Wait* (1978)

Joe	now miss Logan, I don't frighten anybody--
Julia	((yells)) OH MY GOD! ((yells))
Betty	what in god's name was that?
Joe	well that-that was m-- um my-m-m-m-my mrs Farnsworth

In this context, hesitations may be used to express a sense of awkwardness, which verbally contributes to the comical effect of a scene. For example, in extract 63 from *Kramer vs. Kramer* (1979), after spending the night together, Phyllis (Ted's colleague) wakes up in Ted's apartment. While going to the toilet, she meets Ted's son, Billy, in the hallway. Phyllis is completely naked and embarrassed by the meeting. This sense of awkwardness is conveyed by the turn-initial hesitation used to interrupt a moment of awkward silence, trying to articulate something (two false starts in sentences left unfinished: *well I--* and *it's really--*); when uttering the sentence *it's been nice seeing you*, the difficulty and awkwardness of the character is conveyed by the hesitation marker *uh*, which co-occurs with the repetition of the dummy subject *it*.

Extract 63 - *Kramer vs. Kramer* (1979)

Billy	do you like fried chicken?
Phyllis	fried chicken? very much
Billy	so do I
Phyllis	um . well I-- it's really-- . it uh it's been nice seeing you and <i>uh</i>
Billy	bye
Phyllis	I'm-- bye

Thus, hesitations related to this functional domain are commonly employed to add a sense of doubt, uncertainty and trouble to the turn. These meanings are often added at the beginning of the turn, where markers of hesitation proper also fulfil a turn-taking and/or a turn-launching function, in addition to answering questions. In the former case, these elements are used to launch a turn while expressing the character's reluctance (Extract 64) or perplexity for not understanding the real intentions of the interlocutor (for example, when Baxter does not understand that his boss is asking him to use his apartment to meet his mistress in extract 65), as well as the character being puzzled by a situation. This last case is

displayed in extract 66, where Pat's turn is initiated by a hesitation marker expressing confusion towards his dad's actions.

Extract 64 - Juno (2007)

Juno	((sighing)) + em * so I'm not really sure how I'm gonna spit this out

Extract 65 - The Apartment (1960)

Mr. Sheldrake	Baxter you're not reading me, I told you I have plans
Baxter	u:h so do I I'm-- . going to take four aspirins and get into bed so you might as well give the tickets to somebody else

Extract 66 - Silver Linings Playbook (2012)

Dolores	your father wanted you to have this
Pat	he wants me to have this Eagles information?
Dolores	it's his living
Pat	um . tell dad thanks

As briefly mentioned above, hesitation markers are also used by actors to express some sort of difficulty in answering a question and provide unsolicited, troublesome information, as in the case reported below (Extract 67) from the film *Sideways* (2004), where Jack, after meeting Stephanie, confesses to his friend, Miles, his doubts about marrying Christine. Likewise, in extract 68 from *Guess Who's Coming to Dinner* (1967), John, a 37-years-old black man, is talking on the phone with his father. He has to inform his parents about a sensitive topic, namely being engaged to a young white woman. Knowing that this is a delicate topic, he hesitates about how to break the news to his parents, with various phrases that momentarily delay the problematic information.

Extract 67 - Sideways (2004)

Jack	oh well you know just-- the wedding
Miles	what?
Jack	((indistinct hesitating sound)) I've just been doing some thinking about it

Extract 68 - Guess Who's Coming to Dinner (1967)

John	well ((coughs)) . ((sighs)) I guess I should have called you back again because there is this one thing I-I uh I should have-- I've been meaning to write to you about it ((clears throat)) . yeah there was one thing I didn't explain dad . and . I'm afraid it's gonna be kind of a shock

Items pertaining to the hesitation proper domain are also employed to plan or launch an answer that is actually a lie; for instance, in extract 69, the character Mister Roberts asks Frank (his roommate) if Dolan, another member of the ship's crew, had left a letter in their quarter. Frank lies to Doug (i.e. Mister Roberts), saying that he knows nothing about such letter, while in reality he found it, read it and hid it. It may also happen that the tentativeness is conveyed through the prolongation of sounds; this is the case, for instance, in the scene where a soldier is inventing an excuse to avoid working, i.e. he is hearing a ringing bell in his head. The Doctor of the ship knows that he is lying and makes fun of him by asking him how the bell sound is supposed to go: “dong dong, ding ding or ding dong?”. The answer “*uh . din:g u:h DONG?*” includes two hesitation markers, which reflect the character's thinking process to answer the question, while the prolongation of the nasal sound in “ding” conveys that sense of tentativeness commonly associated with the hesitation proper category.

Extract 69 - *Mister Roberts* (1955)

Mr. Roberts	= hey Frank? . Dolan been in here with my letter?
Frank	I uh don't know uh Doug boy I just came in here myself

Likewise, in extract 70 from *The Apartment* (1960), hesitation markers are used in planning a lie; the hesitation marker component can be seen as a value added to the planning component, which is used to convey a temporary moment of difficulty in formulating an excuse to explain the situation to another character, i.e. what happened to Fran Kubelik. In this scene, Baxter comes back to his apartment and finds that Fran passed out on his bed; she tried to kill herself by swallowing a large amount of sleeping pills after being left heartbroken by his lover, Mr. Sheldrake. As soon as he finds out about Fran, Baxter seeks the help of his neighbour, Dr. Dreyfuss. When asked about what happened, Baxter has to lie to the doctor; he cannot reveal to his neighbour that he lets his colleagues borrow his apartment for their extramarital affairs. He has to come up with an explanation, a lie that can justify what happened, while also implying that no other man was responsible for the incident.

Extract 70 - *The Apartment* (1960)

Dr. Dreyfuss	wanna tell me what happened?
Baxter	((sighs)) I don't know I mean I was not even here ... you see we had uh uh some words earlier it-it was nothing serious ((chuckles)) you might call it a lovers' quarrel

Hesitations are sometimes also used to avoid answering questions; for example, in the following extract from the film *Jerry Maguire* (1996), Bob Sugar, one of Jerry's co-workers, enters Jerry's office while he is in the middle of a meeting with Marcee, the wife of one of Jerry's clients. When Jerry introduces Bob to Marcee, she asks him whether he is the person who has called Marcee's house in the past, presumably to steal the client from Jerry. When directly asked about the call, Bob hesitates (hesitation proper) to avoid answering the question; he rather quickly leaves the room, apologizing for the interruption.

Extract 71 - *Jerry Maguire* (1996)

Jerry Maguire	Marcee this is one of our agents this is Bob Sugar who needs to learn to knock
Bob Sugar	pleasure
Marcee	you've called our house right?
Bob Sugar	um I'm sorry to interrupt you guys

4.3.5. Hesitations in Film Speech: The Cognitive Domain

The functions related to the Cognitive Domain are the least frequent in film speech. In this context, hesitations are commonly used to express a cognitive effort made by the speaker to provide specific information to the interlocutor and, indirectly, to the audience. In particular, within this functional domain, hesitations are usually employed to vocalize the thinking process carried out by the speaker. This process may be triggered by a question the character needs to answer, thereby reflecting the mental effort carried out by a character thinking about the information necessary to reply (Extract 72).

Extract 72 - *Driving Miss Daisy* (1989)

Daisy	of course! . er, er er er er, that's the last part, BauER . what letter sounds like er?
Hoke	uh ehm uh, r?

Sometimes, this effort occurs as a consequence of the character trying to remember something; that is, to recall exact information, as in the following extracts. More specifically, in the scene reported in extract 73, after a head concussion, the football player Steve Remo cannot remember his name. When asked if he remembers his name, the use of the hesitation marker aims to convey the thinking effort carried out by Steve to recall the correct information in answering the question. In this case, the hesitation marker also presents a

hesitation proper component, reflecting the character's sense of uncertainty and lack of knowledge. In extract 74, instead, the character is struggling to remember details to answer the interlocutor's question. Interestingly, in this scene from *The Graduate* (1967), the turn-medial hesitation marker preceding the additional information regarding the directions to reach Allan Street co-occurs with a prolongation of the initial sound in the conjunction "and", highlighting the thinking process carried out by the character.

Extract 73 - *Jerry Maguire* (1996)

Doctor	do you know your name?
Steve Remo	uh

Extract 74 - *The Graduate* (1967)

Ben to the Mechanic	= ALLAN STREET WHERE IS IT?
Mechanic	<i>uh</i> Allan uh it's <i>uh--</i> it's <i>uh--</i> i-it's six blocks up <u>a:nd</u> uh three over . <i>uh</i> YOU NEED ANY GAS FATHER?

Likewise, this cognitive effort also manifests when characters are required to execute complex tasks, as for instance, making calculations; this is the case in extract 75, where Juror no. 2 is asked to report the time required by Juror no. 8 to carry out an action.

Extract 75 - *Twelve Angry Men* (1957)

Juror no. 8	what's the time?
Juror no. 2	u:h exactly, forty one seconds

Precise information may relate to details or factual data regarding things that happened in the past, as in extract 76 from *The Kids Are All Right* (2010), where Laser asks Paul how much he used to get paid to be a sperm donor, or as in extract 77, where Jack hesitates while recalling details about a commercial he did in the past.

Extract 76 - *The Kids Are All Right* (2010)

Laser	how much did you get paid?
Paul	why do you want to know that?
Laser	I'm just curious
Paul	I got paid u:h sixty bucks a pop

Extract 77 - *Sideways* (2004)

Miles' mother	and what was that other one you did? the one where you were the jogger
Jack	oh that was for u:hm spray and wash

Interestingly, as illustrated in the examples provided in this section, hesitations fulfilling functions linked to the cognitive domain often appear in answering questions, also occurring as turn-initial hesitations. However, these features more commonly occur in turn-medial position (22 turn-medial tokens as opposed to 8 turn-initial tokens), where they are mainly employed as a moment for the character to stop and reflect to retrieve precise information and details from their memory.

Most notably, hesitation markers are showcased by a character trying to remember the exact term to identify the referent — namely, *fast pass* in extract 78a and *psychiatrist* in extract 78b — or when trying to remember another character’s name (Extracts 78c-e).

Extracts 78 - Turn-Medial Hesitations to Recall Precise Information

78a	<i>In the Bedroom</i> (2001)	uh yeah yeah sometimes, sixty minutes and upwards, unless of course you got the um-- uh fast pass
78b	<i>Kramer vs. Kramer</i> (1979)	she told me she was seeing a-a-a shrink a-a uh a psychiatrist or something
78c	<i>Father of the Bride</i> (1950)	oh you know that boy uh Buckley something or other I don’t know his last name
78d	<i>The Hustler</i> (1961)	you uh-- you ever been to Louisville during Derby week miss uh Packard?
78e	<i>The Apartment</i> (1960)	just before me there was miss Rossi in auditing, and after me there was miss Koch in disability and right before you there was miss uhm oh what's her name? on the twenty fifth floor

In this regard, hesitations also appear as turn-final elements, as exemplified by the following scene from *The Graduate* (1967) (Extract 79), where Mr. Robinson self-interrupts as a symptom of recalling information (i.e. Ben’s name) until Ben himself takes the turn, filling the pause with the necessary details to continue.

Extract 79 - *The Graduate* (1967)

Mr. Robinson to Mrs. Robinson	I uh-- I was just telling uh--
Ben	Ben

Depending on the narrative context, this thinking symptom may be explicitly fictitious, as in *The Hustler* (1961) (Extract 80), where the character of Bert pretends not to remember Findley’s name. In this context, the hesitation at the end of the turn fulfils a double function: on the one hand, it explicates the fact that Bert is trying to recall precise information, while,

on the other, it fulfils a turn-yielding function, asking the interlocutor to step in and contribute with the necessary information.

Extract 80 - *The Hustler* (1961)

Bert to Eddie and Sarah	= <i>uh</i> miss Packard, Eddie Felson James <i>uh--</i>
Findley	Findley

The considerations made thus far have been based on the ‘manual’ observance of hesitations in film speech, a non-trivial procedure built on the self-interpretation of the phenomenon in context. Some of the points made in the qualitative analysis have been further inspected and verified through a computer analysis of cinematic dialogue. More specifically, hesitations have been observed in film speech by means of collocation and concordancing using the AntConc software.

4.4. A Corpus-Linguistic Approach to Hesitations in Film Speech: Collocation and Concordancing

Having established the most common functions hesitations fulfil in film speech, a corpus-linguistic software has been utilized to inspect whether recurrent patterns can be found in the employment of hesitations in film dialogue. This has been achieved through the AntConc software. More precisely, AntConc has been employed using the Collocate Tool and the KWIC (Key-Word-In-Context) Tool.

4.4.1. Collocates of Hesitations in the Corpus

Collocation refers to the non-random combination of words, identifying which other words commonly co-occur with the node in a given co-textual span (Bednarek 2018: 100). More specifically, by conducting a query through the tagging label *hes* in the search bar, it was possible to obtain which words co-occur with hesitation tokens within the corpus sorted by Likelihood.

Firstly, a window span of 4 words to the left and 4 words to the right of the node has been set. This is the range typically used by researchers to investigate the collocates of the node (Jones 2022: 132; see also Sinclair 1991; Sinclair *et al.* 2004). Nevertheless, as shown in the previous sections of the current chapter, because hesitations tend to be functionally tied to their immediate surroundings, a narrower window has then been set (2 words, both to

the right and to the left) in order to obtain a clearer picture of their immediate context. In both searches, the minimum frequency has been set to 5 in order to attain repeated patterns, while the minimum range has been set to 2 in order to focus the analysis on the pragmatic role of hesitations according to their adjacent context.

	Collocate	Rank	Freq(Scaled)	FreqL	FreqR	Range	Likelihood	Effect
1	to	1	7280	120	111	27	34.147	-0.507
2	well	2	2664	87	95	27	29.269	0.600
3	yeah	3	1704	63	63	20	27.985	0.714
4	oh	4	2000	91	46	23	22.249	0.603
5	mister	5	544	12	38	14	21.591	1.028
6	listen	6	432	24	18	15	20.788	1.109
7	hi	7	240	18	10	12	20.223	1.372
8	thanks	8	184	12	10	11	16.614	1.407

Figure 4.8. Collocates of Hesitation Markers in the Corpus (4L-4R)

	Collocate	Rank	Freq(Scaled)	FreqL	FreqR	Range	Likelihood	Effect
1	to	1	3640	45	27	22	69.387	-1.190
2	well	2	1332	54	70	26	55.557	1.044
3	t	3	1940	26	11	21	39.103	-1.243
4	i	4	9992	183	389	28	33.440	0.343
5	yeah	5	852	36	41	16	31.872	1.001
6	listen	6	216	16	15	13	31.540	1.669
7	oh	7	1000	58	28	21	31.136	0.930
8	a	8	3692	48	55	23	29.878	-0.694
9	be	9	816	12	0	10	23.620	-1.618
10	hey	10	312	23	12	12	23.449	1.313
11	mister	11	272	8	24	12	23.448	1.382
12	of	12	2040	29	24	23	20.601	-0.797
13	hi	13	120	12	6	9	19.519	1.732
14	okay	14	424	32	7	13	16.840	1.027
15	really	15	428	3	2	4	15.648	-1.950
16	ll	16	676	20	33	18	14.406	0.796
17	in	17	1780	32	18	16	13.928	-0.684
18	for	18	1120	11	16	16	13.833	-0.905

Figure 4.9. Collocates of Hesitation Markers in the Corpus (2L-2R)

Results show that hesitations in film speech typically co-occur with a range of items: prepositions, discourse markers, interjections, as well as greetings, response forms and verbal elements.

The fact that hesitation markers are commonly used in accordance with *well*, *okay* and *oh* could be in line with the role attributed to the use of the pragmatic features under investigation as possible utterance launchers and planning signals. As a matter of fact, the

discourse marker *well* is described as an element introduced by speakers to give thought to the content of the discussion, a turn initiator “serving to relate a speaker’s response to the ongoing conversation” (Biber *et al.* 1999: 1086-1087); hence, it can mark continuation or it can be employed as an element used to reflect upon the choice of expression (*ibidem*: 1087). *Okay* is frequently used as a positive backchannel, as well as “to indicate a change of topic or to signal that a new stretch of discourse is beginning” (Carter & McCarthy 2006: 116); hence, to help planning and organizing the speech. The interjection *oh*, instead, is described as an element used to initiate utterances, to respond to the sentence uttered by someone else or to introduce or reply to a new piece of information (Biber *et al.* 1999: 1083-1084).

Along these lines, the co-occurrence of hesitations with other discourse markers such as *listen* also uncovers some interesting data regarding the way hesitations are employed in film speech. In spontaneous conversation, *listen* can be used as a discourse marker when employed in the imperative form to grab the listener’s attention; in other words, it is used “to prepare the listener for a new topic or for a shift in the current focus of the topic” (Carter & McCarthy 2006: 211). In short, to focus the listener’s attention on what the speaker considers important (*ibidem*: 218). In addition, hesitation markers commonly accompany the agreement marker *yeah*, generally used as a response form in answers to questions (Biber *et al.* 1999: 1089-1090), as marks of agreement or simply as tools to keep the conversation going; with greetings such as *hi* or *hey* — hence, in opening conversations — and with polite speech-act formulae such as *thanks* (although, in a less immediate context).

The results obtained from the current inquiry seem to be in line with the features previously highlighted in the qualitative functional analysis of hesitations; this seems to suggest that, in film speech, these features routinely come with items that conventionally fulfil initiating or responding functions within an adjacency-pair framework of utterance-response sequences in conversation (Biber *et al.* 1999: 1046), or with elements that have been described as “interactive conversational inserts” such as interjection, response forms and discourse markers (*ibidem*). In addition, hesitations appear to be associated with discourse management functions — favouring, for instance, a change of topic — which is also shared with vocatives or address forms, which are, broadly speaking, “attached to the beginning or end of a clausal or non-clausal unit” (*ibidem*: 1047). This can be evinced by the co-occurrence of hesitations with the vocative *mister*, as a respectful form of address.

Finally, the pragmatic features under study are also used in film speech in conjunction with grammatical features such as the verb *to be*, the negative form of verbs — as signalled by the *t* collocate in Figure 4.9 as a sign of the contracted form ‘verb + not’ —

and future tense verbs, as signalled by the *ll* collocate in string 16 (as the contracted form of ‘will’). Hesitations also co-occur with function words such as the pronoun *I*, the article *a* and, most notably, with prepositions such as *to*, *of*, *in*, and *for*. Given that these elements commonly serve to disclose relationships between lexical items or between larger linguistic units (Biber *et al.* 1999: 55), their co-occurrence with *uh* or *um* is also an indication of the planning function hesitations fulfil in fictional discourse; namely, planning the construction of the utterance, as well as the construction of the meaning of a text (how units should be interpreted in discourse). More precisely, the fact that hesitations accompany the first-person pronoun *I*, for instance, is a reflection of the fact that these elements are commonly employed by speakers as tools to prepare and organize the content of the speech, which concerns self-referential facts, thoughts or feelings. Moreover, *I* in conjunction with hesitations, is often accompanied by the adverb *really* to emphasize and intensify what the speaker is about to say.

To conclude this section, the data presented in the current segment obtained from the AntConc Collocate Tool seem to be in line with the discursual functions of hesitations presented in the previous sections (the functional examination provided in §4.3 according to which hesitations mainly fulfil functions related to the Planning and the Structural Domain), namely allowing transition between adjacent parts of the discourse, as well as signalling transition in the interactive development of the speech. Nevertheless, this aspect needs to be further investigated by looking at the most common patterns in which hesitations occur in the corpus so as to observe their role in larger units, i.e. the clause. This is achieved through the Key-Word-In-Context tool.

4.4.2. KWIC (Key-Word-In-Context) Tool

Another important corpus technique for the qualitative analysis of the language is concordancing, that is, a tool that allows to observe the occurrences of a particular word (the node) with its surrounding co-text and to investigate the meanings and uses of such key item (Bednarek 2018: 102). By using the KWIC format, it was possible to monitor how hesitations are commonly used in a corpus of texts, as well as to identify the most commonly occurring patterns in the target corpus. The analysis of the results in a concordance has also enabled to double-check the qualitative considerations made in the previous sections regarding the employment of pragmatic features in film speech. The analysis has been

carried out by sorting twenty-five random hits of the node *hes* both to the right and to the left (organized by frequency, i.e. the most recurrent patterns in the corpus).

The analysis carried out by looking at the elements coming after the hesitation marker (sorted to the right) does not highlight any recurrent pattern; in fact, the context in which hesitations occur is quite variable. Figure 4.10 demonstrates how hesitations are often used by characters to initiate a sentence in which they describe situations, facts or actions (*I just got held up, I couldn't walk, I was talking, I'll write*), comments (*I know that, I know about*), intentions (*I just wanna make*) or emotions (*I love you, it's terrible*) associated with the speaking character; hence, using the first-person subject pronoun *I*. Moreover, the majority of *hes* seems to occur at the beginning of coordinate clauses, preceding the conjunction *but*, and before the demonstrative *the*. Hesitations also appear to accompany other hesitations, signalling the use of these elements in moments of planning difficulties. Interestingly, the findings seem to support Biber *et al.*'s (1999) statement that “a filled pause is most useful at a point of grammatical incompleteness which is nevertheless a major planning point, such as the beginning of a dependent or coordinate clause” (*ibidem*: 1054). As a matter of fact, hesitations tend to co-occur with features like nominative personal pronouns (*I*) and demonstrative articles, i.e. elements used to mark the beginning of a new clause structure; this location “might be expected to be a fairly major forward-planning site in the course of syntactic composition” (*ibidem*).

	File	Left Context	Hit	Right Context
1	1982_THE ...	ah! ah! . is that right? ((whispering)) oh my god I--	hes	I just got held up hi hes Frank Galvin
2	2004_SIDE...	I would have to say that I prefer the dark	hes	I just wanna make one quick stop hes oh
3	1978_HEAV...	hes you want some tea? hes I-- I-I kno--	hes	I know about Pagglesham and hes-- well that-that
4	2010_THE ...	rock, and hes I thought great ((gets confused and undistinguishable))	hes	I know that hes you and Joni heard, what
5	1974_THE ...	was paralyzed in my left arm and my left leg	hes	I couldn't walk for six months . one doctor
6	2004_SIDE...	waited on me ... hes and while I'm at it .	hes	I guess maybe you should know that, my book
7	1996_JERRY ...	I'm gonna see you this weekend, you know? and	hes--	I love you okay . look hes-- no no, it'
8	1950_FATHE...	please . can I help? it'll speed things up I	hes	I was talking to our accountant he's married
9	1977_ANNI...	brother's crazy, he thinks he's a chicken, and	hes	the doctor says, well why don't you turn
10	1977_ANNI...	it's a negative thing! you-you know bugs are	hes	the entomology is a rapidly growing field! ((chuckling)) +yeah--*,
11	1960_THE ...	what's her name? on the twenty fifth floor well	hes	the mirror i-is broken yes? .. hes just a
12	1978_HEAV...	we got a coup-couple of pretty good backups yeah	hes	the point is we don't need a quarterback, ((
13	1957_12 ...	the question whether we're a hung jury or not	hes	the reason I asked about that was because, you
14	2010_THE ...	well you know I mean I studied architecture in college	hes	but I'm not really an architect I mean
15	1982_THE ...	the staff of anesthesiology-- yes yes I've heard that	hes	but you do hes testify quite a bit against
16	1982_TOOT...	and we'll talk about it I mean hes-- hes	hes	hes go home have you been drinking? can I
17	2016_MANC...	him what time-what time should I come back? hes,	hes	hes thanks mom we'll be right down good
18	2007_JUNO...	kid for like five years no I-- I know . and	hes	it's just that I'm not ready to
19	1955_MISTE...	the old appendix again huh Doc? well yeah it's	hes--	it's terrible! how do they go hes dong
20	1958_SEPAR...	don't see you again before I go I'll	hes--	I'll write and say goodbye no Sibyl . no
21	2007_JUNO...	fourth mh I didn't notice it at all actually	hes	I'm more concerned about when they have to
22	1978_HEAV...	marshmallow would last a little longer, and I'd say	hes	so do I hes sometimes he talked about cookies
23	2010_THE ...	know more about you guys e-e-everything I mean	hes	so what-what do you-what do you do
24	1988_MISSI...	of things I needed to check with you yeah well	hes	actually it was you I wanted to talk to
25	2007_JUNO...	this looks a little stupid have you guys thought of	hes	any names for the baby yet? hes sort of

Figure 4.10. Twenty-Five Random Instances for *hes* Sorted to Right

As for the investigation of the elements that more regularly precede hesitation markers (sorted to the left), Figure 4.11 shows that the node is often used after the personal pronoun *I* — thus, in self-referential constructions aimed at focusing the attention of the interlocutor on facts, thoughts or feelings that concern the speaking character — and the utterance launchers *so* and *okay*, which are also frequently used in spontaneous conversation to mark “the beginning or end of a topic or a transition from one topic or bit of business to another” (Carter & McCarthy 2006: 218). As previously noted, even in this case, hesitations also tend to appear in combination with discourse markers, like *you see* and *listen*, conjunctions (*and* or *but*) and grammatical features such as the demonstrative article *a*, the possessive adjectives *my* and *her*. Hence, as in impromptu conversation, even in film speech, hesitations contribute to the unfolding of the discourse. For instance, the fact that hesitations accompany conjunctions is a reflection of the fact that these elements are commonly employed as tools to extend the body of the utterance through what has been defined as an “‘add-on’ strategy” (Biber *et al.* 1999: 1078) used to “add one clausal unit to another” (*ibidem*) through coordination.

Moreover, once again, some of the results in Figure 4.11 seem to support the claim that hesitations are more commonly used at clause boundaries, signalling the end of the previous utterance and the initiation of a new one, as for instance, in the strings number 8, 20, 22 and 24; in other cases, these features are used to signal a link between one part of a topic and the next part, as for instance, in the strings number 17 and 18.

	File	Left Context	Hit	Right Context
1	1955_MISTE...	if you were to come over alone this afternoon-- I	hes	don't know hes Doug boy I just came
2	1960_THE ...	for a friend .. good night mrs Lieberman oh no I	hes	seem to have dropped my keys ((chuckles)) ah here
3	2001_IN TH...	the ten hes-- Ruth ... what I said just now .. I	hes ..	I-I had no right hes .. no one .. no
4	2010_THE ...	about your social life? no no it's okay I	hes	no I've never been-- never been married nor
5	2010_THE ...	mmy: hes-- hey! don't be a zoomer hes so	hes .	how are you? well how are you? I'm
6	2010_THE ...	another bottle of the seavey cabernet? thank you hey so	hes	have you guys gotten into the whole acai smoothie
7	2007_JUNO_...	chuckle)) that was magical ... ((whispering)) +thank you, thank you* so	hes	I've been listening to that really weird cd
8	1961_THE ...	hes Packard? I'll hes see what I can do	hes ,	hes you were right mister Gordon I mislaid your
9	1982_THE ...	yes yes I've heard that hes but you do	hes	testify quite a bit against other physicians isn't
10	1978_HEAV...	hes beats me I guess we bribed somebody hi ... hes	hes ... ((	clears throat)) hi how you doing everybody? I hes
11	1967_GUES...	told her? did she raise any objections? you see hes	hes	Matt there--there's sort of a special, problem
12	1957_12 ...	guilty I-I mean somebody saw him do it okay	hes	here's what I think and I have no
13	2004_THE ...	today . hey ... I'm glad I did it right, okay	hes	this is our plan, I don't think you
14	2010_THE ...	working on it for a while I also have a	hes	a little organic co op farm just up the
15	1974_THE ...	payphone cabin)) and hes I-I have the material and	hes	I-I'm calling for an appointment no no
16	1967_GUES...	a little girl . happy all through school and college ... but	hes	I don-I don't think I've ever
17	1982_THE ...	hour hello? . oh yeah thanks for calling ((chuckles)) Frank Galvin .	hes	I'm representing Deborah Ann Kaye . well hes I'
18	2004_SIDE...	yeah!, well you know Jack is pretty eager to go	hes	to get up to hes . hes hes you know
19	1982_THE ...	card my sister? you know how beautiful she was .. her	hes-- .	husband left her and he took her kids ((sighs)) .
20	1960_THE ...	to have dropped my keys ((chuckles)) ah here it is!	hes	sure sure hes-- oh! hes look kid I-I
21	2010_THE ...	hey let me hit you back that was nobody just	hes	just clay that's terrible .. hey hes have you--
22	1982_THE ...	locomotion of-- . well in short everything which constitutes her life ,	hes	we will prove, that she was deprived through negligence ..
23	1967_GUES...	will you just stop and think about that? .. oh listen	hes	what are we having for dinner tonight? gotta make
24	1967_THE ...	wait a minute hes-- you talked to Elaine this morning?	hes	you mean she doesn't know that you're
25	1996_JERRY ...	loving it I'm loving life yeah I drop my	hes-- ...	tried to make it on the ice out there

Figure 4.11. Twenty-Five Random Instances for hes Sorted to Left

To conclude, the results obtained from the current analysis seem to be in line with some of the findings previously reviewed on the topic (Chapter One): namely, hesitations tend to be used before function words (Maclay & Osgood 1959: 34) and before structural units (Blankenship & Kay 1964: 369), helping speakers in the process of building up their utterances, i.e. as markers of conversational strategies used to move forward the on-screen narration.

Having observed the way hesitation markers are employed in Anglophone film speech, the next chapter will focus on the way these elements are treated in Italian dubbing through a comparative quantitative and qualitative analysis.

Chapter Five

Hesitations in Italian Dubbing: A Diachronic Analysis

The aim of this chapter is to describe the results derived from the analysis carried out on the Italian section of the corpus. Firstly, the chapter will focus on the discussion of the data obtained from the quantitative analysis by comparing these results with those extracted from the source Anglophone texts. Subsequently, the qualitative analysis of the translation strategies used to render hesitations in Italian dubbing will be reviewed.

5.1. Hesitations in Italian Dubbing over Time

5.1.1. Hesitations in Italian Dubbing over Time: A Quantitative Analysis

This section aims to investigate the use of hesitations in Italian dubbing, as well as to ascertain whether it is possible to observe an evolution in the way these elements are treated in the process of being translated, adapted and performed by the dubbing actors. In order to do so, the original Anglophone soundtracks have been compared with the Italian dubbed version of each title. The Italian section of the corpus comprises a total of 317,392 words. Within the overall Italian corpus, 1342 tokens of hesitation markers have been detected, corresponding to 0.42% of the overall Italian corpus and to 67% of the hesitations found in the Anglophone section of the database.

Table 5.1. Tokens of Hesitations in the Italian Section of the Corpus Compared to the English Section			
Year of Release	Original Title	Tokens in the English Transcription	Tokens in the Italian Transcription
1950	<i>Father of the Bride</i>	100	57
1955	<i>Mister Roberts</i>	52	30
1957	<i>Twelve Angry Men</i>	94	44
1958	<i>Separate Tables</i>	34	29
1960	<i>The Apartment</i>	60	100
1961	<i>The Hustler</i>	41	23
1967	<i>Guess Who's Coming to Dinner</i>	89	73
1967	<i>The Graduate</i>	31	22

1974	<i>The Conversation</i>	70	18
1977	<i>Annie Hall</i>	132	73
1978	<i>Heaven Can Wait</i>	131	127
1979	<i>Kramer vs. Kramer</i>	104	42
1982	<i>Tootsie</i>	91	84
1982	<i>The Verdict</i>	129	69
1988	<i>Mississippi Burning</i>	27	16
1989	<i>Driving Miss Daisy</i>	20	26
1991	<i>The Silence of the Lambs</i>	56	22
1994	<i>Forrest Gump</i>	11	10
1996	<i>Jerry Maguire</i>	60	50
1999	<i>American Beauty</i>	19	13
2001	<i>In the Bedroom</i>	95	61
2004	<i>Sideways</i>	132	67
2007	<i>Michael Clayton</i>	53	31
2007	<i>Juno</i>	87	48
2010	<i>The Kids Are All Right</i>	130	91
2012	<i>Silver Linings Playbook</i>	30	37
2016	<i>Manchester by the Sea</i>	105	60
2019	<i>Joker</i>	18	19

As can be observed in Table 5.1, the Italian versions usually present fewer tokens of hesitations compared to the corresponding English text. *Forrest Gump* (1994) remains the movie with the lowest number of occurrences, while *Heaven Can Wait* (1978) is the only title that approximates the maximum registered in the English component of the corpus (130 occurrences) with 127 tokens. These results seem to be in line with the studies reviewed in Chapter Two, according to which hesitations are customarily reduced in dubbing by either omitting or substituting them with other elements such as repetitions, discourse markers, filler phrases and indistinct noises (Romero-Fresco 2007; Valdeón 2008; Bruti 2019).

More specifically, in the dubbed section of the corpus, five titles contain less than half the number of hesitations found in their English soundtracks (*The Conversation* 26%; *Michael Clayton* 36%; *Silence of the Lambs* 39%; *Kramer vs. Kramer* 36% and *Twelve Angry Men* 47%), while in nine films the Italian version features between 51% and 64% of the hesitations found in the source text¹⁴⁵. In various cases, the reduction is not particularly severe, meaning that the dubbed version includes between 70% and 85% of the occurrences found in the English text¹⁴⁶. In four titles, the number of hesitations contained in the Italian soundtrack is almost equal to the number recorded in the original version, i.e. the Italian

¹⁴⁵ *Sideways* (2004) presents 51% of the original hesitation markers, *The Verdict* (1982) 53%, *Annie Hall* (1977) 55%, *The Hustler* (1961) 56%, both *Father of the Bride* (1950) and *Manchester by the Sea* (2016) 57%, *Mister Roberts* (1955) 58%, *Mississippi Burning* 59%, and *In the Bedroom* (2001) 64%.

¹⁴⁶ The number of hesitation markers found in Italian dubbing compared to the English version approximates to 68% in *American Beauty* (1999), 70% in *The Kids Are All Right* (2010), 71% in *The Graduate* (1967), 82% in *Guess Who's Coming to Dinner*, 83% in *Jerry Maguire* (1996) and 85% in *Separate Tables* (1958).

adaptation of *Heaven Can Wait* (97%), *Tootsie* (92%), *Forrest Gump* (91%) and *Juno* (91%). At the same time, there are cases where the dubbed movie features more occurrences than those registered in the English soundtrack: with respect to *Joker* (2019), the two texts are balanced with only one more token in the dub compared to the English version, while the augmentation amounts to 23% in *Silver Linings Playbook* (2012) and to 30% in *Driving Miss Daisy* (1989)¹⁴⁷. The most important increase, however, is registered in the film *The Apartment* (1960), where the percentage of increment reaches 67%.

Table 5.2. Raw Frequencies and Percentages of Hesitations per Decade in Italian Dubbing						
1950s	1960s	1970s	1980s	1990s	2000s	2010s
160 (57%)	218 (99%)	260 (59%)	195 (73%)	95 (65%)	207 (56%)	207 (73%)

Looking at raw frequencies in the seven sub-sections of the corpus (Table 5.2), it is worth noticing that the number of hesitations included in Italian dubbed film speech is never inferior to half the total (56%) registered in the corresponding Anglophone dialogue. Nevertheless, as previously noted in Chapter Four, raw frequencies must be normalized per 1,000 words (Table 5.3)¹⁴⁸.

Table 5.3. Normalized Frequencies of Hesitations per Decade in Italian Dubbing (1,000 words)			
Decade	Total Number of Words	Raw Frequencies	Normalized Frequencies
1950s	47,890	160	3.341
1960s	43,404	218	5.023
1970s	46,909	260	5.543
1980s	46,010	195	4.238
1990s	44,709	95	2.125
2000s	42,035	207	4.924
2010s	46,435	207	4.458

The analysis of the Italian dubs exhibits some similarities with the data previously reviewed with respect to the source English text: in Italian dubbed film speech as well, the use of hesitation markers tends to fluctuate over time. Nevertheless, these fluctuations do not appear to be statistically significant¹⁴⁹. In other words, although weak increases and

¹⁴⁷ The percentage of increase has been obtained using the following formula: $\frac{\text{current value} - \text{previous value}}{\text{previous value}} \times 100$, where the current value corresponds to the number of tokens found in Italian dubbing, while the previous value corresponds to the number of tokens registered in the original English version of the film.

¹⁴⁸ As already noted in Chapter Four, for the sake of readability, the data here reported provides frequencies normalized to three digits after the decimal point. The exact occurrences normalized per 1,000 words can be found in the Appendices.

¹⁴⁹ Statistical tests have been carried out through RStudio software, using the Conditional Inference Tree (CTree) method. By comparing the dependent variable (normalized frequencies of hesitations) with the

decreases can be observed across decades, these cannot be attributable to a systematic temporal trend.

As can be seen in Figure 5.1, the frequency of occurrence of hesitations in dubbing seems to be partially influenced by the English source text. Some differences, however, can be noticed: while in the English component of the corpus a slight downfall occurs in the 1960s, in the Italian component the employment of hesitations continues to rise until the 1970s. As a matter of fact, although in the 1950s the occurrence of the phenomenon is relatively low — possibly as an aftermath of the linguistic standardization and purism of the Fascist era, which favoured a more neutral and literary style of dialogue writing, avoiding overly colloquial language and unnecessary Anglicisms (Rossi 1999: 54; Mereu Keating 2016)¹⁵⁰ — an increase of approximately 50% emerges in the following decade.

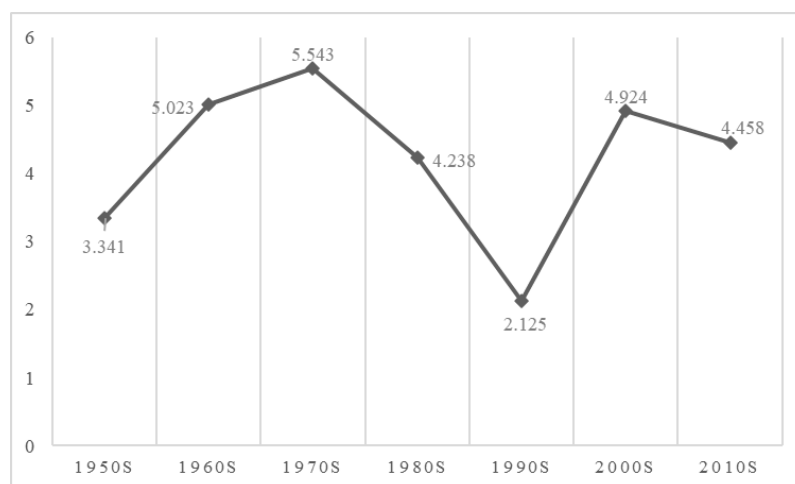


Figure 5.1. Normalized Frequencies of Hesitations in Italian Dubbing over the Years (1,000 words)

From the 1970s onwards, the trend of variation recorded in Italian dubbing seems to comply with the trend witnessed in Figure 4.1 (see §4.1.1): after the positive peak found in the 1970s, in both graphs (i.e. Figure 4.1 and Figure 5.1) there is a progressive decline that reaches its lowest point in the 1990s; then, frequencies augment in the 2000s and slightly decrease again in the 2010s. This seems to suggest that, also in Italian dubbing, the 1970s is the decade in which dialogues appear to be most ‘hesitant’, in the sense that the features of orality

predictor variable (decade), the test yielded a single node, i.e. no significant splits (the graph can be found in the Appendices). Therefore, no statistically significant data emerge in relation to how hesitations fluctuate over the years.

¹⁵⁰ As pointed out by studies conducted on dubbing in fascist Italy, “standard Italian in films could be exploited as a tool to educate the public in the standard pronunciation of the national language and extend the potential and effective, or passive and active, use of it” (Mereu Keating 2016: 73). This resulted in a preference for a formal style of dialogue writing, along the lines of literary language, which impacted both Italian productions and dubbed movies until the end of the 1950s (Rossi 1999: 54).

discussed here tend to be utilized more frequently than in other periods. As pointed out by Paolinelli & Di Fortunato (2005: 18), at that time, the language of both Italian cinema and dubbing became less formal, reintroducing elements that were previously avoided such as, for instance, dialectal variation. During this period, Italian dubbing moved away from the literary style that had defined it up to that point, favouring more natural dialogues and less artificial performances by dubbing actors (*ibidem*: 19).

After the highest peak, there is a progressive decrease, reaching the lowest value in the 1990s. Even in this case, the decline may be attributable to the limited number of hesitations already found in the source English text, which leads to a more linear and less spontaneous style in dubbing. The reduction registered in the 1980s, however, is not as severe as the one recorded in the English language (a 24% decrease in Italian dubbing compared to the 38% decrease registered in Anglophone film speech during the same decade)¹⁵¹; the decrease is far more dramatic in the following decade, where the percentage of reduction reaches 50% (falling by 42% in original English dialogue). It could also be hypothesized that, in the 1990s, original scripts tended to be heavily domesticated, resulting in a polished, transparent, fluent language style according to which hesitations were more likely to be disregarded.

New attention seems to be paid to hesitation markers in more recent dubbings: a new stability is found in the new millennium where, despite the slight decline registered in the 2010s, the frequency of occurrence maintains medium-high values. Thus, hesitations seem to be reappraised in dubbed film speech, possibly as a consequence of the rise of DVDs and, later, streaming platforms, which favoured the circulation and consumption of cinematic products easily available in their original version (see, among others, Matamala 2018; Chiaro 2019; Valdeón 2022)¹⁵² and, in turn, a greater sensitivity towards maintaining and reproducing the orality of the source text.

As pointed out in the previous chapter, the variations registered across decades in terms of employment of hesitation markers in Italian dubbed film speech may result from the influence of various stylistic, narrative and contextual factors. These aspects can impact

¹⁵¹ The percentage of reduction between decades has been obtained using the formula: $\frac{\text{previous value} - \text{current value}}{\text{previous value}} \times 100$, where current value refers to the more frequent decade taken into account as against a previous decade. The percentage value of reduction has been obtained by using the normalized frequencies (1,000 words).

¹⁵² As remarked by various scholars, technologies have moved towards a greater empowerment of audiences; most notably, DVDs, Blu-ray technology and streaming platforms have gradually allowed viewers to decide which audiovisual translation mode is more in line with their personal needs, regardless of the national dubbing/subtitling tradition (Matamala 2018: 71; Chiaro 2019: 171; Valdeón 2022: 372).

the strategies and conventions shared within the dubbing industry with regard to the translation of orality features from English into Italian. In this respect, the influence film genre may have on the occurrence of hesitations in Italian dubbing has been examined.

5.1.2. Dubbing Hesitations and the Genre Variable: A Diachronic Analysis

In this section, the genre variable is reviewed in order to ascertain whether film genre impacts on the attention paid to hesitations in Italian dubbed adaptations. In order to accomplish this, the normalized frequency (per 1,000 words) of occurrence of hesitations in each dubbed title has been computed (Table 5.4). Subsequently, by grouping together the normalized frequencies of hesitations in both comedy and non-comedy films, the mean score of the two genre categories per each decennium has been calculated.

Year	Film	Genre	Number of Words	Hesitation Tokens	Normalized Frequencies
1950	<i>Father of the Bride</i>	Comedy	11,668	57	4.885
1955	<i>Mister Roberts</i>	Comedy	12,536	30	2.393
1957	<i>Twelve Angry Men</i>	Non-comedy	12,822	44	3.432
1958	<i>Separate Tables</i>	Non-comedy	10,864	29	2.669
1960	<i>The Apartment</i>	Comedy	13,124	100	7.62
1961	<i>The Hustler</i>	Non-comedy	8,545	23	2.692
1967	<i>Guess Who's Coming to Dinner</i>	Comedy	13,726	73	5.318
1967	<i>The Graduate</i>	Non-comedy	8,009	22	2.747
1974	<i>The Conversation</i>	Non-comedy	8,534	18	2.109
1977	<i>Annie Hall</i>	Comedy	14,883	73	4.905
1978	<i>Heaven Can Wait</i>	Comedy	11,983	127	10.598
1979	<i>Kramer vs. Kramer</i>	Non-comedy	11,509	42	3.649
1982	<i>Tootsie</i>	Comedy	15,495	84	5.421
1982	<i>The Verdict</i>	Non-comedy	11,102	69	5.314
1988	<i>Mississippi Burning</i>	Non-comedy	9,862	16	1.622
1989	<i>Driving Miss Daisy</i>	Comedy	9,551	26	2.722
1991	<i>The Silence of the Lambs</i>	Non-comedy	8,961	22	2.455
1994	<i>Forrest Gump</i>	Comedy	11,434	10	0.875
1996	<i>Jerry Maguire</i>	Comedy	15,610	50	3.203
1999	<i>American Beauty</i>	Non-comedy	8,704	13	1.494
2001	<i>In the Bedroom</i>	Non-comedy	9,169	61	6.653
2004	<i>Sideways</i>	Comedy	11,484	67	5.834
2007	<i>Michael Clayton</i>	Non-comedy	12,032	31	2.576
2007	<i>Juno</i>	Comedy	9,350	48	5.134
2010	<i>The Kids Are All Right</i>	Comedy	10,344	91	8.797
2012	<i>Silver Linings Playbook</i>	Comedy	15,880	37	2.33
2016	<i>Manchester by the Sea</i>	Non-comedy	12,325	60	4.868
2019	<i>Joker</i>	Non-comedy	7,886	19	2.409

The mean score obtained from the analysis of the standardized frequencies based on genre shows that, as highlighted for English dialogues, hesitations tend to be more present in dubbed comedies than they are in dubbed non-comedies (Figure 5.2). This is true also for those decades that presented the opposite trend in the English component of the corpus, i.e. the 1980s and the 1990s. Unlike Anglophone film speech, in Italian dubbing the phenomenon remains consistently more prevalent in comedies. Interestingly, the difference between the two genre categories is almost non-existent in the 1990s, where the mean score obtained for comedies is only 3% higher than the mean score obtained for non-comedy films¹⁵³. More specifically, the disparity in favour of comedy films is rather scarce in the 1950s (19%), in the 1980s (17%) and in the 2000s (19%), while it reaches a higher percentage in the most recent dubbings, where the mean score recorded for comedies is 53% higher than the value recorded for non-comedies. Remarkably, there are two decades in which the genre variable seems to play a more significant role: in the 1960s and in the 1970s, the mean score witnessed for comedies is respectively 144% and 169% higher than the mean score registered for non-comedy films.

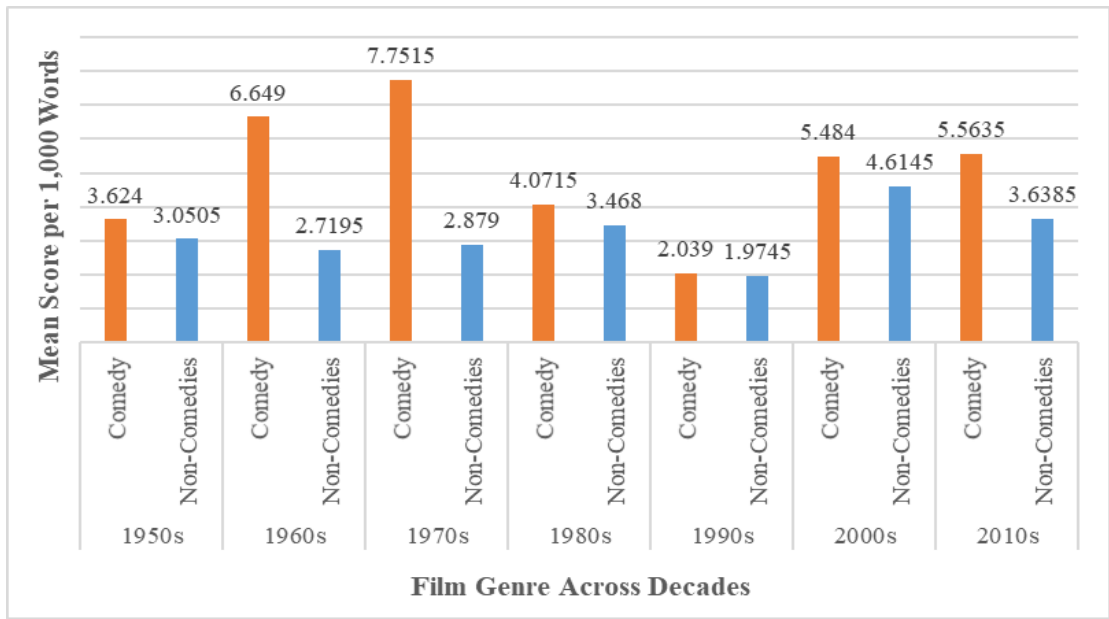


Figure 5.2. Hesitations in Dubbed Comedies and Non-Comedies over the Years

These high peaks could be attributed to the specificities of two comedies, which exhibit some of the highest values registered in the Italian dubbed corpus: *The Apartment* (1960), with a normalized frequency of 7.62 (simplified from 7.61962816), and *Heaven Can Wait* (1978),

¹⁵³ To find out how much higher the mean score of one genre category is compared to the other, the following formula has been used: $\frac{\text{higher mean score} - \text{lower mean score}}{\text{lower mean score}} \times 100$.

with a normalized frequency of 10.598 (simplified from 10.5983477)¹⁵⁴. It is possible to hypothesize that, in the Italian adaptations of these two titles, the protagonists' characterization and the situations in which characters find themselves encouraged the use of pragmatic markers such as hesitations.

More specifically, in *The Apartment* (1960), the protagonist C.C. Baxter (Jack Lemmon), an employee at a large insurance company, advances his career by lending his apartment to his superiors, who are involved in extramarital affairs; among them there is also the president of the insurance company (Mr. Sheldrake), who has a relationship with the woman of Baxter's dreams, Fran Kubelik. Baxter is quite a lonely character, goofy and funny, but also weak; he is "ambitious [...], but in an oddly passive way" (Scott 2021: online)¹⁵⁵ and, because of his secret, he often finds himself in equivocal situations, which contribute to the comical aspect of the film. The goofiness of the character and the comical misunderstandings are emphasized in the Italian dubbed version, where hesitations are 67% more present than they are in the original version¹⁵⁶. On the other hand, the film *Heaven Can Wait* (1978) narrates the story of Joe Pendleton (Warren Beatty), a backup quarterback for the football team Los Angeles Rams, who reincarnates in the body of a billionaire, Leo Farnsworth, killed by his wife and his accountant. In this movie, the protagonist finds himself having to pretend to be Leo Farnsworth, thus living his day-to-day life. Thereby, Joe is repeatedly facing unfamiliar situations in which he has to dissimulate, acting as if he knew what to do. Even in this case, the combination of misunderstandings and Joe's ignorance and foreignness to Leo's world constitutes the comical essence of the film.

Based on these premises, the positive peaks found in the 1960s and in the 1970s may be a reflection of the fact that, in the adaptations of the abovementioned titles, hesitations are employed with the aim to exaggerate the meanings stereotypically associated with these pragmatic features such as uncertainty, clumsiness, vagueness and lies, dubiousness and confusion, so as to characterize characters and the situations they find themselves in.

It could also be hypothesized that, during those years, Italian dubbing may have been influenced by the cinematic products distributed in Italy between the end of the 1950s and the end of the 1970s. These are the years of the so-called "commedia all'italiana" (Italian-

¹⁵⁴ Normalized frequencies reach medium-high levels also in the other two comedy films selected for those decades; most notably, 5.318 (simplified from 5.31837389) in *Guess Who's Coming to Dinner* (1967) and 4.905 (simplified from 4.90492508) in *Annie Hall* (1977).

¹⁵⁵ Retrieved online at the following link: <https://scottsm589.medium.com/the-apartment-a-masterpiece-wilderwise-102fb32f2b61>.

¹⁵⁶ Further information regarding the translation strategies used in the Italian dubbed adaptation of *The Apartment* (1960) is provided in section §5.2.

style comedy), i.e. comedies characterized by a type of satire that focuses on customs and, albeit not always explicitly, politics. These comedies present a realistic approach and report ordinary events with the aim of exposing the bane of contemporary society: satirical stories that could also be told in the form of tragedy (D’Amico 2008: 127)¹⁵⁷. The realistic approach adopted in these films also applies to language, which pandered to what was emerging from contemporary linguistic usage (De Mauro 1986: 162), reproducing spontaneous speech through a heterogeneous linguistic repertoire, combining codes, registers and dialectal inflections (Messina 2017: 103). In this regard, language was used for characterization, emphasizing clichés and stereotypes, which enhance the ironic and humorous effect of the film (*ibidem*).

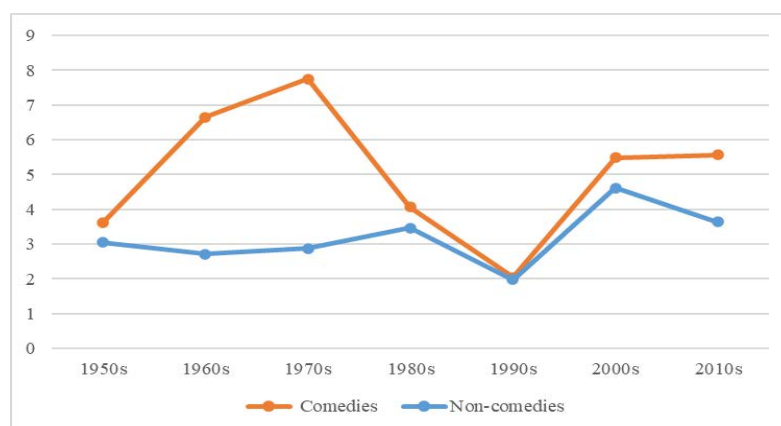


Figure 5.3. Variation of Hesitations in Dubbed Comedies and Non-Comedies (1950s-2010s)

As can be observed in Figure 5.3, in Italian dubbing, film genre seems to have a greater impact on the occurrence of hesitations compared to Anglophone film speech (see Figure 4.3 in §4.1.2). This observation is further supported by the statistical test ANOVA (analysis of variance). While the CTree analysis does not highlight any statistically significant difference (see the graph in the Appendices)¹⁵⁸, the ANOVA test yields an interesting finding: the genre variable presents a p-value of 0.0919. Although this result does not reach the conventional threshold for statistical significance ($p < 0.05$), the proximity of this value to the significance level seems to suggest that there could be a potential trend. In other words,

¹⁵⁷ As reported by D’Amico (2008: 152), the 1960s were particularly fruitful years for comedy in terms of national production; suffice it to say that between 1961 and 1971, 71 titles classified as “commedia all’italiana” were produced.

¹⁵⁸ The graph in the Appendices shows the single node obtained from the examination of the potential interactions between the normalized frequencies of hesitations (i.e. dependent variable) and the two predictor variables, decade and genre. Since the CTree results in a single node, none of the predictor variables meet the statistical criteria to explain the fluctuations registered over the years.

as opposed to Anglophone film speech (where the p-value for the genre variable is 0.640), there might be a correlation between film genre and the frequency of occurrence of hesitations in dubbing¹⁵⁹. From the graph above, this appears to be particularly true in older dubbings, where the values registered in comedies are visibly higher than the values registered in non-comedy films. This could be interpreted as a sign of the fact that, in older comedies, hesitations tend to be used as a dominant stylistic element for purposes related to the characterization of the protagonist; hence, a more stereotyped use of hesitation markers is found in Italian dubbing, where these features are employed as an expressive tool to boost the comical dimension of the film.

Looking at the variation of the two genre categories over the years, non-comedies present a more stable trend, especially in older dubbings, while comedies exhibit more extreme variations (see both Figure 5.3 and Table 5.5). As a matter of fact, while a strong increase is found in older comedies (+83% from the 1950s to the 1960s, followed by another +17% in the 1970s)¹⁶⁰, the situation remains rather stable in the non-comedy category with a slight decrease of 11%, followed by a modest increase (+6%), which becomes more significant in the ensuing decade (+20% in the 1980s). In non-comedy films, the most severe variations occur between the 1980s and the 2000s, with a reduction of 43% (1980s > 1990s), followed by a sharp increase of 134% (1990s > 2000s). On the other hand, in comedy films, the decrease toward the record low begins immediately after the positive peak of the 1970s with a decline of 47% (1970s > 1980s), followed by another 50% (1980s > 1990s).

Table 5.5. Percentage Variation of Hesitation in Dubbed Comedies and Non-Comedies		
Decades	Percentage Variation in Comedies	Percentage Variation in Non-Comedies
1950 > 1960	+83%	-11%
1960 > 1970	+17%	+6%
1970 > 1980	-47%	+20%
1980 > 1990	-50%	-43%
1990 > 2000	+169%	+134%
2000 > 2010	+1%	-21%

¹⁵⁹ ANOVA has been calculated using the software RStudio. In so doing, the p-value for the two predictor variables (i.e. decade and genre) has been computed. The p-value corresponds to 0.4924 for the decade variable and to 0.0919 for the genre variable, while the ratio between the two predictor variables is 0.8763. Further details on the data obtained from the ANOVA test are provided in the Appendices.

¹⁶⁰ The percentage variation is obtained by using the following equation: $\frac{\text{more recent mean score} - \text{previous mean score}}{\text{previous mean score}} \times 100$.

The situation between the two genre categories appears more balanced during the decades spanning 1980 to 2000 (Figure 5.3). The similar trends displayed by both comedies and non-comedies over these twenty years may be a reflection of the fact that, at that time, Italian dubbing was characterized by a more consistent style, which realigns with the source text regardless of the genre of the film being dubbed (see the similar trend in Figure 4.3 in §4.1.2). In both genre categories, the phenomenon decreases between the 1980s and the 1990s (with comedies exhibiting a more marked decline) and increases in the new millennium. Most notably, comedies remain above non-comedies; nevertheless, the difference between the two is not as marked as it was in the past. In the new millennium, a new spike occurs for both genre categories, with an increase of 169% in comedies and 134% in non-comedies. Interestingly, while new stability is found in comedy films, non-comedies present a slight decline; hence, in more recent dubbings, the gap based on genre seems to be reaffirmed.

All things considered, non-comedies appear more stable over the years; in comedies, the phenomenon was more prevalent in the past, with more extreme variations as opposed to more recent dubbings, where a new balance is achieved. The differences between the two genre categories may be interpreted as a sign of the fact that, in dubbed non-comedy films, the phenomenon is mainly used to convey a sense of naturalness and spontaneity to dialogues, while in dubbed comedy films, hesitations are used as a more dominant stylistic tool for expressive, comical and characterization purposes.

Having examined the incidence of occurrence of hesitation markers in Italian dubbed film speech over the years, it has been crucial to further extend the diachronic investigation by focusing on the specific translation strategies used in Italian dubbing to treat these pragmatic markers.

5.2. Dubbing Hesitations into Italian: A Diachronic Analysis of Translation Strategies

The analysis addressed the treatment of hesitations in dubbing by focusing on the translation strategies identified in the Italian section of the corpus, i.e. Close Rendering, Omission, Substitution, and Addition (see §3.7). By comparing the original English text with the Italian dubbed soundtrack, these are the strategies consistently found in the translated component of the corpus. Starting from raw data, the percentage of occurrence of each translation

strategy has been calculated out of the total number of occurrences found in the source text for each decade¹⁶¹.

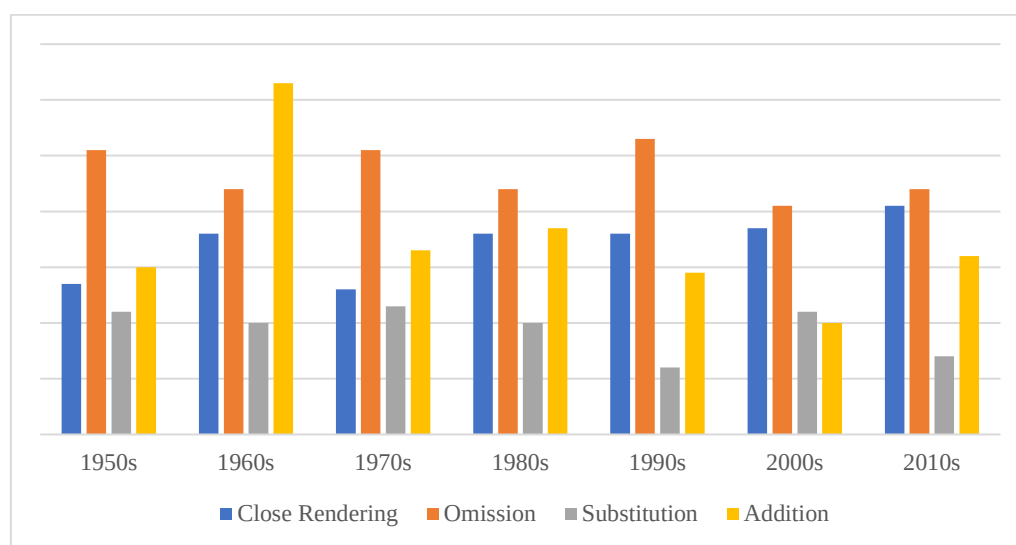


Figure 5.4. Translation Strategies in Italian Dubbing

The results obtained from the analysis are in line with previous investigations, according to which, in dubbing, hesitations tend to be neglected (Romero-Fresco 2007: 199; Valdeón 2008: 127, 133). Indeed, omission is the most frequent strategy in six out of the seven decades, the only exception being the 1960s, where Addition appears as the most used strategy. This result, however, is influenced by the Italian adaptation of the film *The Apartment* (1960), where hesitation markers outnumber the occurrences registered in the English soundtrack (100 tokens in dubbed Italian versus 60 tokens in English). More specifically, out of the 60 tokens included in the original text, 26 are rendered in Italian dubbing, 22 are omitted, and 12 are substituted by other phenomena; at the same time, 74 hesitations are added. As pointed out in the previous section, the substantial increase found in this title could be ascribed to the use of hesitations as a characterization tool: the protagonist finds himself in unanticipated situations, which cause him to make frequent use of hesitation markers.

With regard to Addition, it is interesting to notice that, until the 1970s, this appears as the second most used strategy in dubbing; however, the situation changes in the 1990s. As a matter of fact, while in the 1980s Addition and Close Rendering are almost balanced — the former being slightly more used than the latter — in the 1990s Close Rendering

¹⁶¹ The percentage of use of a specific translation strategy in each decade was calculated using the following formula: $\frac{\text{raw occurrences of a translation strategy in the TT}}{\text{raw frequency of hesitations in the ST}} \times 100$.

emerges as the preferred strategy next to Omission. This tendency continues in the following decades. As a result, it could be claimed that, in earlier movies, the Italian dubbing industry appeared more focused on adapting orality features to the system of the target language by adding hesitation markers in locations where no hesitation is present in the source text. In more recent dubbings, instead, greater attention seems to be paid to the orality of the original text. In other words, Italian dubbing seems to have become more receptive to the way orality features are used in the source English soundtrack. In more recent dubbings, the hesitation markers present in the Anglophone text are maintained or slightly adapted to the target system — for instance, slightly changing their location within the text, presumably to make the character's line flow in a way that is more in accordance with Italian prefabricated orality and with synchronization requirements. In addition, in more recent films, the gap between the rate of employment of the first and the second most used translation strategies is not as wide as it was in earlier films: in the 2000s, Close Rendering includes 135 tokens (37%) and Omission 152 (41%), while in the 2010s, 117 (41%) and 125 (44%), respectively.

To sum up, Omission is the most frequently used strategy, while Substitution is the most underused. As for the other two strategies taken into account, a shift occurs over the years. While in earlier movies the addition of hesitation markers not present in the original soundtrack seems to be preferred, more recent adaptations tend to opt for the same, or slightly adapted, rendering of the original text. As a consequence, it could be claimed that, in more recent dubbings, the original English dialogue is most likely to impact on the use of these pragmatic markers, leading to a greater sensitivity towards maintaining and reproducing the orality of the source text. The four translation strategies presented here will be further discussed henceforward.

5.3. Dubbing Hesitations into Italian: Examples from the Corpus

In the following sections, a more detailed discussion of the way hesitations are treated in Italian dubbing is presented. Through the use of examples extracted from the corpus, the phenomenon of hesitations will be examined with respect to the four translation strategies mentioned above, namely Close Rendering, Omission, Addition, and Substitution.

5.3.1. Dubbing Hesitations: Close Rendering

Close Rendering includes those cases in which hesitation markers are rendered in the Italian soundtrack through vocalizations, that is, vowel sounds or schwa sounds, which can either

occur on their own or as cliticized to a previous item. As can be seen from the graph below (Figure 5.5), in older dubbings this strategy is not particularly utilized, with the exception of the 1960s, where Close Rendering is used as much as it is employed in more recent dubbed films (the 1980s and the 1990s). Indeed, since the 1980s, this strategy exhibits a more stable percentage of use, with a growing tendency over the last ten years.

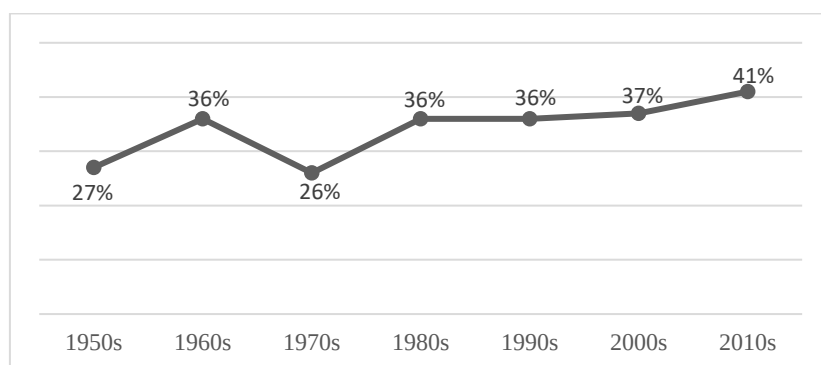


Figure 5.5. The Use of the Close Rendering Strategy over Time

In particular, Close Rendering includes those instances in which a hesitation marker present in the source text is maintained and reproduced in the target version. This can be observed in the examples below: in extract 1 from *Joker* (2019), the original hesitation — a silence filling pause while looking for a joke on a notebook — is reproduced with the addition of the prolongation of the nasal sound at the end; in extract 2 from *The Apartment* (1960), the original hesitation which follows the first-person subject pronoun *I* is rendered in Italian within the repetition of the self-referential item *mi* (the first-person singular object pronoun), maintaining the same function of the original element, i.e. to plan the lying answer to a question (Baxter did not drop his keys, his keys were left under the doormat by one of Baxter's colleagues, who used the apartment to meet his mistress).

Extract 1 - *Joker* (2019)

	English Dialogue	Italian Dubbing
Arthur	((mimicking his mum's voice)) you should enjoy it one day you'll have to work for a living . ((back to his voice)) no I won't ma I'm gonna be a comedian ((laughter)) ... and um .. here's one .. you know, I was just thinking the other day, why are the rich people so:--	((mimicking his mum's voice)) dovresti fartela piacere, un giorno dovrai LAVORARE per vivere . ((back to his voice)) non servirà mamma, no perché io farò il comico ((laughter)) ... ehm: .. ecco questa .. sapete, giusto l'altro giorno stavo pensando, perché le persone ricche--

Extract 2 - *The Apartment* (1960)

	English Dialogue	Italian Dubbing
Mrs. Dreyfuss	something the matter?	le è successo qualcosa eh?
Baxter	oh no I uh seem to have dropped my keys	no no mi eh mi deve essere caduta la chiave

As can be inferred from the following extracts, when maintaining the hesitation of the source text, the original function is oftentimes also reproduced in the target version. In extract 3, for instance, *uh* is preserved in answering a question with specific information; in extract 4, the vocalization is maintained as a holding device used to organize different speech acts within the turn, i.e. moving from congratulating someone to launching a leave-taking; in extract 5, instead, the hesitation uttered in dubbing reproduces the original planning moment in which the character reflects on how to phrase the subsequent question.

Extract 3 - *Forrest Gump* (1994)

	English Dialogue	Italian Dubbing
Nixon	well where are you staying?	bene e dove alloggia?
Forrest	uh it's called the hotel Ebbott	uh si chiama l'albergo Ebbott

Extract 4 - *American Beauty* (1999)

	English Dialogue	Italian Dubbing
Carolyn to Jane	= honey . oh! I'm so proud of you you know I watched you very closely, you didn't screw up once! ((chuckles)) okay, u:h we have to go	= tesoro, oh! come sono fiera di te sai ti ho guardata con attenzione, non hai sgarrato un passo! ((chuckles)) bene, ehm dobbiamo andare

Extract 5 - *The Verdict* (1982)

	English Dialogue	Italian Dubbing
Sally	what does it mean? I mean you uh . have other tactics?	ma che succede? voglio dire eh . ha qualche altra risorsa?

Close Rendering seems to be especially used in contexts where hesitations fulfil a function related to the Hesitation Proper Domain, as can be observed in extract 6 from *Heaven Can Wait* (1978). In this scene, Joe is talking with Mr. Jordan, some kind of guardian angel who is only visible to the protagonist. When the maid finds Joe talking on his own, his turn reflects the character's awkwardness in trying to come up with something to say that could explain his weird behaviour. In the Italian adaptation, the three markers of hesitation proper are reproduced, echoing the character's embarrassment and dithering attempt at explaining the situation.

Extract 6 - *Heaven Can Wait* (1978)

	English Dialogue	Italian Dubbing
Joe to maid	= ((chuckles)) uh I-I-- the uh -- ... uh hi I-I'm-I'm just--	= ((chuckles)) uh io l-la-- eh -- eh uh no e-- io-- vado, ciao, ciao

A similar case is also reported in the following extract from *Tootsie* (1982). In this scene, Michael, in the guise of Dorothy, is spending the weekend with the colleague Julie and her father, Les. When Julie leaves the room to go to bed, Dorothy hesitates about joining her; she does not want to get unready in front of Julie, risking her finding out that Dorothy is actually a man. This sense of panic is conveyed in the original text by three hesitation markers that co-occur with false starts and reformulations. The moment of difficulty experienced by the character is also rendered in Italian through three vocalizations, which emulate the hesitation proper function of the original ones.

Extract 7 - *Tootsie* (1982)

	English Dialogue	Italian Dubbing
Julie	well I'm going to bed, what do you say Dorothy? do you wanna hit the hay as they say on the farm?	io vado a letto, tu che fai Dorothy? anche tu al fienile? come dicono qua?
Dorothy	what? oh well I think I'll-- you uh -- oh uh both of-- ((coughs)) you two uh both of you go to bed I'll stay up	cosa? ah beh io-io credo che eh tu eh tutti e due ((clears throat)) eh voi andate pure a letto io resto al piano

On the other hand, there are instances where the hesitation reproduced in the target soundtrack performs a slightly altered function compared to the one fulfilled in the source text. This occurs, for instance, in the following scene from *Sideways* (2004). In the English soundtrack, the hesitation marker represents a planning moment in which the speaker stops and redrafts the sentence, which is actually a lie (i.e. Planning Domain). In its Italian translation, the hesitation marker corresponds to a turn-holding device (i.e. Structural Domain) rather than a planning device, that is, a tool used to maintain the turn and add the explanation of the reason why the two friends, Miles and Jack, are spending the weekend in Santa Ynez Valley, California. They are not there to celebrate Miles' book being published; they are there for Jack's bachelor party.

Extract 8 - *Sideways* (2004)

	English Dialogue	Italian Dubbing
Jack	it's getting published that's u:h -- it's actually why we're up here celebrating	e verrà pubblicato em per questo siamo qui a festeggiare

Close rendering may require the original hesitator to be adapted to the translated text, slightly changing how it takes place within the turn. In the scene reported in extract 9, for instance, the original hesitation marker is rendered in Italian through an indistinct sound, which follows the prolongation of the vowel at the end of the verb *fosse* in “come se fosse zero” (literally, “as if she were zero”) and conveys the momentary planning in looking for the right term.

Extract 9 - Tootsie (1982)

	English Dialogue	Italian Dubbing
Michael	((offscreen)) she's really a very very attractive girl ((onscreen)) and she's no dummy either but for the life of me, I cannot understand what she's doing hanging around with that director he treats her like she's just uh , nothing!	((offscreen)) è davvero una ragazza molto molto attraente ((onscreen)) e neanche stupida ma io, giuro sulla mia vita, non riesco proprio a capire come faccia a stare con quel regista, la tratta come se fosse: ((indistinct sound)) zero

Moreover, the adaptation process may require dubbing actors to deliver the hesitation in a slightly different location compared to the source text, adapting the phenomenon to the norms of the target language. In extract 10, for instance, the hesitations at the beginning of the English turn have been adapted in the Italian version through a cluster that contains two hesitation markers — one hesitation proper and one hesitation used to mitigate Baxter's refusal to lend his apartment to Mr. Dobisch, preceding the apology (*mi dispiace*) rather than following it (*I'm sorry*) — in combination with the interjection *ah*, which launches the turn (instead of the original hesitation *uh*), and the interjection *beh*, followed by a soft chuckling sound.

Extract 10 - The Apartment (1960)

	English Dialogue	Italian Dubbing
Baxter	uh ((soft chuckle)) I'm sorry uh mister Dobisch I like to help you fellas out but it's sort of late so why don't we uh make it some other time?	ah eh beh ((soft chuckle)) eh mi dispiace signor Dobisch io l'aiuterei volentieri ma è un po' tardi adesso, perché non rimandiamo a un altro momento?

Likewise, in Extract 11, it is possible to observe a hesitation being relocated at the beginning of the turn, inducing a change of pragmatic function in the Italian dub. Most notably, while in the source text the hesitation marker corresponds to a planning element interrupted by another character, in the target text the turn-initial item corresponds to a tool used to launch the answer to Natalie's question.

Extract 11 - In the Bedroom (2001)

	English Dialogue	Italian Dubbing
Natalie	mrs Fowler?	la signora Fowler?
Matt	no <i>uh</i> --	<i>uh</i> no, no è--
Customer	can I get a pack of Marlboro Reds?	un pacchetto delle solite Natalie

Similarly, the adaptation process may also require separating two adjacent hesitation markers. In the following scene (Extract 12), two contiguous vocalizations are uttered in the source text; in Italian dubbing, one hesitation marker is moved at the beginning of the turn as a turn-taking device, while the other hesitation is moved after the vocative *Matt* to plan the continuation of dispreferred content, i.e. announcing something to some extent problematic to the interlocutor.

Extract 12 - Guess Who's Coming to Dinner (1967)

	English Dialogue	Italian Dubbing
Mrs. Drayton	you see <i>uhm uh</i> Matt there-there's sort of a special, problem	<i>eh</i> vedi, Matt <i>uh</i> c'è anche-- c'è anche un altro piccolo, problema
John	see, I've got to fly to New York tonight and on to Switzerland tomorrow night	sì, sì stasera devo andare a New York, e domani parto per la Svizzera

Nevertheless, it must be pointed out that the Italian adaptation may sometimes alter viewers' perception of the character, leading to a slightly different impression compared to the one conveyed in the source text. This commonly occurs when hesitations are employed in conjunction with other phenomena typical of spontaneous language, as can be observed in extract 13. In the target text, the two hesitations uttered in the English soundtrack are reproduced, adapting them to the translated version, where disfluency phenomena and other elements such as *oh ma* (corresponding to "oh but") are added. Here, hesitations co-occur with the repetition of the first-person singular pronoun *io*, the filler sentence *beh eh bene* (corresponding to "well oh okay"), as well as with the reformulation *ma è-- era così* (back translation: "that's-- that was just [...]"). All these features make the dubbed character sound more hesitant, as if he were drowning in verbiage, producing a turn that is more fragmented in Italian compared to the source text.

Extract 13 - The Apartment (1960)

	English Dialogue	Italian Dubbing
Baxter	I uh I understand . well it was uh just an idea I hate to see a ticket go to waste	oh ma <u>io-io</u> eh <u>io--</u> beh mi rendo conto beh eh bene <u>ma è--</u> era <u>così</u> un'idea eh dispiace sciupare un biglietto

Close Rendering also involves cases of explicitation; that is, an original indistinct hesitating sound is made more explicit in dubbing through the employment of vowel or schwa sounds. For example, in the case presented in extract 14a, the indistinct hesitating sound has been rendered through the vocalization *eh*, uttered with the repetition and prolongation of the initial vowel sound to mitigate the bad-news announcement: Pat's father (Pat Senior) is informing his son (Pat) that his ex-wife has sold the house. Moreover, as can be observed in the transcription, the second hesitation marker has been displaced to a different location, moved to the end of the false start ("ecco lei" to translate "you know she") as a tool to reformulate the sentence. Likewise, in extract 14b, the original indistinct sounds are made more explicit through the hesitation marker *eh* followed by a stuttering repetition of the article *la*, used to render the same sense of concern expressed in the original version. In extract 14c, instead, the original indistinct hesitating sound used to express disagreement is rendered in Italian through a schwa sound anticipated at the beginning of the unit *un momento* ("wait a minute") and following the interjection *beh* ("well"), which is also added in dubbing. Finally, in extract 14d, the interjection *ah* is inserted with a final cliticized schwa sound.

Extracts 14 - Indistinct Hesitating Sounds Rendered in Italian Dubbing

	English Dialogue	Italian Dubbing
14a	<i>Silver Linings Playbook</i> (2012) ((indistinct hesitating sound)) uh y-you know she-- Nikki sold the house she left .. didn't your mother tell you she--	e-e:h ecco lei-- uh Nikki ha venduto la casa se n'è andata .. tua madre non te l'ha detto?
14b	<i>The Apartment</i> (1960) ((indistinct sounds)) the vice squad?	eh l-la-la polizia?
14c	<i>Tootsie</i> (1982) wait a minute ((indistinct sound)) her line-her line is, your husband--	beh uh un momento lei-- lei doveva dire suo marito--
14d	<i>Michael Clayton</i> (2007) ((indistinct sound)) yeah b- because I-I-I meant what I said	ahuh sì per-perché io-io parlavo sul serio

As can be inferred from extract 14d, when adapting prefabricated orality to the norms of the target language, vocalizations may be rendered as cliticized to inserts in Italian dubbing.

This can also be observed in the following extract, where schwa sounds are added at the end of the prolonged interjections *ah*, uttered by Jules while planning the answer to Paul's questions:

Extract 15 - *The Kids Are All Right* (2010)

	English Dialogue	Italian Dubbing
Paul	<i>uh</i> . so Nic you-- you're a doctor I know that but what are-what are you into Jules? what do you-what do you do?	beh . e così Nic tu-tu sei medico lo so ma tu di cosa ti occupi Jules che-che-che cosa fai?
Jules	<i>uhm, u:hm</i> well you know I mean I studied architecture in college	a: <i>huh</i> , a: <i>huh</i> beh sai ho cominciato l'università ho studiato architettura

To conclude, it is relevant to point out that Close Rendering also concerns hesitations occurring as free-standing items in the source text. As illustrated by the following extracts, these elements tend to be maintained as a form of backchannel to a bad-news announcement (Extract 16) or, for example, to reproduce characters' intention to answer a question, i.e. a thinking process carried out by speakers to declare that they are reasoning about a potential response (Extract 17).

Extract 16 - *Manchester by the Sea* (2016)

	English Dialogue	Italian Dubbing
Dr. Betheny	total . for most people with the particulars of your condition <i>um</i> the average life expectancy is, five to ten years	in tutto, per la maggior parte delle persone con la sua patologia <i>mh</i> l'aspettativa di vita è dai-- cinque ai dieci anni
Joe	<i>uh</i>	<i>uh</i>

Extract 17 - *The Kids Are All Right* (2010)

	English Dialogue	Italian Dubbing
Paul	Laser, do you have any questions that you wanna-- ask?	((chuckles)) Laser c'è qualche domanda che vorresti-- farmi?
Laser	<i>uh</i>	<i>uhm</i>
Paul	go ahead shoot anything	dai spara forza

As can be inferred from the examples above, even in those instances where the hesitation marker of the ST is maintained and reproduced in the TT, the dubbed version maintains some degree of autonomy from the English text. In other words, while in many cases the vocalization of the source text is faithfully rendered in the Italian dub, in many others the Italian version seems to reproduce the hesitation marker by adapting it to the target system through various techniques, i.e. by adding prolongations next to the vocalization or by

resorting to a slight change of location within the translated text, as well as to explication and cliticization to other words.

5.3.2. Dubbing Hesitations: Omission

Omission is the most frequently used strategy; most notably, even if the percentage of employment tends to fluctuate over time, in most cases it represents the preferred solution to render (or not to render) the hesitation marker of the source text. Nevertheless, as illustrated in the graph below (Figure 5.6), the percentage of use of Omission has decreased over time, reaching a percentage of employment that remains between 41% and 44% in more recent dubbings compared to the values registered in previous decades such as 51% in the 1950s/1960s and 53% in the 1990s.

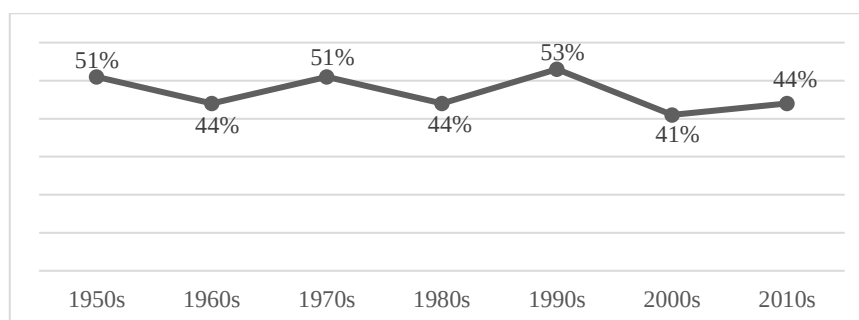


Figure 5.6. The Use of the Omission Strategy over Time

Omission or zero translation implies the complete elimination of the pragmatic marker present in the source text (Extracts 18). As can be inferred from the extract from *Kramer vs. Kramer* (1979), omission may be attributed to the shifts that commonly occur in translation. In this example, the original *whatever she's having* is rendered as *quello cos'è?* (“what is that?” or “what you got there?”) (Extract 18c); more specifically, the type of sentence uttered by the character is changed in the adaptation, shifting from a declarative sentence to a question. This is ascribable to the need to synchronize the line with the speaker’s gesture of pointing at Joanna’s glass while uttering the question (i.e. kinesic synchrony). Similarly, in extract 18e from *Sideways* (2004), a transformation occurs in the type of question asked by Jack: while in the source text the turn is aimed at requesting the interlocutor to put the seat belt on, using the imperative verb (*put on*), in the dubbed version the question is aimed at checking if the interlocutor already has the seat belt on and, if not, implicitly asking him to do so (literally: “did you put on your seat belt, Miles?”). Since characters are not visible on

the screen, this shift is not attributable to synchronization; rather, it occurs as the result of linguistic transposition.

Extracts 18 - Omission of Hesitations in Italian Dubbing

		English Dialogue	Italian Dubbing
18a	<i>Separate Tables</i> (1958)	worried? why? has uh anybody been talking about me?	preoccupata? perché? qualcuno le ha detto qualcosa di me?
18b	<i>The Apartment</i> (1960)	look Baxter, we're making out the monthly efficiency rating and I'm putting you in the top ten now uh you don't want to louse yourself up do you?	senta Baxter, stiamo facendo il rapporto mensile di efficienza e la metterò fra i primi dieci quindi non vorrà rovinarsi la carriera vero?
18c	<i>Kramer vs. Kramer</i> (1979)	u:h whatever she's having	quello cos'è?
18d	<i>Mississippi Burning</i> (1988)	yeah well uh actually it was you I wanted to talk to	ecco beh, veramente era con lei che io volevo parlare
18e	<i>Sideways</i> (2004)	((offscreen)) u:h hey put on your seat belt okay?	((offscreen)) hai messo la cintura Miles?
18f	<i>In the Bedroom</i>	I uh I-I need some things from the bathroom	dovrei-- dovrei prendere delle cose in bagno
18g	<i>The Kids Are All Right</i> (2010)	it just seemed like a lot more fun than uh donating blood ((chuckles)) ... no I um I loved the idea of helping people uh people-- who were in need you know wanted to have kids and couldn't--	mi sembrava più divertente che donare il sangue ((chuckles)) ... no mi piaceva l'idea, di aiutare altre persone, persone, che magari, avrebbero voluto dei figli ma non potevano averli

The omission of a hesitation marker can also correspond to the elimination of a whole turn, as in the case reported from *Heaven Can Wait* (1978) in extract 19. In the Italian version, Joe's turn includes a hesitation marker used to fill a silent pause, providing a backchannel to the information received from the previous speaker (Man 10). In Italian, this backchannel is omitted: Joe appears speechless while other characters say something in the background. Similarly, in the movie *Manchester by the Sea* (2016), Patrick's start to the turn (*uh*), interrupted by the camera shot moving toward the hockey coach, is silenced in Italian; in other words, no hesitation can be heard in the dub. In this case, however, the omission does not have implications for the dubbed version since Patrick is far from the screen — thereby, there are no synchronization requirements — the shot simply moves to another character without reproducing Patrick's start to the answer (Extract 20).

Extract 19 - *Heaven Can Wait* (1978)

	English Dialogue	Italian Dubbing
Man 10	<i>uhm</i> m-mister Farnsworth, that's an ecological suit brought against our canning factory for destroying porpoises	<i>uh</i> signor Farnsworth, quella è una denuncia ecologica fatta contro la nostra industria alimentare per avere ucciso dei marsuini
Man voice	((clears throat))	perfino quello fanno
Joe	<i>uh</i>	
Man 12	oh for God's sake he knows that	oh per la miseria lo sa questo

Extract 20 - *Manchester by the Sea* (2016)

	English Dialogue	Italian Dubbing
CJ	you all right, brother?	tutto okay fratello?
Patrick	<i>um--</i>	
Hockey coach	take these sticks back would you?	raccogliete queste mazze ok? e tornate in fila

As it will be discussed in the next section, the omission of the original hesitation marker is sometimes compensated by the addition of a new hesitation in the target text. At the same time, the elimination of the original item may be compensated by a variety of phenomena, as illustrated by the following extracts from *Twelve Angry Men* (1957) and *Jerry Maguire* (1996). In the former, three hesitations are uttered in the source text. In the Italian dub, one is omitted in favour of a brief silent pause of 1.00800 seconds, signalled by the single dot (.); one is omitted and substituted by an intonative break, signalled by the comma symbol (,); the third one is omitted and compensated for through the repetition of the verbal element *ho* (Extract 21a). In the latter, in lieu of the original hesitation, an intonative break occurs in dubbing with a change of intonation¹⁶², while the medium pause present in the English soundtrack (about 2.403507 seconds, signalled by the two dots) is reduced (to approximately one second, i.e. 0.832000 seconds) in favour of the prepositional phrase *per cui* (literally, “for which”), along with the addition of a false start when uttering the initial sound in *non* (“not”), which is used to render Jerry’s momentary planning process (Extract 21b).

¹⁶² Through the Praat software, it was possible to investigate the Italian dubbed segment “sensazione, per cui [...]” (backtranslation: “that feeling that [...])). The analysis shows a progressive lowering pitch in pronouncing the word *sensazione* (“feeling”), accompanied by a decrease in intensity, which indicates a prosodic closure (the speaker is concluding an intonation unit). After the intonation break (signalled by the comma symbol), the pitch resumes at a higher value, suggesting an intonational re-opening when uttering “per cui”; then, toward the end of the segment, it gradually falls again, outlining a new melodic closure. As for the intonative change, the higher value pitch registered during the onset of the prepositional phrase suggests an emphatic continuation of the speech, i.e. after the intonative break, the speaker resumes the concept of the previous sentence.

Extracts 21 - Omission of Hesitations Compensated by other Phenomena

		English Dialogue	Italian Dubbing
21a	<i>Twelve Angry Men</i> (1957)	[...] there's always some character ready with an idea see ((chuckles)) and it kills me, it's the weirdest thing in the world the way they sometimes precede their idea with a phrase see like-- , some account exec will get up and he'll say uh , ok uh here's an idea let's uh run it up the flagpole and see if anyone salutes it ((laughter)) I mean it's idiotic but it's funny	[...] c'è sempre un tizio pronto con un'idea ((chuckles)) è incredibile ((chuckles)) mi fanno crepar dal ridere come annunciano le loro idee certe volte perché, vede, per esempio si alza un contabile e <u>dice</u> . <u>ragazzi</u> , <u>ho-- ho</u> un'idea <u>facciamola</u> , <u>sventolare</u> e vediamo se qualcuno la saluta ((chuckles)) è idiota ma è buffa
21b	<i>Jerry Maguire</i> (1996)	you know that-- feeling? uh you're not-- .. completely embarrassed yet but you glimpse tomorrow's, embarrassment?	hai presente quella, sensazione, <u>per cui</u> . n-- non sei totalmente imbarazzato ancora ma già, intravedi l'imbarazzo del giorno dopo?

On the other hand, omission may even be compensated by the insertion of a prolonged sound, as in extract 22. In this example, the hesitation within the repetition of the pronoun *I* is not rendered in Italian; nevertheless, a prolongation of the final vowel in *cosetta* is added (“a little something”), together with the stuttering repetition of the *p* sound in *perché* (“because”) and the addition of the filler word *ecco* (corresponding to “well” or “you know” in English).

Extract 22 - Guess Who's Coming to Dinner (1967)

	English Dialogue	Italian Dubbing
John	well ((coughs)) . ((sighs)) I guess I should have called you back again because there is this one thing I-I uh I should have-- I've been meaning to write to you about it ((clears throat)) . yeah there was one thing I didn't explain dad . and . I'm afraid it's gonna be kind of a shock	beh ((chuckles)) . uh forse avrei dovuto richiamarvi <u>p-perché</u> c'è una cosetta : che <u>ecco</u> io-io-- avrei--, ve l'avrei detto per lettera ((soft chuckle)) . si c'è una cosa che non vi ho spiegato uh papà uh . e non vorrei che fosse un colpo per voi

Omission commonly occurs when two adjacent hesitations are present in the source text or when multiple hesitations are uttered near one another. In those cases, it is quite common for one hesitation marker to be maintained in Italian dubbing and for the other element to be omitted, as in the following extract from *The Conversation* (1974). In this scene, the turn-initial hesitation marker is maintained, while the second one is omitted in favour of the discourse marker *senta* (“listen”), followed by the verbal request *mi passi* (“put me

through”), which together cover the space of the original monosyllabic *yes* followed by the vocalization *u:h*.

Extract 23 - *The Conversation* (1974)

	English Dialogue	Italian Dubbing
Harry Caul	<i>uh</i> yes <i>u:h</i> extension seven forty six please	<i>uh</i> <u>senta</u> , <u>mi</u> <u>passi</u> l'interno quarantasei per favore?

Thus, omission could also be attributable to the use of longer sentences in Italian compared to the English soundtrack (see Comastri 2002 and Musacchio 2004 on this), as in the case of *quando avevo cinque anni* (nine syllables), which translates the original *when I was five* (four syllables) in extract 24. Given the longer sentence in Italian, three out of the four hesitations have been eliminated to guarantee isochrony between the two versions, i.e. turns of the same length.

Extract 24 - *The Conversation* (1974)

	English Dialogue	Italian Dubbing
Harry Caul	when I was five <i>u:h uh u:h uh</i> my father introduced me to a friend of his	quando avevo cinque anni <i>uh</i> mio padre mi presentò a un suo amico

Nevertheless, it is important to point out that the omission of hesitation markers in the Italian dubbed version may have consequences on the way lines are delivered, especially when these pragmatic markers accompany other phenomena such as false starts and repetitions. In these instances, the omission of hesitations may result in more linear utterances. This is the case, for instance, in the extracts shown hereunder. In the scene from *Annie Hall* (1977), the omission of the turn-medial hesitation in combination with the repetition of the initial sound when uttering the pronoun *they*, as well as the omission of the false start *they-- like they*, make Alvy's character sound franker in Italian; less doubtful about how to say things and how to continue with his utterance (extract 25). Likewise, in extract 26 from *Tootsie* (1982), Dorothy's line is built in a narrower way in Italian. In this scene, Dorothy, Michael's alter ego, is spending the weekend with Julie and her father. When they get to the house in the mountains, Michael (aka Dorothy) learns that he is supposed to share a room with Julie. Worried about the possibility that Julie may find out his secret (i.e. Dorothy is, in reality, a man), the protagonist utters a stuttering sentence characterized by “retrace-and-repair sequences” (Biber *et al.* 1999: 1062) together with the hesitation *uh*. Both the hesitation and

the reformulations are not rendered in the Italian version, where the sentence is far more linear and less panicky compared to the original one. The character's nervousness is mainly communicated through the actor's facial expressions visible on the screen and through the vocal performance of Dustin Hoffman's dubbing actor, Ferruccio Amendola, who conveys a nervous tone to the line. Therefore, it can be suggested that the omission is offset by visual cues.

Extract 25 - *Annie Hall* (1977)

	English Dialogue	Italian Dubbing
Alvy	I can't believe this family, Annie's mother, she's really beautiful, and, they're talking swap meets and boat basins and-- the old lady at the end of the table ((clears throat)) is a, classic Jew hater, and uh th-th-they really look American you know very healthy <u>they-- like they</u> never get sick or anything, nothing like my family, you know the two are like oil and water	stento a credere a questa famiglia ((sniffs)) la mamma di Annie, è veramente bella, e, parlano di aste e di barche in Darsena e, la vecchia là in fondo alla tavola ((clears throat)) è la-- ((chewing sounds)) classica mangia ebrei, e hanno il classico aspetto americano sapete molto sano, come se non dovessero mai star male, non hanno proprio niente dei miei sono uguali come l'olio e l'aceto

Extract 26 - *Tootsie* (1982)

	English Dialogue	Italian Dubbing
Dorothy	wait what-- are we-- uh what-- are we sharing?	aspetti siamo in camera insieme?

In line with the view of Omission being a domesticating technique (Chaume 2012: 146), this strategy is generally used to simplify the content of a turn and make the dubbed sentence more linear (for example, avoid repeating vocalizations in the case of two adjacent hesitation markers) and to avoid unnecessary elements such as the reiteration of information offset by visual cues. In many instances, Omission is attributable to synchronization, for instance, to guarantee isochrony between the original and the dubbed performance. Along these lines, hesitations omitted and compensated through the insertion of other phenomena also exhibit a domesticating tendency, i.e. favouring linguistic and paralinguistic items that bring the dub closer to the naturalness of Italian viewers. In other words, the use of the Omission strategy highlights a general tendency of Italian versions to be rather independent from the ST with respect to the pragmatic meanings originally conveyed through hesitations.

5.3.3. Dubbing Hesitations: Addition

Addition involves the insertion in the target text of hesitation markers that are not present in the source text; in other words, by comparing the original English dialogue with the Italian dubbed soundtrack, no hesitation marker occurs in the source text in that same location where a hesitation marker is used in the target text. As previously noted, this strategy is particularly employed in the adaptations of specific titles where hesitations collaborate to the comical effect of the on-screen narrative. With the exception of the peak in the 1960s, this translation strategy exhibits a rather balanced cycle, characterized by a gradual growth, followed by a decline and an upturn that rebalances it. Interestingly, the major decline occurs at the time when Addition is replaced by Close Rendering as the second most used strategy to render hesitations in Italian dubbing.

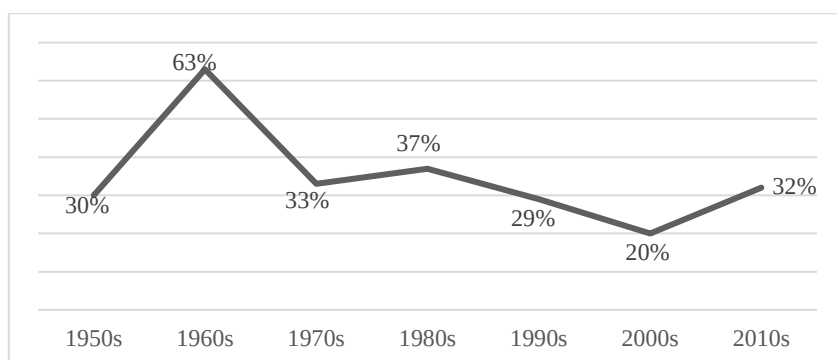


Figure 5.7. *The Use of the Addition Strategy over Time*

As reported in the following extracts, Addition takes place in contexts where no hesitation marker is present in the corresponding Anglophone turn; in these cases, a hesitation may be added at the beginning of the utterance as a turn-taking device used, for instance, to ask a question (Extract 27a) or as a turn-launching device, like in extract 27b, where an ironic hesitator is used to launch the turn. When added in the middle of a turn, hesitations may be inserted as holding devices (i.e. to continue saying something) or as planning tools, used to move forward with the character's telling and/or explanation (Extract 27c). As will be further examined below, this can also be related to synchronization; more specifically, Addition is a strategy used by dubbing actors to fill gaps when on-screen actors' lips are moving (27a). Moreover, it can also be used to add sounds when characters are silent, but their lips are visibly parted (27b and 27c). In the latter case, Addition is possible but not necessary.

Extracts 27 - Additions of Hesitations in Italian Dubbing

		English Dialogue	Italian Dubbing
27a	<i>Twelve Angry Men</i> (1957)	you're two?	uh lei è numero due eh?
27b	<i>The Hustler</i> (1961)	how do you want me to play it? I'm broke	e:h che vuoi che faccia sono scannato
27c	<i>Silver Linings Playbook</i> (2012)	it's insanity he-he spikes the ball at the one yard line, the one fucking yard line! I mean I-- . get into the end zone! dummy I mean--, he celebrates before he's even in	eh beh è una follia eh ferma la palla a una yard dalla meta una sola fottuta yard insomma eh . portala almeno nella end zone cretino! insomma eh-- festeggia ancora prima di aver segnato?

Sometimes, hesitations are inserted to add a sense of tentativeness to a character's turn, as in the following scene from *The Apartment* (1960). In this extract, the hesitation *eh* in conjunction with the repetition of its initial sound is added as a hesitation proper to convey a sense of discomfort toward what Dr. Dreyfuss said to Baxter (Extract 28). Likewise, in extract 29 from *Mister Roberts* (1955), instead of the original moment of silence, a hesitation marker and a nasal sound are added in the Italian soundtrack to verbally express Pulver's reluctance to answer Captain's question.

Extract 28 - *The Apartment* (1960)

	English Dialogue	Italian Dubbing
Dr. Dreyfuss	and sometimes there's a: twi night double header ((tuts)) a nebbish like you!	a volte sembra una doppia punzonatrice alternata ((tuts)) eh! uno piccino come lei!
Baxter	yeah well I'll see you doc	eh già e-eh ci vediamo

Extract 29 - *Mister Roberts* (1950)

	English Dialogue	Italian Dubbing
Captain	how long have you been aboard Pulver?	da quanto tempo è a bordo Pulver?
Pulver	fourteen months sir	uh .. mh ... quattordici mesi

The addition of hesitations may even occur as a compensation strategy when the original hesitation marker is rendered differently in the target text. In these cases, the hesitation appears in a position other than where the source filled pause originally occurred in the English turn; a new location, perhaps more in line with the way these pragmatic markers are employed in (prefabricated) Italian. This can be observed in the following extracts: in 30a,

for instance, a new hesitation is uttered at the very beginning of the turn to launch the answer to the question “Who’s that?”, compensating for the omission of the original planner within the repetition of *just* (*solo* in Italian), which is instead rendered through voice inflection; i.e. in correspondence of *solo--*, the waveform retrieved through Praat shows a slight falling intonational contour combined with decreasing intensity, suggesting a local prosodic closure before continuing the utterance. In extract 30b, a new hesitation is added as a turn-taking device, which compensates for the omission of two elements: a hesitation proper and a thinking filler used to vocalize the character looking for an excuse not to let the other person use the binoculars. Similarly, in extract 30c, a turn-initial hesitation is inserted to launch a response to the interlocutor correcting the speaker about the state of affairs (“We’re not, it’s a protest not a lawsuit”), which compensates for the substitution of the original turn-holding device with the filler expression *insomma* (corresponding to “I mean” in English). In extract 30d, the Italian turn features the reiteration of *ho* followed by *accompagnato* (5 syllables) in lieu of the original hesitation within the repetition of *I drove*, along with additional elements such as *eh sì è mia* (“oh yes it’s mine”); the vocalization is thus recreated at the beginning of the turn as a tool to launch the answer to the question “Is that Ben’s car in front?”. Interestingly, in Italian, Ben’s turn appears more hesitant and fragmented compared to the source text, as can be inferred from the initial *sì*, *eh sì* (backtranslation: “yes oh yes”), as well as from a series of stuttering repetitions (*ho-ho*, *la-la-la*, *l’ho-l’ho*).

Nevertheless, it may also happen that the new element maintains a similar function to the one fulfilled by the original omitted item; this is the case in the extract reported from *Annie Hall* (1977). In this turn, the original hesitation marker is used in the English version to implicitly humble the content of what the invitation involves (negative politeness)¹⁶³, while in the Italian version, the new hesitation marker is used as a tool to mitigate the actual act of inviting Alvy and Annie to join Tony Lacey and his friends, i.e. both functions are related to the Interpersonal Domain.

Extracts 30 - Addition of Hesitations in Different Locations

		English Dialogue	Italian Dubbing
30a	<i>The Kids Are All Right</i> (2010)	that was nobody just <u>uh--</u> just Clay	u:h nessuno solo-- solo Clay
30b	<i>Mister Roberts</i> (1955)	Captain’s order miss you know uh uh bifocals	eh non si può parlare signorina scusi sono bifocali

¹⁶³ As pointed out by Leech (2014:11), “Neg-politeness typically involves indirectness, hedging, and understatement”.

30c	<i>Heaven Can Wait</i> (1978)	yeah but that's not the point the point is that we're doing something wrong uh everybody's suing us, look we got a lawsuit here from a refinery, we got a nuclear lawsuit, we got a guy here a gu-a guy named Porpoise is suing us	eh sì beh questo non è il punto il punto è che facciamo qualcosa di storto <u>insomma</u> ci fanno tutti causa ecco abbiamo una causa qui per una raffineria un'altra per la centrale nucleare c'è un tizio qui ecco, un tizio di nome Marsuini ci fa causa
30d	<i>The Graduate</i> (1967)	yes sir . I drove uh I drove mrs Robinson home, she wanted me to drive her home so I, drove her home	uh sì, eh sì è mia, ho-ho accompagnato la-la-la sua signora a casa, ha voluto che l'accompagnassi e io l'ho- l'ho accompagnata
30e	<i>Annie Hall</i> (1977)	<i>um</i> we-we're going back to the Pierre we're stayng at the Pierre and we're gonna-- meet Jack and Angelica and have a drink there and, if you'd like to come, we'd love to have you and, we can just sit and talk, nothing, uh not a big deal just relax just be very mellow	uh noi ora torniamo in hotel noi siamo al Waldorf e là ci vediamo con Jack e Angelica prendiamo un drink là e se volete venire eh ci fate piacere, facciamo due chiacchiere nient'altro, niente di eccezionale, per distendersi si sta un po' in allegria

Furthermore, various cases of over-extension can be attested in the corpus. In particular, vocalizations seem to be added in dubbing to substitute a range of elements typical of spontaneous language. As illustrated in the following extracts, hesitation markers are commonly used to substitute monosyllabic items such as interjections.

Extract 31 - Hesitations that Substitute Interjections

		English Dialogue	Italian Dubbing
31a	<i>Silence of The Lambs</i> (1991)	oh uh Jack Gordon	uh uh Jack Godon
31b	<i>Forrest Gump</i> (1994)	[...] since it was just me and mama and we had all these empty rooms, mama decided to let those rooms out, mostly to people passing through like from-- oh , Mobile, Montgomery place like that, that's how me and mama got money . mama was a real smart lady	[...] siccome eravamo solo io e mamma e c'erano tutte quelle stanze vuote mamma, aveva deciso di affittarle, soprattutto a gente di passaggio, tipo uh da Mobile, Montgomery, posti come questi. mamma e io facevamo i soldi così . mamma era una donna molto in gamba
31c	<i>Tootsie</i> (1982)	they have dinner-- oh c-can I just start again? I just I-I wanna-- I didn't get kicked off right	so che a di che-- uh posso ricominciare? che voglio-- non sono partito bene

31d	<i>Sideways</i> (2004)	oh we're pretty wiped, you know I think we're just gonna-- go back to the motel and crash	uh siamo un po' a pezzi io credo che ce ne--, ce ne torneremo al motel a collassare

At the same time, hesitations are also employed in Italian dubbing as an alternative to disfluency phenomena such as, for instance, repetitions (*I'm-I'm*; *they-they*; *a-a-a* and *ca-can* in extracts 32a, 32b, 32c and 32d, respectively), false starts and reformulations (extracts 32e and 32f, respectively).

Extracts 32 - Hesitations Added to Replace Disfluency Phenomena

		English Dialogue	Italian Dubbing
32a	<i>The Kids Are All Right</i> (2010)	and I'm-I'm married	e uh io sono sposata
32b	<i>The Conversation</i> (1974)	well they-they took the: crosshairs of the telescope and they lined it up on the mouths--	beh eh puntando col collimatore dal teleobiettivo mi hanno inquadrato esattamente le bocche dei due ragazzi
32c	<i>Annie Hall</i> (1977)	it's different cause you keep your own APARTMENT, because you know it's there! we don't have to go to it we don't have to deal with it but, it's like a-a-a free-floating life raft, that we know that we're not married	è diverso finché ti tieni il tuo appartamento, perché capisci, è là! noi non ci andiamo noi lo ignoriamo è come una uh uh zattera di libertà che è lì, è il sapere che non siamo sposati
32d	<i>Kramer vs. Kramer</i> (1979)	uh yeah ca-can she come in here? =	uh sì uh lei può entrare? =
32e	<i>The Apartment</i> (1960)	well ((chuckle)) that's very kind of you but I'm not feeling well I've-- see I've got this cold I'm gonna go right home	beh ((soft chuckle)) grazie lei è molto gentile ma non sto molto bene uh vede ho un raffreddore me ne vado dritto a casa
32f	<i>Silver Linings Playbook</i> (2012)	wait wait no no no what are you doing? it's-- I'm behind on my syllabus	ehi aspetta no no no no che cosa fai? eh sono indietro con il programma

Alternatively, as revealed from the examples below, vocalizations are also introduced in dubbing to replace a wide variety of features: they are inserted in lieu of a backchannel, delivering a hesitant answer in the target text (Extract 33a), or as an alternative to the attention signal *hey*, used to grab the new interlocutor's attention when launching the turn in extract 33b. Moreover, these elements are also introduced as an alternative to other sounds present in the English soundtrack, among others: clearing throat sounds (Extract 33c),

chuckling sounds (used as a turn-holding device in extract 33d), as well as nasal (Extract 33e) and glottal sounds, like in extract 33f, where a hesitation is added in Italian, explicating the character's uncertainty about what to say.

Extract 33 - Hesitations Replacing Various Phenomena in Italian Dubbing

		English Dialogue	Italian Dubbing
33a	<i>Sideways</i> (2004)	uh-uh	eh sì
33b	<i>Joker</i> (2019)	= hey Bobby would you put the lights on? =	= eh Bobby, puoi illuminare lì? =
33c	<i>Sideways</i> (2004)	oh just a little local pinot, you know a little burgundy that old cotes de Beaune ((chuckles)) ((clears throat))	no oh solo un po' di pinot nero locale, un goccio di borgogna un buon côtes de beaune ((chuckles)) em--
33d	<i>Tootsie</i> (1982)	oh god! oh it's Ron ((sighs)) oh Dorothy god bless you ((chuckles)) wish me luck now	oh dio . è Ron ((sighs)) oh Dorothy dio ti benedica eh fammi gli auguri
33e	<i>Guess Who's Coming to Dinner</i> (1967)	mh I-I: t-uh take it they've told you all about their plans?	eh io eh io immagino che vi avranno già detto dei loro piani
33f	<i>In the Bedroom</i> (2001)	((indistinct glottal sound))	((whispering)) + eh--*

In the context of film dialogue, hesitations are commonly added in exchanges where indistinct talking takes place in the original soundtrack; that is, inaudible or unclear speech used for a realistic effect, described by Kozloff (2000: 47) as “verbal wallpaper”. In dubbing, this incidental dialogue is often made explicit, as in the following extract from *Father of the Bride* (1950). Stanley's Italian turn involves the addition of apologizing units (literally: “sorry” and “no nothing nothing”) together with a turn-initial hesitation, uttered when the character accidentally bumps into one of the workers preparing the house for his daughter's wedding (Extract 34). The same also occurs in the extract from *The Kids Are All Right* (2010), where Paul asks Tanya — who is having dinner with Joni and her friend Sasha — to talk; while in the source text Paul's turn is not audible as he and Tanya walk away from the table, dialogue is added in Italian (backtranslation: “uh I wanted to ask you--”), including a turn-launching hesitation (Extract 35).

Extract 34 - *Father of the Bride* (1950)

	English Dialogue	Italian Dubbing
Man 5	GANGWAY-- OH!	oh
Stanley	((indistinct))	uh scusate ((soft chuckle)) no niente niente
Man 5	gangway!	pista! . pista!

Extract 35 - *The Kids Are All Right* (2010)

	English Dialogue	Italian Dubbing
Paul	I'm sorry but I have to take . the lovely Tanya away from a moment	<i>eh</i> mi dispiace ma devo portare via, l'adorabile Tanya per qualche minuto
Tanya	all right, it was lovely talking to you, ladies	va bene, ragazze è stato molto bello parlare con noi
Sasha & Joni	you too	anche per noi
Paul to Joni	((very low)) thanks =	grazie =
Paul to Tanya	= ((offscreen)) ((indistinct))	= ((offscreen)) uh ti volevo chiedere--

To conclude, it is important to point out that the addition of hesitation markers in the Italian dubbed soundtrack often serves as a means of achieving synchronization, that is, to synchronize the sound with the lip movements visible on the screen. Indeed, in the English version, characters may slightly open or move their lips without actually producing any sound; in the dubbed version, however, this lack of correspondence between what audiences see on the screen and what they hear (or not hear) may seem odd, impairing their suspension of disbelief. As a consequence, brief vocalizations are included. In extract 36, for example, a hesitation marker is added at the end of the turn to vocalize the character wanting to add something, slightly opening his mouth, but self-interrupts, leaving the sentence hanging. Similarly, in extract 37, a hesitation as a free-standing turn is added in dubbing to fill the actor's lip movements, conveying the idea of John wanting to take the floor to contribute to the communicative exchange, but he is blocked by Joey calling her mom.

Extract 36 - *Manchester by the Sea* (2016)

	English Dialogue	Italian Dubbing
Lawyer	Lee--	Lee uh--

Extract 37 - *Guess Who's Coming to Dinner* (1967)

	English Dialogue	Italian Dubbing
Joey to John	((chuckles)) what's the matter with you? ((laughter)) =	((laughter)) ma si può sapere che hai? ((laughter)) =
John		uh--
Joey to Mrs. Drayton	= mom!	= mamma?

As can be inferred from these examples, this may cause the insertion of lines that are not present in the source dialogue, making the dubbed version wordier, as in the following

extract from *Driving Miss Daisy* (1989), where Hoke starts to say something (*beh uh--*; backtranslation: “well uh--”) but stops when Idella takes the floor. An illustration of this can be found in extract 39 from *Mister Roberts* (1955), where Ensign Frank Thurlowe Pulver, the laundry and morale officer, is trying to impress a group of nurses. While giving the nurses a tour of the ship, Pulver pretends to be the executive officer, i.e. the second-in-command, giving unusual orders to the other members of the crew and being in charge of tasting the cook’s food. Before tasting the soup, Frank turns slightly towards one of the nurses. In the English version, this movement is not accompanied by any sound, whereas in the dubbed version, the turn *e-eh scusi* (backtranslation: “u-uhm sorry”) is added, which includes the addition of a hesitation with the repetition of the initial vowel sound.

Extract 38 - *Driving Miss Daisy* (1989)

	English Dialogue	Italian Dubbing
Hoke	((laughter)) yassum ((laughter)) <i>uh</i> ((clears his throat)) Idella said we running short on outa coffee and Dutch Cleanser	((laughter)) sì signora ((chuckles)) <i>eh</i> Idella dice che siamo a corto di caffè e anche di sapone
Daisy	we are?	siamo?
Hoke		<i>beh uh--</i>
Idella	yassum we’re low on silver polish too	e manca anche il liquido per l’argenteria

Extract 39 - *Mister Roberts* (1955)

	English Dialogue	Italian Dubbing
Pulver	<i>u:h</i> flemish up the lines out there, will you Dotty?	ah <i>ehm</i> mi mandi a rapporto il tenente, eh capo?
Dowdy	what?	cosa?
Pulver	yeah, new man	sì, è uno nuovo ... =
Pulver to Nurse		= <i>e-eh</i> scusi
	((Pulver tastes food))	((Pulver tastes food))
Pulver to Cookie	more salt Cookie, put some more salt there =	((groans)) più sale cuoco, metti più sale qua dentro =

As pointed out above in relation to the Omission strategy, the addition of hesitations to achieve synchronization can be conceived as a way to orient the text toward the target language, hence the target viewers. This also concerns instances where hesitations are used as a tool to explicate meanings or as a compensation strategy, i.e. inserting a new hesitation in sentences or locations where no vocalization is present in the English text to adapt Italian dubs to the way these pragmatic markers are used in (prefabricated) Italian. On the other

hand, the source English text seems to influence the TT in those instances where hesitations are added to the dub regardless of synchronization requirements and in cases of over-extensions, i.e. hesitations used in Italian to substitute other orality features such as interjections or disfluencies. Therefore, the addition of hesitations may reveal a greater or lesser influence of the source English language depending on the possibility vs necessity (i.e. due to synchronization) to add vocalizations to the TT.

5.3.4. Dubbing Hesitations: Substitution

Substitution is the least frequently used strategy within the overall corpus, the only exception being the 2000s, where it is slightly more frequently used than Addition (22% as against 20%). Nevertheless, as can be seen in Figure 5.8, this strategy appears to be consistently used in dubbing, maintaining for the most part a percentage of employment between 20% and 23%. Two drops occur during the decades spanning 1990 to 2010, possibly attributable to the boost detected for other solutions such as Addition and Close Rendering.

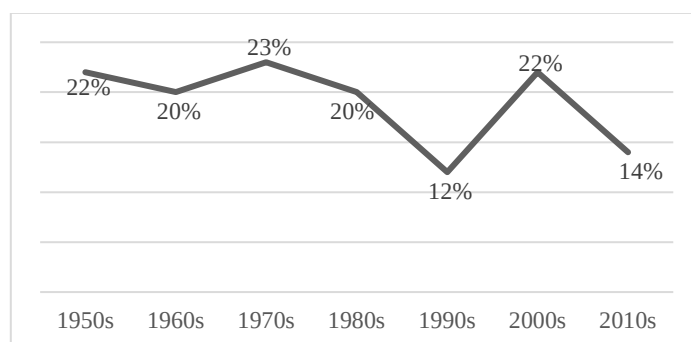


Figure 5.8. The Use of the Substitution Strategy over Time

Substitution involves the original hesitation marker being replaced with a different element in the target text. Sometimes, in the translated soundtrack, the original hesitation is substituted by a different kind of sound, for instance, a nasal sound. As can be inferred from the examples below, nasal sounds occur both as single items within the turn (40a), as well as cliticized to another item, i.e. pronounced as a continuation of the final sound of a previous element such as, for instance, the end of a pronoun (40b, where *la* translates “you”), a conjunction (40c) or a verbal element (*è* in 40d; backtranslation: “is”).

Extracts 40 - Hesitations Replaced by Nasal Sounds

		English Dialogue	Italian Dubbing
40a	<i>Mississippi Burning</i> (1988)	well yes you can I'll tell you I just um -- I hate the way I look you know? uh what do you think? [...]	ah sì ecco vi dirò io ecco-- mm non mi posso più vedere sapete io-- lei che cosa dice eh? [...]
40b	<i>Separate Tables</i> (1958)	= I'll uh book you in just for tonight then . dinner is at seven o'clock you'll hear the gong ... Edgar will show you to your room	= lamm la registrerò solo per stanotte . la cena è alle sette sentirà il gong .. Edgar le mostrerà la sua stanza
40c	<i>Jerry Maguire</i> (1996)	and uh -- . I love you	emm -- io ti amo
40d	<i>Juno</i> (2007)	no I-- uh more commercial stuff	no èmm roba molto più commerciale

In some cases, this may cause a change of function in the target version; in extract 41, for instance, the nasal sound inserted in Italian does not replicate the launch of the request performed by the hesitation marker in the source text, i.e. asking Stanley to prepare three old fashioned. In the dubbed scene, the nasal sound constitutes a backchannel to Stanley offering some martinis to the guest, who then mitigates the refusal by specifying that the girls want something different to drink. Likewise, in the extract 42, while in the English soundtrack the hesitation marker *uhm* is used with a turn-holding function, that is, to maintain the turn and add other ingredients to the list, in the Italian version the nasal sound *mh* is used as a tool to remark Miles' statement, i.e. to express that he is confident about his guess.

Extract 41 - *Father of the Bride* (1950)

	English Dialogue	Italian Dubbing
Stanley	hello hello hi hi	salve! salve! ti andrebbe un bel martini?
Man 4	uh the girls want three old fashioned	mm le donne vogliono tre gin fizz

Extract 42 - *Sideways* (2004)

	English Dialogue	Italian Dubbing
Miles	hm a little citrus ((sniffs)) maybe some strawberries uhm ((sniffs)) passion fruit mm? ((sniffs)) and-- ((sniffs)) ah there's just like, the faintest-- ((sniffs)) soupçon of like uh asparagus and ((sniffs)) there's just a flutter of like a-like a nutty edam cheese	hm ((sniffs)) degli agrumi ((sniffs)) c'è un po' di fragola mh ((sniffs)) f-frutto della passione((sniffs)) e:-- ah ci sento anche una flebile, soupçon come di asparagi, e anche una certa sfumatura di-- sembra-sembra formaggio olandese

In Italian dubbing, hesitations are commonly substituted by the lengthening of (usually) a vowel sound. This occurs, for instance, in extract 43a, where *uh* is substituted by two prolongations in “può dirci” (“can you tell us”), followed by a nasal sound (*mm*) that replaces the planning repetition of *what*. A similar case is also reported in extract 43b from the film *Father of the Bride* (1950), where hesitations are substituted by two prolongations within “e questo” (“and that”), followed by a nasal sound (*mh*) used to plan the self-correction, i.e. replacing the demonstrative *questo* (masculine ending) with the feminine form *questa*.

Extract 43 - Hesitations Substituted by Prolongations and Nasal Sounds

		English Dialogue	Italian Dubbing
43a	<i>Forrest Gump</i> (1994)	you had quite a trip can you uh tell us what-- . what was China like?	ha fatto <i>u:n</i> grosso viaggio può: dirci: <i>mm</i> com'è questa Cina?
43b	<i>Father of the Bride</i> (1950)	uh this uh -- this twenty four Maple Drive uh is it a club? or a country estate?	e: questo:-- mh questa Villa Fiorita è--? è un circolo o una tenuta?

Lengthening may involve verbal elements (Extracts 44a-c), demonstratives (Extract 44d), as well as adverbial features, as in extract 44e. This last example illustrates an interesting case where the lengthening at the end of “solo” (*just*) results in a nasal sound (*mm*), which works as a holding feature. In other words, the character has not finished talking; the speaker (Mark) is momentarily suspending the sentence to plan while looking for the right term.

Extracts 44 - Hesitations and Prolongations

		English Dialogue	Italian Dubbing
44a	<i>Twelve Angry Men</i> (1957)	well I uh thought we'd sit in order you know by jury numbers <i>uh</i> one two three four five and so on around the table, if that's okay with you gentlemen	<i>eh ho:--</i> pensato di sederci in ordine di numero uno sono io <i>uh</i> due tre quattro cinque e così via attorno al tavolo, se siete tutti d'accordo signori
44b	<i>Tootsie</i> (1982)	can I uh call you back? thank you	<i>eh . posso:</i> richiamarti? . grazie
44c	<i>The Hustler</i> (1961)	= would you uh like another cup?	= lo prende: ? un altro caffè lei?
44d	<i>The Verdict</i> (1982)	if she HAD eaten uh an hour before being admitted to the hospital, then the inducement of a general anaesthetic <i>u:h</i> the kind that you gave her would have been--, negligent?	se invece avesse mangiato-- . u:n' ora prima di-di essere ricoverata in ospedale, allora praticarle l'anestesia generale che le avete praticato sarebbe stata--, negligenza
44e	<i>Juno</i> (2007)	I didn't do anything .. just um ((sighs)) ... I've been thinking!	non ho fatto niente .. ho solo:mm ((sighs)) ... riflettuto

As previously mentioned in §5.3.3, hesitation markers are often interchanged with other orality features, like interjections. As illustrated in the extracts below, interjections can replace hesitations to launch the turn (45a-b), to retake the turn after a moment of silence (45c) or as a backchannel sound. The latter case occurs, for instance, in extract 45d, where Frank Galvin is talking on the phone; the original *uh* has been replaced by an interjection, probably used to provide some feedback to the interlocutor on the phone (ah-ah); nevertheless, the interlocutor's speech is not audible to either the original or the target audience.

Extract 45 - Hesitations Replaced by Interjections in Italian Dubbing

		English Dialogue	Italian Dubbing
45a	<i>Mister Roberts</i> (1955)	uh thanks . woul-would <i>uh</i> you like to come out to the ship this afternoon?	ah eh beh grazie tante <i>eh</i> non vorrebbe:-- non vuole venire a bordo nel pomeriggio?
45b	<i>The Apartment</i> (1960)	uh look kid I-I put in a good word for you with Sheldrake in personnel	ah sa che ho messo una buona parola per lei con Sheldrake del personale?
45c	<i>Manchester by the Sea</i> (2016)	((chuckles)) ((whispering)) +just fucking with you* ... yeah .. um . do you think I should call my mom? to tell her about Dad?	((chuckles)) molto bene ... ciao . a:h .. credi che dovrei chiamare mia madre? per dirle di papà?
45d	<i>The Verdict</i> (1982)	[...] well <i>uh</i> we've had a strange change of strategy . uh w-- I was just wondering, if you could <i>um</i> -- well this kinda short notice at any rate	[...] . dunque c'è stato un gua- - un cambiamento di startegia sì . ah-ah beh io mi stavo chiedendo <i>semm eh</i> se lei poteva insomma ((soft chuckle)) beh% è breve come preavviso comunque--

Interjections are commonly employed to substitute hesitations when the latter are used as free-standing items. For example, in the following cases (Extracts 46), interjections are employed as backchannel signals in response to the information provided by the interlocutor.

Extract 46a - Sideways (2004)

	English Dialogue	Italian Dubbing
Maya	about two thirty ... hey, did you hear about this bordeaux tasting dinner down in santa barbara on Saturday night?	alle due e mezza ... sabato a santa barbara c'è una cena con degustazione di bordeaux
Miles	uh	ah

Maya	it's kind of pricey but, if you wanted to go I'd be into it .. why don't you just stay for the weekend?	costa un po' ma se ti va di andarci io ci sono .. vuoi restare per il weekend?
Extract 46b - Heaven Can Wait (1978)		
	English Dialogue	Italian Dubbing
Betty	I want to know what you're going to do about Pagglesham	io voglio sapere che intende fare per Pagglesham
Joe	uh--	ah, eh?
Betty	PAGGLESHAM! PAGGLESHAM! it's what we're talking about!, [...]	PAGGLESHAM! PAGGLESHAM! è di questo che stiamo parlando! [...]*

In some instances, vocalizations are also substituted by different types of sounds, for example, a sigh (Extract 47a), a character clearing his or her throat (Extract 47b), as well as a glottal sound (signalled by the % symbol in extracts 47c-d).

Extracts 47 - Hesitations Substituted by other Sounds

		English Dialogue	Italian Dubbing
47a	<i>Mississippi Burning</i> (1988)	we got a real peaceful community down here Anderson, course they're just like any other folks I reckon <i>uh</i> you push them too far, the way I figure it uh , it's like three sticks of old dynamite, you shake it up, and we're gonna be scraping bodies off the street	abbiamo una pacifica comunità quaggiù Anderson, certo siamo come tutti gli altri quando ci pestano troppo i piedi, per come la penso io-- ((sighs)) siamo un po' come i vecchi tubi di gelatina, scuotili un po', e raschieremo brandelli di carne dalla strada
47b	<i>Heaven Can Wait</i> (1978)	hi ... <i>u:h</i> um ... ((clears throat)) . hi how you doing everybody? [...]	salve ... <i>u:h</i> ((clears throat)) ((indistinct sounds)) ... salve come state tutti quanti? [...]
47c	<i>Father of the Bride</i> (1950)	hello? hello? operator w-w-will you get me Jake's Garage please? . well uh well keep on trying will you? please? kee-- <i>uh</i> we-we-we've got a wedding and-and no way to get to the church . no cars, will you do that? . thanks very much	pronto? pronto? centralino mi-mi-mi dia il garage Jackson prego . beh% -- beh provi ancora la prego, provi ancora <i>um ehm</i> do-- abbiamo un matrimonio e dobbiamo sbrigarci ad andare i-in chiesa, non ancora? ci pensa lei? . grazie infinite
47d	<i>Twelve Angry Men</i> (1957)	well this is just off the top of my head butuh , well if I were the boy and I'd done the stabbing and everything I'd take a chance and go back for the knife [...]	ecco la prima risposta che mi viene ma% , se fossi stato io a commettere il delitto sarei senz'altro ritornato a prendere il coltello [...]

At the same time, hesitations are commonly substituted in favour of disfluency phenomena such as repetitions. In these cases, the disfluency mainly involves grammatical elements such as subject pronouns (48a), personal and relative pronouns (48b-c, respectively), articles (48d), and conjunctions (48e).

Extracts 48 - Hesitations Substituted by Repetitions

		English Dialogue	Italian Dubbing
48a	<i>Jerry Maguire</i> (1996)	oh Jer I'm-- I'm three months away from the pay increase ((whispering)) + I-- uh *, sorry	oh, Jerry io-- mi mancano tre mesi per l'aumento di stipendio e io-- io-- .. mi dispiace
48b	<i>Forrest Gump</i> (1994)	[...] he uh wrote you a-- . a letter .. he says I can't read it, I'm not supposed to so I'll just leave it here for you [...]	[...] sai ti-ti ha scritto u-una lettera .. dice che non posso leggerla, non devo farlo perciò la-la lascio qui per te [...]
48c	<i>Tootsie</i> (1982)	to Michael um who's been my friend for six years oh, was it that long? and who is my coach and he's just great he's a great coach he's a great actor he's a great guy [...]	a Michael che-che è mio amico già da sei anni oh ((chuckles)) è tanto tempo, che è il mio maestro e che è grande! ecco è un grande maestro un grande attore ed è un gran ragazzo! ((chuckles)) [...]
48d	<i>Annie Hall</i> (1977)	((coughs)) thank you I-I don't know why they would have me at this kind of rally cause-- ((clears throat)) excuse me I'm not essentially a-- political, comedian at all, I-- interestingly had um dated, a woman in the Eisenhower administration [...]	((clears throat)) grazie io-io non so perché mi abbiano richiesto per questo congresso ((clears throat)) perché io non sono essenzialmente un comico, politico per niente, di natura politica ho avuto solo u-u-una relazione con-con una donna durante, l'amministrazione di Eisenhower [...]
47e	<i>Silver Linings Playbook</i> (2012)	[...] I then realised that oh oh wow you know I've been dealing with this my whole life . and uh and-- without any supervision I've been doing it all on my own uh with no help, and um you know I basically have been living like white knuckling it this whole time	[...] e allora mi sono reso conto che wowowo wow ecco io combatto con questo da tutta la vita . e-e : senza alcuna assistenza ho affrontato tutto da solo, e senza un aiuto e-- e-e e insomma praticamente ho vissuto a pugni stretti tutta la vita

In addition to grammatical features — i.e. conjunctions, pronouns, etc. (Extracts 49a-b) — hesitations are substituted by a variety of arbitrary elements such as the filler word *ecco* (“well”) (Extract 49c), adverbs like *allora*, *poi* (both corresponding to “then”) and *anzi* (“in fact”) (Extracts 49d-f), the telephone opening *pronto* (“hello”) used to attract the interlocutor’s attention and make sure that the communicative channel between speakers is

still open (phatic function in extract 49g), or items such as *sì/no* (“yes/no”) as both adverbial elements (49h-i) and response forms (49j).

Extract 49 - Hesitations Substituted by Various Linguistic Features

		English Dialogue	Italian Dubbing
49a	<i>Sideways</i> (2004)	uh me? no, no no I'd just fuck 'em up, yeah yeah, no that was the one, unpolluted part of my divorce, no kids	chi io? no, no no no no gli insegnerei solamente a sbagliare sì, sì, è l'unica parte, incontaminata del mio divorzio, niente figli
49b	<i>The Conversation</i> (1974)	I was paralyzed in my left arm and my left leg uh I couldn't walk for six months	o avuto una paralisi al braccio sinistro e alla gamba sinistra e , non ho potuto camminare per sei mesi
49c	<i>Sideways</i> (2004)	oh uh well I like them both, but um if pressed I would have to say that I prefer the dark	oh eh sono tutt'e due ottime, ma ecco , se fossi costretto a scegliere andrei sul cioccolato
49d	<i>In the Bedroom</i> (2001)	are you-are you saying that-- .. that I'm the one, responsible? .. is that it? well uh -- .. well let me tell you something, let me tell you something, you got it backwards! [...]	vorresti-vorresti dire che-- .. in qualche modo è mia, la responsabilità? .. è così? beh allora .. ti dico una cosa se permetti, ti dico una cosa se permetti, è tutto il contrario semmai! [...]
49e	<i>The Kids Are All Right</i> (2010)	well it was like-- a lot of money to me at the time and uh with inflation it's like ninety dollars today . hey ... I'm glad I did it	beh per me, all'epoca erano parecchi soldi e poi -- con l'inflazione oggi sarebbero più di novanta dollari, ehi ... sono felice di averlo fatto
49f	<i>Annie Hall</i> (1977)	I'm not one of those characters you know although I'm balding slightly on top that's about the WORST you can say about me I-- uhm I think I'm going to get better as I get, older you know	io-io non mi preoccupa invecchiare non sono uno di quei tipi sapete lo so quassù mi si apre una piazzetta ma peggio di questo per ora non mi è successo io, anzi credo che migliorerò invecchiando ecco sapete
49g	<i>Jerry Maguire</i> (1996)	((offscreen)) Dorothy? uh hello?	((offscreen)) Dorothy? pronto? pronto?
49h	<i>Annie Hall</i> (1977)	uh you know what I've got to-- get there and begin whining soon otherwise I-- hey are you busy Friday night?	sì sai devo guadagnarmi-- . il caviale quotidiano, hai- hai da fare venerdì sera?
49i	<i>Mister Roberts</i> (1955)	now wait a minute Doug listen I-I wouldn't send a letter like that in if I were you boy you're gonna get yourself in trouble	aspetta Doug aspetta io ti consiglio di non mandare una domanda con un tono simile ti vuoi mettere nei guai e parlare

		saying things like disharmony aboard this ship . um you know if-if you didn't say disharmony or ha-harmony--	di disarmonia a bordo della nave? no sai se-- se ne parli di- - disarmonia o armonia o quel che credi--
49j	<i>American Beauty</i> (1999)	no no, no uh actually I'd like to fill out an application	no no, no no, no a dire la verità vorrei compilare una domanda d'impiego

As previously mentioned, the dubbed version of a cinematic product has oftentimes been recognized as more verbose, with lines being added or extended. In this regard, hesitations are sometimes omitted in favour of the insertion of extra items such as, for instance, the introduction of *è così* (“that’s it”) in extract 50a and *perché puoi* (literally, “because you can”) followed by a sniffing sound in extract 50b. This may also lead to a hesitation occurring as a free-standing item to be completely eliminated for the sake of a more talkative turn. For example, this is illustrated in extract 51 from *Heaven Can Wait* (1978), where Joe’s attempt to take the floor is rendered through the sentence *ma no non--* (“but no don’t--”), left unfinished because Betty refuses to cede the floor.

Extract 50 - Hesitations Substituted by Discursive Elements

		English Dialogue	Italian Dubbing
50a	<i>Silver Linings Playbook</i> (2012)	because uh-- .. it's just not right lumping you and I together it's- - [...]	perché è così . è: che non è giusto accomunare te e me <i>e:huh</i> [...]
50b	<i>Annie Hall</i> (1977)	what do you mean where am I? what-- where do you think I am? I'm-I'm at the-I'm at the Los Angeles airport, I flew in ((sniffs)) I-- well I flew in to see YOU ((<i>indistinct hesitating stutter</i>)) hey, listen can we not debate this on-on the telephone because I-- [...]	come sarebbe dove sono? dove credi che sia? so-so-sono qua sono all'aeroporto di Los Angeles ho preso l'aereo ((sniffs)) io-io sono venuto per vederti perché puoi-- ((<i>sniffs</i>)) ehi senti potremmo non discuterne al-al-al telefono perché io-- [...]

Extract 51 - *Heaven Can Wait* (1978)

	English Dialogue	Italian Dubbing
Betty	very funny mister Farnsworth if you think I'm gonna turn around and go back to Pagglesham you're wrong	divertente signor Farnsworth ma se lei crede che io giri la schiena e me ne torni a Pagglesham si sbaglia
Joe	uh--	ma no non--
Betty	I know you're not afraid of me [...]	lo so che non ha paura di me [...]

In line with the view that Substitution is a translation strategy that orients the translated text toward the target language (see Pedersen 2011: 101), the examples illustrated above seem to be in accordance with the tendency of dubbed dialogues to be domesticated. More specifically, original vocalizations are substituted by a variety of features probably more in line with (prefabricated) Italian, i.e. hesitation markers are replaced by linguistic and paralinguistic items that serve to the naturalness of the dubbed text.

Thus far, the analysis of how hesitations are treated in Italian dubbing has highlighted the fact that vocalizations are most often omitted in the target versions. On the one hand, this may be attributable to the tendency of the Italian language to give preference to other elements to convey some of the meanings expressed through hesitations in the English language. As pointed out in Chapter One, phenomena such as prolongations are more likely to be used in spontaneous Italian instead of vocalizations (Tissi 2000; Di Napoli 2020; Schettino & Eklund 2023). In this regard, the source English text appears to influence Italian adaptations to some degree, as suggested by the fact that, in Italian dubbed film speech, hesitations are often added to the text and used to replace a variety of phenomena occurring in the source soundtrack (i.e. hesitations are utilized in lieu of interjections, repetitions, false starts, etc.). Nevertheless, in most cases, the Italian version seems to prefer solutions that bring the dubbed product closer to the target language and to the naturalness of Italian viewers so as to guarantee their immersion in the story. When vocalizations are not transposed to the target version, a wide variety of features are employed to compensate for and reproduce the sense of orality conveyed by hesitations in the source soundtrack. These include lengthenings, interjections, filler words, discourse markers, repetitions, and small sentences that make the dubbed version more verbose.

In this respect, prosody also plays a role. For example, turn-medial hesitations can be omitted in favour of voice inflections, indicating that characters are momentarily suspending their sentence but intend to continue saying something. In these cases, intonation and vocal performance allow the speaker to temporarily withhold the continuation of the utterance without adding any vocalization to the target text, as in the following case from *Mississippi Burning* (1988). Figure 5.9 shows the waveform of the pronunciation of the Italian preposition *di* — which occurs in the same position as the original hesitation — highlighted in red (sounding section). The waveform retrieved through the speech analysis software Praat shows regular, closely spaced, and fairly uniform oscillations; the spectrogram (the grey central part) also illustrates a regular pitch with a slight upward trend

(blue line) and a rather constant intensity (yellow line). Overall, the use of a continuous vowel sound allows the speaker to momentarily suspend the sentence and withhold the turn while carrying out a planning process before continuing with the utterance.

Extract 52 - *Mississippi Burning* (1988)

	English Dialogue	Italian Dubbing
Anderson	two . well I was, as I remember . it didn't last very long I was never home, I guess she got fed up with uh . phone calls from Miami and postcards from Des Moines [...]	due grazie, beh lo ero, a quanto mi ricordo . durò poco non ero mai a casa . io credo che fosse stufa di-- . telefonate da Miami e cartoline da De Moines [...]

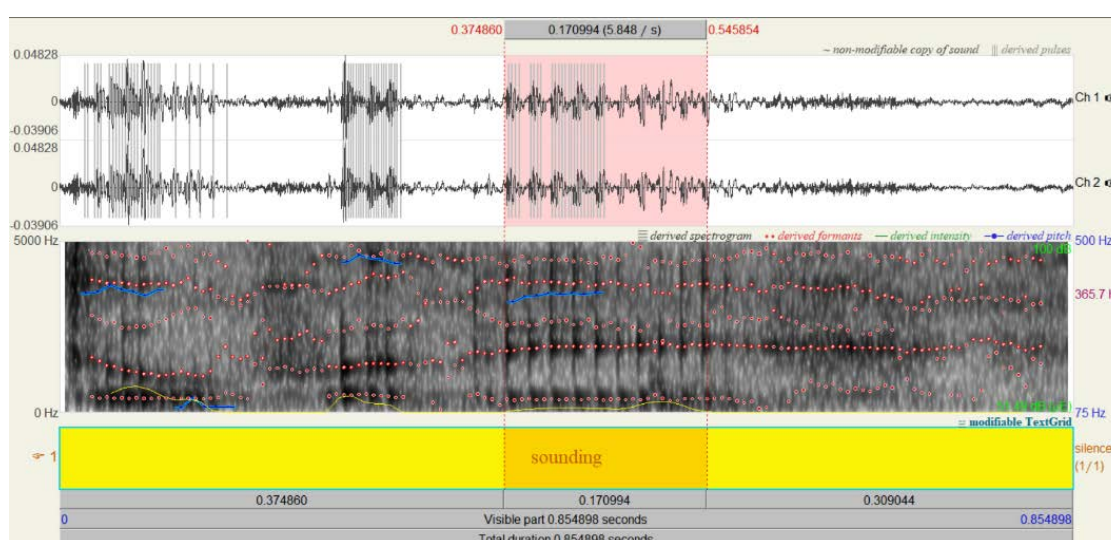


Figure 5.9. Waveform of *di* in Extract 52

Likewise, Figure 5.10 shows the waveform of the enunciation of *perché* (“because”) — which occurs in the same position as the original hesitation in extract 53 from *American Beauty* (1999) — highlighted in red. The waveform displays regular ongoing oscillations; the spectrogram shows a periodic sound with a stable formant structure (red dots), moderate intensity values (yellow line) — i.e. a non-emphatic segment — and a pitch (blue line) that exhibits a slight downward trend. Interestingly, toward the end of the segment, the falling registered in pitch is accompanied by a slight decrease in intensity as well. This downward trend seems to correspond to a local prosodic closure; hence, by concluding an intonation unit, the speaker is momentarily suspending her sentence in order to plan how to continue.

Extract 53 - American Beauty (1999)

	English Dialogue	Italian Dubbing
Carolyn	((cries)) no I'm glad because uh ... because you're old enough now to learn the most important lesson in life [...]	((cries)) no sono contenta perché-- . ((inhales)) perché ora sei abbastanza grande da imparare la più importante lezione della vita [...]

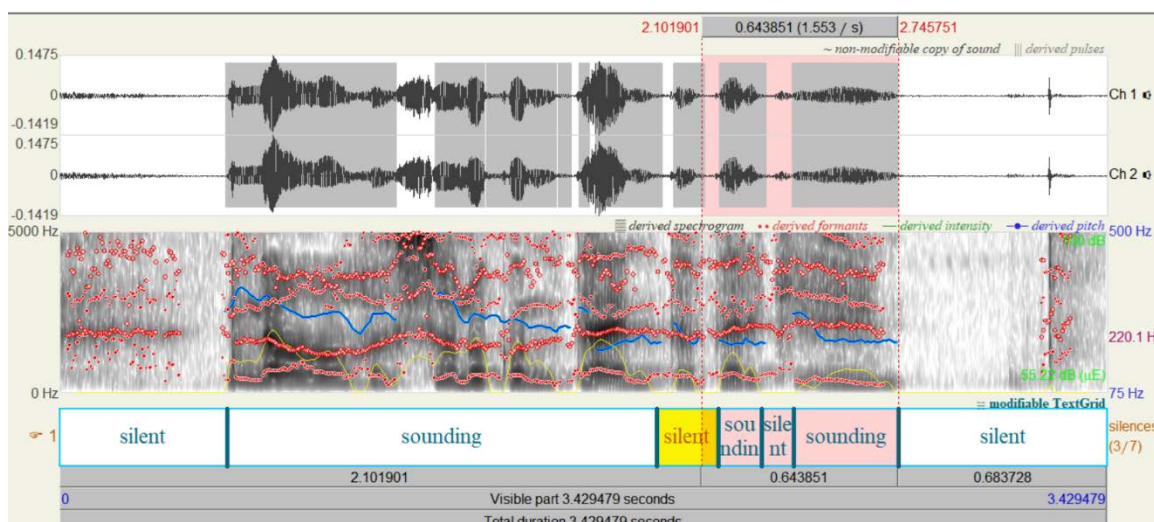


Figure 5.10. Waveform of *perché* in Extract 53

Finally, the following example from *The Graduate* (1967) illustrates a case of a turn-initial hesitation omitted in dubbing (Extract 54). In particular, as can be seen from the extract below, the sentence analysed here (i.e. *è-- in fondo al corridoio*) is an answer to a question. Even in this case, it is voice inflection that substitutes the information originally conveyed through the hesitation marker *uh*. The waveform of the articulation of *è* (which occurs in the same position as the original hesitation), highlighted in red, shows regular and continuous oscillations, with an initially larger amplitude and a slight decrease toward the end. Figure 5.11 illustrates a voiced and periodic vowel segment with a stable formant structure (red dots), a continuous pitch characterized by a slight downward tendency (blue line), and a moderate and fairly stable intensity (yellow line) that, toward the end of the segment, exhibits a gradual decrease. In this case, the downward trend registered in both pitch and intensity can be interpreted as a sign of a local prosodic boundary within the utterance; in other words, the initial vowel segment *è--* corresponds to an autonomous sound unit within the utterance. The sentence analysed here thus presents a bipartite structure: the copula *è*, followed by a momentary break (0.633171 seconds; this has not been interpreted as an unfilled pause but as a transitory break in moving from a prosodic closure to a re-opening with the following phrase) that separates it from the prepositional phrase (i.e. *in fondo al corridoio*). Hence, the prosodic realization of *è* serves a discursive and planning function,

i.e. introducing the prepositional phrase that conveys the informational content of the utterance. Therefore, the selected segment (highlighted in red) exhibits acoustic characteristics attributable to low informational prominence; the segment realizes prosodically the function of an introductory unit that prepares the interlocutor for the relevant content, i.e. the answer to the interlocutor's question.

Extract 54 - *The Graduate* (1967)

	English Dialogue	Italian Dubbing
Mrs. Robinson	oh, I guess this isn't the bathroom is it?	oh ((soft chuckle)) non mi pare che questo sia il bagno vero?
Ben	it's uh down the hall	è-- in fondo al corridoio

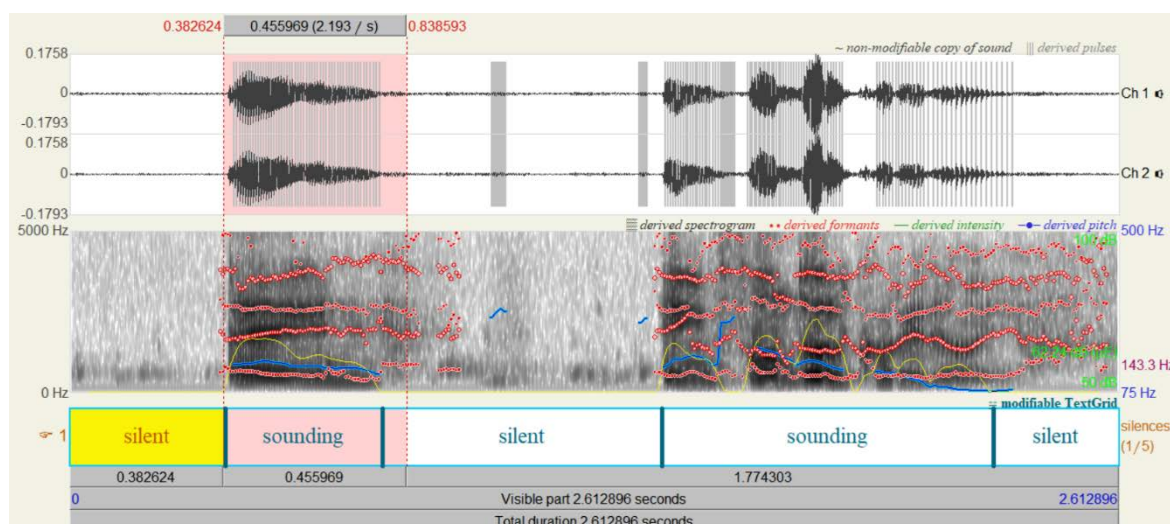


Figure 5.11. Waveform of *è* in Extract 54

Conclusions

This research aimed at examining the diachronic evolution of orality in original and dubbed film speech by focusing on a phenomenon still understudied in AVT, i.e. hesitations.

The analysis of hesitations in prefabricated dialogues has required a crucial initial step: the theoretical description of the phenomenon with respect to its distribution, functions and meanings in spontaneous speech (Chapter One). Drawing upon Brennan & Williams (1995), Poyatos (2002a, 2002b) and Culpeper & Kytö (2010), in this study, hesitations have been described as paralinguistic markers capable of conveying a variety of pragmatic meanings and as a phenomenon that needs to be investigated as a category in its own right. The next step was to present the specific context in which hesitations were to be examined, i.e. film speech. In particular, in order to survey occurrences of hesitations in pre-planned dialogues, audiovisual texts were analysed by focusing on prefabricated orality in both original and dubbed products, observing dialogue in relation to its compliance with the codes of naturalness and realism and examining its diachronic evolution (Chapter Two).

As pointed out by scholars like Pavesi (2019b), Werner (2021) and Zanotti (2022), in AVT, diachronic studies are still underrepresented, especially with regard to translation norms and patterns in adapting orality features for an international audience. Among such features, hesitations remain underexamined with respect to the degree of adherence of film speech to spontaneous conversation, as well as from a translational point of view. In order to fill this void, a corpus of 28 original American English-language films distributed over the span of 70 years and then dubbed into Italian was compiled. In devising the methodology for the investigation, it was paramount to guarantee the corpus to be balanced and comparable; hence, four films per decade were included: two comedies and two non-comedies.

Drawing upon the results discussed in Chapter Four and Chapter Five, insights obtained from the quantitative and qualitative analysis of hesitations in original and dubbed film speech will be presented hereafter with respect to the five research questions addressed in §3.1.1:

RQ1: What are the mechanisms of codification of orality in English-language film speech with respect to hesitation markers?

RQ2: What functions do hesitations fulfil within the mechanisms of interaction in the cinematic context?

As for the way orality is codified with respect to hesitation markers in original Anglophone dialogues, the data reviewed in Chapter Four seems to be in line with Forchini's (2012) findings, according to which spontaneous and movie conversations "do not differ to a great extent" (*ibidem*: 119). More specifically, insightful findings were obtained in relation to where and how hesitations are uttered in film speech compared to spontaneous conversation: in film speech, hesitations more frequently occur at the beginning and in the middle of a turn; on the other hand, they rarely occur at the end of a turn and as free-standing items. Most notably, in line with the data obtained from spontaneous conversation (Tottie 2014a: 393; Boonsuk *et al.* 2019), hesitation markers appear in the following order from most to least frequent: turn-medial, turn-initial, turn-final and single-turn.

Based on the functional categories identified by previous research with regard to hesitations in both spontaneous conversation and fictional dialogue (see Chapter One), the multiple functions hesitations fulfil in film speech were organized in five domains: Structural, Planning, Interpersonal, Cognitive, and Hesitation Proper Domain (see §3.6). The qualitative analysis of hesitations has pointed out how these features mainly fulfil functions categorized within the Structural and the Planning Domains. In particular, with regard to the two most frequent locations, turn-initial hesitations most commonly fulfil functions within the Structural Domain, while turn-medial hesitations typically fulfil functions within the Planning Domain. In accordance with spontaneous speech, the 'manual' observance of hesitations in context has underlined that in film speech these elements are routinely uttered at the beginning of the turn as turn initiators, i.e. as turn-taking and turn-launching devices (see Boomer 1965 and Hawkins 1971; Stenström 1990: 227; Kjellmer 2003: 184), often co-occurring with utterance launchers, as is also the case in impromptu conversation (Biber *et al.* 1999: 1054); moreover, they are often employed in launching answers to questions (Tottie 2014a: 402) and to help the speaker perform multiple speech acts within the same conversational turn. In addition, hesitations are usually employed in the middle of a turn as a turn-holding device and, more commonly, as a planning device (Biber *et al.* 1999: 1054; Tottie 2014a; Boonsuk *et al.* 2019), i.e. to plan and organize how to construct the following part of the sentence (Clark 2004), giving the speaker enough time to think about how to continue and extend the body of his or her speech.

Hence, in film speech, hesitations mainly fulfil discourse functions. This was corroborated by the empirical analysis carried out through AntConc (see §4.4 for collocation and concordancing), indicating that, as in spontaneous conversation, hesitations often accompany features attributable to the interactive nature of verbal exchanges (Biber *et al.* 1999: 1046), i.e. discourse particles and conversational inserts such as *well*, *okay* and *oh*, *yeah* and greetings. In this regard, the planning function of the pragmatic markers studied here also emerges from the fact that they tend to be used in conjunction with verbal elements, pronouns, articles and prepositions, i.e. elements that bring together larger linguistic units and clauses (*ibidem*: 1054-1055), hence major planning points. Moreover, as pointed out by Kjellmer (2003: 182) in relation to spontaneous speech, concordancing also provides evidence of the fact that, in film speech, hesitations tend to be used next to each other.

These findings, in combination with the discursual functions more frequently found in the corpus (i.e. Planning and Structural Domains), corroborate the idea that hesitations are used as discourse tools; that is, they are employed by speakers in the process of building up their utterances, signalling transition in adjacent parts of the discourse and in the interactive development of the dialogue.

These considerations also allow to further reflect upon the status of hesitations as a whole. The results obtained from this study seem to support the claim that hesitations do not hinder communication; on the contrary, they are employed as markers of conversational strategies used to move forward with the verbal interaction and the on-screen narration. As a matter of fact, based on the data reported in Chapter Four, the meanings associated with the Hesitation Proper Domain — whose functions concern hesitations being used as disfluencies, i.e. a blockage to fluent speech — are not inherent to these pragmatic markers; rather, they represent an additional value based on the specific situation depicted on-screen and aimed at specific characterization purposes: in *Mister Roberts* (1955), for example, to characterize Ensign Pulver as being intimidated by the ship captain, or in *Tootsie* (1982) to characterize the protagonist, Michael Dorsey, in situations where he is dressed up as his alter ego (Dorothy Michaels) and risks being discovered.

At the same time, results allow to reflect upon the debate concerning hesitations and delays described in §1.4.2 (Clark & Fox Tree 2002). The qualitative analysis presented in Chapter Four has highlighted the multiple pragmatic meanings hesitations can convey; although hesitations can be employed to procrastinate the continuation of the sentence (vocalizations commonly followed by pauses), data do not seem to support the claim that the basic meaning of hesitation markers is that of signalling the speaker's intention to

introduce a delay. The ‘delay effect’ rather appears to be added to other structural and/or planning functions — for example, to delay the launch of a turn or the next part of the sentence, allowing speakers to stop and think about how to continue (i.e. planning process). In line with Schegloff’s (2010: 171) view, when this occurs, hesitation markers seem to embody the delay rather than announcing it.

RQ3: Does the way hesitations are used in film speech change over time? If so, does this provide some information regarding the evolution of film dialogue with respect to orality features? Does the genre of the film influence the frequency of occurrence of the phenomenon?

As for the diachronic evolution of Anglophone film speech with regard to orality features, quantitative data show that there is no linear increase in the amount of hesitation markers contained in cinematic dialogue. Frequencies tend to increase and decrease over time; these fluctuations, however, are not statistically significant. Most notably, the frequency of occurrence of hesitations in more recent titles does not differ much from the frequency registered in older titles. As discussed in Chapter Four, regardless of the sharp increase found in the 1970s/2000s and the sharp decrease registered in the 1990s, the level of employment of hesitations in Anglophone film speech has remained rather stable: normalized frequencies correspond to 5.064 in the 1950s and to 5.555 in the 2010s; in other words, in more recent titles the phenomenon is only 9% more present than it was in older titles. Thus, hesitations are not a characteristic feature of orality of more recent films; they appear to be an integral and lasting aspect of film dialogue, a phenomenon that was already used in older productions.

These results are in line with Zago’s (2016) findings concerning colloquiality in English film speech as regards original American films and their remakes. In particular, the author claims that “the degree of colloquiality is already significant in the original films of the corpus” (*ibidem*: 167), that is, in the 1950s and in the 1960s. Colloquiality thus appears as “a specificity of American cinema, which seems to have shown a noticeable degree of proximity to spontaneous colloquiality since at least the ‘50s” (*ibidem*). Similarly, Veirano Pinto (2014) states that the results obtained from her investigation “seem to point in the direction of the stability of the register, suggesting that the verbal language incorporated in mainstream North American films has not changed substantially in nearly a century [...]” (*ibidem*: 141).

This stability has also been observed in relation to genre; more specifically, as reported in §4.1.2, although hesitations tend to be more present in comedies than they are in non-comedies, film genre does not seem to impact the rate of occurrence of hesitation markers in a significant way. On the contrary, the two genre categories exhibit similar trends in the way frequencies change through the years. Differences emerge in relation to the extent of variations from decade to decade: while non-comedies exhibit a more regular trend, comedies exhibit more extreme variations over the years. In this respect, the results obtained from this investigation are in line with Forchini's (2012) findings, showing that, although comedies are slightly more similar to spoken language than non-comedy films (*ibidem*: 101), "movie genre does not significantly influence the resemblance of movie conversation to face-to-face conversation" (*ibidem*: 96).

The stability found within the language of Anglophone films with respect to both decade and genre seems to suggest that fluctuations are not attributable to a structural transformation of film language over time depending on genre; rather, they seem to be influenced by factors like actors' style of performance and/or directing choices, linguistic and stylistic preferences in relation to characters and the narration of events, as well as linguistic fads and inputs attributable to the socio-cultural context (i.e. contemporary speech) and to changes occurring within the movie industry. Most notably, film speech is a type of discourse that exhibits stable, repetitive and conventionalised patterns that make it similar to the type of discourse it wants to reproduce, i.e. impromptu conversation (Zago 2021: 179). As pointed out by previous scholars (Forchini 2012; Zago 2021), the stability found in the corpus seems to suggest that spontaneous and cinematic conversation do not differ to a great extent: movie language is not as artificial as it is often assumed to be; it actually exhibits a significant amount of spontaneity (Forchini 2012: 119).

RQ4: How are hesitations treated in Italian dubbing? Can regularities in translation behaviour be observed?

As for the way hesitations are rendered in Italian dubbing, four translation strategies were identified, i.e. Close Rendering, Omission, Substitution, and Addition (see §3.7 in Chapter Three). The analysis of the Italian section of the corpus corroborates the tendency of hesitations being overlooked in dubbing. In particular, the quantitative analysis reveals that dubs do indeed present fewer tokens of hesitations compared to the corresponding English text, while the qualitative analysis reveals that Omission is the most frequently used strategy

throughout the years. The results obtained from this study thus seem to be in line with previous research, according to which pragmatic markers are typically omitted in dubbing (see Galiano 2024) and hesitations are commonly reduced by either omitting or substituting them with other elements such as repetitions, discourse markers, filler phrases and indistinct noises (Romero-Fresco 2007; Valdeón 2008; Bruti 2019). Interestingly, in the corpus, Substitution is the least frequently used strategy; nevertheless, it exhibits the most consistent percentage of employment over the years. Thus, even if not as common as other strategies, Substitution has been consistently used over decades. In particular, hesitations tend to be interchanged with monosyllabic items such as interjections and replaced by a range of elements: different types of sound (for example, a nasal sound, a sigh or a glottal sound), disfluency phenomena (for example, repetitions), filler words, adverbs or phrases, and the lengthening of (usually) a vowel sound. Interestingly, prolongations have been identified as an intrinsic feature of dubbing in other languages besides Italian; for example, studies conducted on dubbed Spanish (Baños Piñero 2014: 415; Sánchez-Mompeán 2020: 116-117) identify lengthening as a conscious technique used by dubbing actors to achieve isochrony and good lip-synch. Thus, the use of prolongations at the end of a word to render hesitations could be understood as a characteristic of dubbing language.

Results obtained from the quantitative and qualitative analysis of hesitations in the dubbed section of the corpus appear consistent with established perspectives concerning systematic linguistic and translational behaviours observed in Italian dubbed film speech, such as SL influence and TL orientation (see §2.6 on this). In particular, quantitative data suggest that the occurrence of hesitations in dubbing seems to be partially influenced by the English text; this can be observed, for instance, in those cases where hesitations are added to the Italian text to replace interjections in the English soundtrack. Thus, although vocalizations are not as common as other phenomena in spontaneous Italian, they are sometimes added to the Italian text to substitute monosyllabic items such as interjections that could easily be maintained in the TT. The influence of the SL is even more noticeable with respect to the fact that, with the exception of the 1960s, frequencies of occurrence of hesitations in Italian dubbing follow the same trends registered with regard to English dialogues; that is, the phenomenon increases and decreases in dubbing in a similar way to what happens in Anglophone film speech over the years. This appears to be in line with Motta's (2015: 960-961) view, according to which changes attested in Italian dubbing reflect changes of the texts being dubbed, i.e. they reflect changes occurring within American

productions. In general, as for English film speech, also in Italian dubs, the small variations noticed over the years cannot be attributed to a systematic temporal trend.

At the same time, the qualitative analysis of hesitations in dubbing also seems to confirm the well-established view that Italian dubbing is oriented toward the target language system (Pavesi 2009: 129); this is specifically found when vocalizations are relocated, omitted, added or substituted to respect synchronization requirements related to lip-sync, kinesic synchrony and isochrony. To put it in Chaume's (2012: 69) words: "The implementation of these three types of synchronization [...] is the result of a conscious agenda to domesticate the translated text, so that viewers are oblivious to the fact that what they are witnessing on screen is a translation". Thus, even though there are instances where Italian is affected by the English language, the way hesitations are treated in dubs is often influenced by technical requirements (i.e. synchronization) and changes oriented toward the target language, i.e. slightly changing the location of hesitations in the dubbed text so as to approximate the rendering to Italian prefabricated orality or by resorting to solutions more closely linked to the Italian pragmatic dimension such as prolongations and filler words.

RQ5: Does the way hesitations are rendered (or not rendered) in Italian dubbing change over time? If so, does this provide some information regarding the evolution of Italian dubbing with respect to orality features? Does the genre of the film influence the frequency of occurrence of the phenomenon in dubbed Italian?

According to the results obtained from the qualitative analysis, a change occurs with respect to the strategies used in dubbing over the years: while in earlier dubs Addition was the preferred strategy next to Omission, from the 1990s Close Rendering becomes the second most used strategy. Thereby, while in earlier dubs the Italian dubbing industry appeared more focused on adapting orality features to the system of the target language through the addition of hesitation markers (i.e. domestication to make dialogue sound natural), in more recent dubbings greater attention seems to be paid to the orality of the original text. In other words, Italian dubbing — and possibly audiences as well — seems to have become more receptive to the way orality features are used in the source English soundtrack. These findings are in line with some of the considerations made by Minutella (2021b) with respect to the presence of pragmatic Anglicisms in the form of interjections such as *okay* and *wow* in Italian dubbing. More specifically, the author points out how Italian dubbing has become

more absorbent of the source text when it comes to pragmatic markers, which are more frequently used in dubbed animated films compared to the past to perform a variety of functions (*ibidem*: 84-85). Similar considerations have also been made with regard to familiarizers: Forchini (2013) notes that these elements are usually omitted in dubbing; nevertheless, when translated, more recent dubs mirror in a more precise way the function these elements fulfil in the original language as opposed to older dubs that were not as functionally accurate. Thus, more recent dubbings seem to be paying more attention to pragmatic equivalence compared to the past (*ibidem*: 522).

It must be noted that the considerations presented above with regard to Italian dubbing are not aimed at examining how good the Italian adaptation of each film is or how good Italian dubbing is compared to the past; rather, this study focuses on investigating the rendering of pragmatic phenomena through a quantitative and qualitative examination of hesitations compared to the source English text. In this regard, the main difference between English and Italian emerges in relation to the genre variable. More specifically, in Italian dubbed film speech, the genre of the film seems to influence the presence of hesitations more than it does in Anglophone dialogue; in other words, there might be a correlation between film genre and the frequency of occurrence of hesitations in dubbing. As discussed in Chapter Five, hesitations tend to be more present in dubbed comedy films (especially earlier dubs) where these elements tend to be used to characterize characters and the situations they find themselves in, emphasizing clichés and stereotypes that, in turn, favour the ironic and humorous effect of the film. Thereby, a difference between the way hesitations are used in English and Italian dialogues can be noticed: as for the English language, hesitation markers are mainly used to fulfil discursive functions (Structural and Planning Domains), reproducing a sense of naturalness and spontaneity regardless of the genre of the film. As for dubbed Italian, while in non-comedy films the phenomenon is mainly used to convey a sense of naturalness and spontaneity to dialogues, in comedies hesitations are used as a dominant stylistic element to exaggerate the meanings stereotypically associated to them — i.e. uncertainty, clumsiness, lies, comical misunderstandings etc. — and boost the comical dimension of the film.

The results discussed above aimed at describing the tendencies noticed in original English and in Italian dubbed film speech in relation to the pragmatic phenomenon of hesitations through a systematic and empirical approach. Findings have allowed to illustrate the way these elements are used in English-language dialogues and how they are treated in

Italian dubs, also outlining a change of focus with respect to these pragmatic features within the Italian dubbing industry. This investigation wanted to overcome some of the limits found in previous research by compiling a comprehensive and representative corpus of film speech within the limits imposed by the specific type of texts studied here, i.e. verbatim transcriptions of spoken dialogue. The methodological decision to opt for the manual analysis of verbatim transcriptions was deliberately made so as to prioritize depth over breadth. Indeed, while Automated Speech Recognition (ASR) tools would have allowed a larger and inevitably more representative dataset, they would not have guaranteed the same level of accuracy as manual transcription. Nevertheless, there are multiple ways in which this study could be further developed. First, the investigation could be extended by increasing the number of titles included in the corpus. This could have relevant implications for statistical tests; as a matter of fact, the sample studied here is relatively limited: 28 titles in total, including 4 films per decade, distributed in two genre categories. Although this dataset is bigger than previous corpora built on transcribed texts manually analysed, the statistical power of the data is quite low; as a result, real differences may not emerge as ‘significant’ even though they actually are. In particular, the near-significance outcome obtained for the genre variable (p-value: 0.0919) in the Italian section of the corpus may reflect the limited significance power of the data due to the small sample size or limited variability within groups. Additional control variables, for instance, could provide more robust results; in the case of Italian dubbing, future studies on a larger sample could help clarify whether genre significantly influences the frequency of hesitations in dubbed film speech. In this respect, the corpus could be expanded to include more recent works, among which also titles produced and distributed via streaming platforms.

In addition, significant findings could be obtained by resorting to archival materials and a genetic approach. Studies conducted on archival materials (Mereu Keating 2012, 2016, 2019a-b; Zanotti 2014a, 2018, 2019, 2025) have underlined how textual resources can provide insightful details on the translation and adaptation process of audiovisual texts. Zanotti (2014a), for instance, resorted to archival documents to shed some light on the role of the different agents involved in the translation process of dubbing scripts and how the final translated product comes into being. Moreover, archival documents can reveal how translation and dubbing practices have evolved through the years (see Mereu Keating 2019a and 2019b on the emergence of dubbing in Italy); Mereu Keating (2016), for instance, used

the *Italia Taglia* database¹⁶⁴ and other archival resources to investigate the evolution of Italian dubbing and, more specifically, the changing norms governing the translation of original English-language movies in Fascist Italy. Archival resources could thus be employed to investigate whether the actors' performances alter the way dialogues were originally scripted with respect to hesitation markers and whether dubbing scripts differ from the original ones in the way hesitation markers are codified in the written text. This would allow to examine the role played by the various agents involved in the dubbing process in relation to how hesitations are treated in Italian adaptations, i.e. whether these pragmatic markers were maintained, omitted, altered or substituted at a pre-production phase or whether the script was changed while recording the dubbed soundtrack. This can be achieved by comparing the manuscript (the pre-planned written text) with the words actually uttered by actors and recorded on the soundtrack. In this regard, scholars have pointed out how improvisation can play a central role (Taylor 2004; Baños Piñero 2024), affecting the frequency of *uh* and *um* in fictional dialogue and the distribution of their functions (Tonetti Tübben & Landert 2022). Moreover, archival materials may also provide reliable information regarding who translated and directed the dubbing of some of the films analysed in the corpus. Details concerning practitioners are not always easy to retrieve, especially with respect to early dubbings; nevertheless, it would be interesting to inspect whether some of the trends noticed in the dubbed section of the study are motivated by stylistic or specific translation approaches adopted by translators, dialogue writers and/or dubbing directors.

Drawing upon Pavesi & Perego (2006) and Sánchez-Mompeán (2017, 2020a), it would be fruitful to survey professionals through questionnaires distributed among dubbing studios and dubbing actors so as to examine whether the agents involved in the adaptation process are aware of these pragmatic markers and the role they fulfil in the dialogue, and whether the way vocalizations are treated in dubbing is the result of a conscious translation strategy or is a performance-based result. Along these lines, drawing upon Schettino *et al.*'s (2022) investigation on how hesitation phenomena are perceived by Italian speakers and their role in guaranteeing the naturalness of prefabricated speech, surveys could also be used to extend this research through audience reception studies, i.e. to examine whether the rendering, or lack of rendering, of these features impacts viewers' perception of the scene, the characters and the relations between them. In this respect, findings could also have useful implications for practitioners, raising awareness on the pragmatic roles of oral features in

¹⁶⁴ The *Italia Taglia* database is available at the following link: <https://www.italiataglia.it/>.

film speech; in particular, this could concretely help counteract the tendency to over-standardize the translated text and enhance naturalness in dubbing. In turn, this investigation may contribute to the training of future AVT professionals in achieving credible and natural-sounding dialogue in dubbing (see Baños Piñero 2021 on this). As pointed out by Spiteri Miggiani (2025), one of the main challenges for trainees who are learning translation and adaptation for dubbing concerns the unnecessary loss of nuances “embedded in the prosodic features of the original dialogue, that may have an impact on the overall reception of the dialogue” (*ibidem*: 87). Since this may also involve hesitation markers, raising awareness on the meanings of these often-overlooked features could help identify the most successful strategies to render them in the TL.

Moreover, drawing on Schettino *et al.*’s (2024) finding that hesitation phenomena increase the naturalness effect of synthesised utterances, this investigation holds potential for advancing research on AI-mediated dubbing and naturalness, as well as for developing AI-driven tools. In the latter case, insights about the pragmatic markers studied here could contribute to improving the overall performance of ASR tools by training them on more accurate labelling methods and annotation formats so as to obtain a more precise representation of spoken language. In this respect, the Functional Framework developed for this study — drawing on previous literature — could represent a valuable tool for creating annotated corpora. Future researchers may resort to this framework for a range of applications: to investigate orality and pragmatic markers in spontaneous conversation and telecinematic language — also focusing on a corpus of other types of audiovisual products, such as TV series, to test whether different productions yield different results — but also in language learning to teach the communicative functions of pragmatic features and develop students’ pragmatic competence (see Wang 2021 on hesitation markers in the speech of English learners at different levels).

In conclusion, future developments could be aimed at replicating the functional analysis carried out for the English section of the corpus on Italian dubbing, i.e. a more accurate quantitative and qualitative analysis of the functions hesitations fulfil in dubbed fictional dialogue. In this respect, it would be interesting to compare the role hesitations fulfil in Italian dubbed film speech with their occurrences in spontaneous Italian, on the one hand, and in original Italian film speech, on the other. Therefore, similarly to the *Pavia Corpus of Film Dialogue* (Freddi & Pavesi 2009; Pavesi 2014; Zago 2019), this corpus could be expanded through a comparable component of original Italian films.

To conclude, this study wanted to contribute to ongoing discussions on naturalness and the representation of orality in scripted speech and its dubbed adaptations. Through the development of a diachronic parallel corpus, it provides a longitudinal perspective on how norms have evolved with respect to the codification and translation of pragmatic markers over time. By acknowledging the value of these often-overlooked features, this research further highlights the complex nature of audiovisual texts and their adaptation process, bringing to the fore the nuanced meanings that emerge through (dubbing) actors' performances.

References

- Aarts B. 2001. Corpus linguistics, Chomsky and fuzzy tree fragments. In Mair C., Hundt M. (eds.), *Corpus Linguistics and Linguistic Theory*. Amsterdam: Rodopi, pp. 5-13.
- Agost R. 1999. *Traducción y doblaje: palabras, voces e imágenes*. Barcelona: Ariel.
- Alvarez-Pereyre M. 2011. Using film as linguistic specimen: Theoretical and practical issues. In Piazza R., Bednarek M., Rossi F. (eds.), *Telecinematic Discourse: Approaches to the Language of Films and Television Series*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 47-67. DOI: <https://doi.org/10.1075/pbns.211.05alv>.
- Anthony L. 2022. What can corpus software do?. In O’Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 103-125.
- Anthony L. 2023. *AntConc (Version 4.2.2)* [Computer Software]. Tokyo: Waseda University. Available online: <https://www.laurenceanthony.net/software.html>.
- Austin J.L. 1962. *How to Do Things with Words: The William James Lectures Delivered at Harvard University in 1955*. Oxford: Oxford University Press.
- Baker M. 1993. Corpus linguistics and translation studies: Implications and applications. In Baker M., Francis G., Tognini-Bonelli E. (eds.), *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 233-250.
- Baker M. 1996. Corpus-based translation studies: The challenges that lie ahead. In Somers H. (ed.), *LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 175-186.
- Baker P. 2006. *Using Corpora in Discourse Analysis*. London/New York: Continuum.
- Baker P., Hardie A., McEnery T. 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Baldick C. 2015. *The Oxford Dictionary of Literary Terms* (4th edition). Oxford: Oxford University Press.
- Baños Piñero R. 2013. ‘That is so cool’: Investigating the translation of adverbial intensifiers in English-Spanish dubbing through a parallel corpus of sitcoms. *Perspectives:*

- Studies in Translation Theory and Practice*, 21(4), pp. 526-542. DOI: <https://doi.org/10.1080/0907676X.2013.831924>.
- Baños Piñero R. 2014. Orality markers in Spanish native and dubbed sitcoms: Pretended spontaneity and prefabricated orality. *Meta*, 59(2), pp. 406-435. DOI: <https://doi.org/10.7202/1027482ar>.
- Baños Piñero R. 2021. Creating credible and natural-sounding dialogs in dubbing: Can it be taught?. *The Interpreter and Translator Trainer*, 15(1), pp. 13-33. DOI: <https://doi.org/10.1080/1750399X.2021.1880262>.
- Baños Piñero R. 2024. *La oralidad prefabricada en la ficción audiovisual original y doblada: Siete Vidas y Friends*. Castelló de la Plana: Universitat Jaume I.
- Baños Piñero R., Bruti S., Zanotti S. 2013. Corpus linguistics and audiovisual translation: In search of an integrated approach. *Perspectives: Studies in Translation Theory and Practice*, 21(4), pp. 483-490.
- Baños Piñero R., Chaume F. 2009. Prefabricated orality: A challenge in audiovisual translation. In Marrano M.G., Nadiani G., Rundle C. (eds.), *inTRAlinea Special Issue: The Translation of Dialects in Multimedia*. Available online at: http://www.intralinea.org/specials/article/Prefabricated_Orality.
- Baños Piñero R., Díaz-Cintas J. 2024. Exploring new forms of audiovisual translation in the age of digital media: Cybersubtitling and cyberdubbing. *The Translator*, 30(1), pp. 129-144. DOI: <https://doi.org/10.1080/13556509.2023.2274119>.
- Bard E.G., Lickley R.J. 1998. Disfluency deafness: Graceful failure in the recognition of running speech. In *Proceedings of the 20th Annual Meeting of the Cognitive Science Society*. Madison: University of Wisconsin-Madison, pp. 108-113.
- Barik H. 1968. On defining juncture pauses: A note on Boomer's "Hesitation and Grammatical Encoding". *Language and Speech*, 11(3), pp. 156-159.
- Baumgarten N. 2005. *The Secret Agent: Film Dubbing and the Influence of the English Language on German Communicative Preferences. Towards a Model for the Analysis of Language use in Visual Media*. PhD Thesis. Hamburg: University of Hamburg.
- Bednarek M. 2010. *The Language of Fictional Television: Drama and Identity*. London/New York: Continuum.
- Bednarek M. 2012. "Get us the hell out of here": Key words and trigrams in fictional television series. *International Journal of Corpus Linguistics*, 17(1), pp. 35-63. DOI: <https://doi.org/10.1075/ijcl.17.1.02bed>.

- Bednarek M. 2015. Corpus-assisted multimodal discourse analysis of television and film narratives. In Baker P., McEnery T. (eds.), *Corpora and Discourse Studies*. London: Palgrave Macmillan, pp. 63-87.
- Bednarek M. 2017. The role of dialogue in fiction. In Locher M.A., Jucker A.H. (eds.), *Pragmatics of Fiction*. Berlin/Boston: De Gruyter Mouton, pp. 129-158.
- Bednarek M. 2018. *Language and Television Series: A Linguistic Approach to TV Dialogue*. Cambridge: Cambridge University Press.
- Bednarek M. 2019. *Creating Dialogue for TV: Screenwriters Talk Television*. London/New York: Routledge.
- Bednarek M. 2020. The Sydney Corpus of Television Dialogue: Designing and building a corpus of dialogue from US TV series. *Corpora (Online)*, 15(1), pp. 107-119. DOI: <https://doi.org/10.3366/cor.2020.0187>.
- Bednarek M. 2023. *Language and Characterisation in Television Series: A Corpus-informed Approach to the Construction of Social Identity in the Media*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Bednarek M., Veirano Pinto M., Werner V. 2021. Corpus approaches to telecinematic language. *International Journal of Corpus Linguistics*, 26(1), pp. 1-9. DOI: <https://doi.org/10.1075/ijcl.00034.int>.
- Bernardini S. 2022. How to use corpora for translation. In O’Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 485-498.
- Betz S., Kosmala L. 2019. Fill the silence! Basic for modelling hesitation. In Rose R.L., Eklund R. (eds.), *Proceedings of DiSS 2019, The 9th Workshop on Disfluency in Spontaneous Speech*. Budapest: ELTE Faculty of Humanities, pp. 11-14.
- Biber D. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber D., Johansson S., Leech G., Conrad S., Finegan E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Blankenship J.D., Kay C. 1964. Hesitation phenomena in English speech: A study in distribution. *WORD*, 20(3), pp. 360-372.
- Bonsignori V. 2009a. Translating English non-standard tags in Italian dubbing. In Giorgio Marrano M., Nadiani G., Rundle C. (eds), *inTRAlinea Special Issue: The Translation of Dialects in Multimedia*. Available online at: <http://www.intralinea.org/specials/article/1709>.

- Bonsignori V. 2009b. Transcribing film dialogue: From orthographic to prosodic transcription. In Freddi M., Pavesi M. (eds.), *Analysing Audio-visual Dialogue. Linguistic and Translational Insights*. Bologna: Clueb, pp. 185-200.
- Bonsignori V. 2015. “Mind you, that’s just, like, a guestimation”: A diachronic analysis of morphological creativity in American teen talk and dubbing. *Lingue e Linguaggi*, 15, pp. 53-68. DOI: <https://doi.org/doi:10.1285/i22390359v15p53>.
- Bonsignori V., Bruti S. 2014. How people greet each other in TV series and dubbing. In Pavesi M., Formentelli M., Ghia E. (eds.), *The Languages of Dubbing: Mainstream Audiovisual Translation in Italy*. Bern: Peter Lang, pp. 89-111.
- Bonsignori V., Bruti S., Masi S. 2012. Exploring greetings and leave-takings in original and dubbed language. In Remael A., Orero P., Carroll M. (eds.), *Audiovisual Translation and Media Accessibility at the Crossroads: Media for All 3*. Leiden: Brill, pp. 357-379.
- Bonsignori V., Filmer D. 2023. Formulaic forms of address as (im)politeness markers in Prime Minister’s Questions: Margaret Thatcher versus Theresa May. *International Journal of English Linguistics*, 13(5), pp. 72-86. DOI: <https://doi.org/10.5539/ijel.v13n5p72>.
- Boomer D.S. 1965. Hesitation and grammatical encoding. *Language and Speech*, 8(3), pp. 148-158.
- Boonsuk Y., Ambele E.A., Buddharat C. 2019. Position of hesitation markers in everyday informal conversation in English. *ABAC Journal*, 39(3), pp. 129-140.
- Bortfeld H., Leon S.D., Bloom J.E., Schober M.F., Brennan S.E. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, 44(2), pp. 123-147. DOI: <https://doi.org/10.1177/00238309010440020101>.
- Bosseaux C. 2015. *Dubbing, Film and Performance: Uncanny Encounters*. Oxford: Peter Lang.
- Bosseaux C. 2018. Investigating dubbing: Learning from the past, looking to the future. In Pérez González L. (ed.), *The Routledge Handbook of Audiovisual Translation*. London/New York: Routledge, pp. 48-63.
- Brennan S.E., Williams M. 1995. The feeling of another’s knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34(3), pp. 383-398. DOI: <https://doi.org/10.1006/jmla.1995.1017>.
- Brock A. 2015. Participation frameworks and participation in televised sitcom, candid camera and stand-up comedy. In Dynel M., Chovanec J. (eds.), *Participation in*

- Public and Social Media Interactions*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 27-47.
- Brown G., Yule G. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
- Brown P., Levinson S.C. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Bruti S. 2009. Translating compliments and insults in the Pavia Corpus of Film Dialogue: Two sides of the same coin?. In Freddi M., Pavesi M. (eds.), *Analysing Audiovisual Dialogue: Linguistic and Translational Insights*. Bologna: Clueb, pp. 143-163.
- Bruti S. 2019. Spoken discourse and conversational interaction. In Pérez-González L. (ed.), *The Routledge Handbook of Audiovisual Translation*. London/New York: Routledge, pp. 192-208.
- Bruti S. 2020. Corpus approaches and audiovisual translation. In Bogucki Ł., Deckert M. (eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan, pp. 381-395. DOI: https://doi.org/10.1007/978-3-030-42105-2_19.
- Bruti S., Pavesi M. 2008. Interjections in translated Italian: Looking for traces of dubbed language. In Prat M., Martelli A., Pulcini V. (eds.), *Investigating English with Corpora: Studies in Honour of Maria Teresa Prat*. Monza: Polimetrica, pp. 207-222.
- Bruti S., Zanotti S. 2017. Exploring the sensory dimension of translated films: An analysis of *The King's Speech*. *Palimpsestes*, 30, pp. 155-178.
- Bruti S., Zanotti S. 2018. Representations of stuttering in subtitling: A view from a corpus of English language films. In Ranzato I., Zanotti S., (eds.), *Linguistic and Cultural Representation in Audiovisual Translation*. London/New York: Routledge, pp. 228-262.
- Bubel C. 2006. *The Linguistic Construction of Character Relations in TV Drama: Doing Friendship in Sex and the City*. Doctoral thesis, Saarland University.
- Bubel C. 2008. Film audiences as overhearers. *Journal of Pragmatics*, 40(1), pp. 55-71.
- Bublitz W. 2017. Oral features in fiction. In Locher M.A., Jucker A.H. (eds.), *Pragmatics of Fiction*. Berlin/Boston: De Gruyter Mouton, pp. 235-264.
- Bucaria C. 2007. Humour and other catastrophes: Dealing with the translation of mixed-genre TV series. *Linguistica Antverpiensia*, 6, pp. 235-254.
- Bucaria C. 2008. Dubbing dark humour: A case study in audiovisual translation. *Lodz Papers in Pragmatics*, 4(2), pp. 215-240. DOI: <https://doi.org/10.2478/v10016-008-0014-2>.

- Calogero L. 2024. Translating hesitations in film speech: The case of Hugh Grant. *Lingue e Linguaggi*, 67, pp. 63-79. DOI: <https://doi.org/doi:10.1285/i22390359v67p63>.
- Cameron D. 2001. *Working with Spoken Discourse*. London: SAGE.
- Candea M., Vasilescu I., Adda-Decker M. 2005. Inter- and intra-language acoustic analysis of autonomous fillers. In *Proceedings of DiSS 2005, Disfluency in Spontaneous Speech Workshop*. Aix-en-Provence, France, pp. 47-52.
- Carter R., McCarthy M. 2006. *Cambridge Grammar of English: A Comprehensive Guide. Spoken and Written English. Grammar and Usage*. Cambridge: Cambridge University Press.
- Cataldo V., Schettino L., Savy R., Poggi I., Origlia A., Ansani A., Sessa I., Chiera A. 2019. Phonetic and functional features of pauses, and concurrent gestures, in tourist guides' speech. In Piccardi D., Ardolino F., Calamai S. (eds.), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale [Audio Archives at the Crossroads of Speech Sciences, Digital Humanities and Digital Heritage]*, Studi AISV 6. Milan: Officinaventuno, pp. 205-231. DOI: <https://doi.org/10.17469/O2106AISV000013>.
- Chafe W.L. 1980. Some reasons for hesitating. In Dechert H.W., Raupach M. (eds.), *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. Berlin/New York: De Gruyter Mouton, pp. 169-180.
- Chaume F. 2004a. Film studies and translation studies: Two disciplines at stake in audiovisual translation. *Meta*, 49(1), pp. 12-24.
- Chaume F. 2004b. *Cine y traducción*. Madrid: Cátedra.
- Chaume F. 2004c. Synchronization in dubbing: A translational approach. In Orero P. (ed.), *Topics in Audiovisual Translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 35-52. DOI: <https://doi.org/10.1075/btl.56.07cha>.
- Chaume F. 2004d. Discourse markers in audiovisual translating. *Meta*, 49(4), pp. 843-855. DOI: <https://doi.org/10.7202/009785ar>.
- Chaume F. 2007. La retraducción de textos audiovisuales: razones y repercusiones traductológicas. In Zaro Vera J.J., Ruíz Noguera F. (eds.), *Retraducir: una nueva mirada. La retraducción de textos literarios y audiovisuales*. Malaga: Miguel Gómez Ediciones, pp. 49-63.
- Chaume F. 2012. *Audiovisual Translation: Dubbing*. Manchester: St. Jerome Publishing.

- Chaume F. 2016. Dubbing a TV drama series. The case of *The West Wing*. in *InTralinea Special Issue: A Text of Many Colours – translating The West Wing*. Available online at: <http://www.intralinea.org/specials/article/2194>.
- Chaume F. 2018a. Is audiovisual translation putting the concept of translation up against the ropes?. *Journal of Specialised Translation*, 30, pp. 84-104. DOI: <https://doi.org/10.26034/cm.jostrans.2018.197>.
- Chaume F. 2018b. The retranslation and mediated translation of audiovisual content in multilingual Spain: Reasons and market trends. *Status Quaestionis*, 15, pp. 12-27. DOI: <https://doi.org/10.13133/2239-1983/14569>.
- Chaume F. 2019. Audiovisual translation in the age of digital transformation: Industrial and social implications. In Ranzato I., Zanotti S. (eds.), *Reassessing Dubbing: Historical Approaches and Current Trends*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 103-124.
- Chaume F. 2020. Dubbing. In Bogucki Ł., Deckert M. (eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan, pp. 103-132.
- Chepinchikj N., Thompson C. 2016. Analysing cinematic discourse using conversation analysis. *Discourse, Context & Media*, 14, pp. 40-53.
- Chiaro D. 2000. The British will use question tags, won't they? The case of *Four Weddings and a Funeral*. In Taylor C. (ed.), *Tradurre il Cinema*, Atti del Convegno del 29-30 novembre 1996. Trieste: Università degli Studi di Trieste, pp. 27-39.
- Chiaro D. 2008. Issues of quality in screen translation. In Chiaro D., Heiss C., Bucaria C. (eds.), *Between Text and Image: Updating Research in Screen Translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 241-253.
- Chiaro D. 2009. Issues in audiovisual translation. In Munday J. (ed.), *The Routledge Companion to Translation Studies*. London/New York: Routledge, pp. 141-165.
- Chiaro D. 2019. To Europe with love: Woody Allen's liquid society. In Ranzato I., Zanotti S. (eds.), *Reassessing Dubbing: Historical Approaches and Current Trends*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 169-187.
- Chmiel A., Mazur I. 2021. A homogenous or heterogeneous audience? Audio description preferences of persons with congenital blindness, non-congenital blindness and low vision. *Perspectives: Studies in Translation Theory and Practice*, 30(3), pp. 552-567.
- Christenfeld N. 1995. Does it hurt to say um?. *Journal of Non-Verbal Behavior*, 19(3), pp. 171-186.

- Clark H.H. 2004. Pragmatics of language performance. In Horn L.R., Ward G. (eds.), *The Handbook of Pragmatics*. Chichester: Wiley-Blackwell, pp. 365-382.
- Clark H.H., Clark E.V. 1977. *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Clark H.H., Fox Tree J.E. 2002. Using uh and um in spontaneous speaking. *Cognition*, 84(1), pp. 73-111. DOI: [https://doi.org/10.1016/S0010-0277\(02\)00017-3](https://doi.org/10.1016/S0010-0277(02)00017-3).
- Clayman S.E. 2002. Sequence and solidarity. In Lawler E.J., Thye S.R. (eds.), *Advances in Group Processes: Group Cohesion, Trust and Solidarity*. Bingley: Emerald Publishing, pp. 229-253. DOI: [https://doi.org/10.1016/S0882-6145\(02\)19009-6](https://doi.org/10.1016/S0882-6145(02)19009-6).
- Clayman S.E. 2013. Turn-constructal units and the transition-relevance place. In Stivers T., Sidnell J. (eds.), *The Handbook of Conversation Analysis*. Chichester: Wiley-Blackwell, pp. 150-166. DOI: <https://doi.org/10.1002/9781118325001.ch8>.
- Coleridge S.T. 1817. *Biographia Literaria; or Biographical Sketches of my Literary Life and Opinions*. London: Rest Fenner. Retrieved online at: <https://www.proquest.com/books/biographia-literaria-biographical-sketches-my/docview/2138581407/se-2>.
- Collard P., Corley M., MacGregor L.J., Donaldson D.I. 2008. Attention orienting effects of hesitations in speech: Evidence from ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), pp. 696-702.
- Comastri F. 2002. Un esperimento nella creazione di un testo elettronico parallelo: Codifica e allineamento di *A Brief History of Time* di Stephen Hawking. in *TRALinea*, 5. Available online at: https://www.intralinea.org/archive/article/Un_esperimento_nella_creazione_di_un_testo_elettronico_parallelo?utm_source=chatgpt.com.
- Corley M., MacGregor L.J., Donaldson D.I. 2007. It's the way you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, 105(3), pp. 658-668.
- Corley M., Stewart O.W. 2008. Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, 2(4), pp. 589-602.
- Cornu J.F. 2019a. The significance of dubbed versions for early sound-film history. In O'Sullivan C., Cornu J.F. (eds.), *The Translation of Films, 1900-1950*. Oxford: Oxford University Press, pp. 191-220.
- Cornu J.F. 2019b. Conclusion: The public face of film translation history. In O'Sullivan C., Cornu J.F. (eds.), *The Translation of Films, 1900-1950*. Oxford: Oxford University Press, pp. 291-296.

- Crochiquia A., Eriksson A., Madureira S. 2024. Dubbing in animated films: Challenges and the impact of voice design. In De Iacovo V., De Paolis B.M., Mereu D. (eds.), *La voce nei media e nelle nuove tecnologie: produzione e percezione [The voice in the media and new technologies: production and perception]*, Studi AISV 12. Milan: Officinaventuno, pp. 25-45. DOI: <https://doi.org/10.17469/O2112AISV000003>.
- Culpeper J. 2001. *Language and Characterisation: People in Plays and Other Texts*. London/New York: Routledge.
- Culpeper J., Kytö M. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
- Cutting J.E. 2023. Changes in conversations across a century of popular cinema. *Art & Perception*, 11(3-4), pp. 319-337. DOI: <https://doi.org/10.1163/22134913-bja10051>.
- Cutting J.E., Candan A. 2015. Shot durations, shot classes, and the increased pace of popular movies. *Projections*, 9, pp. 40-52. DOI: <https://doi.org/10.3167/proj.2015.090204>.
- D'Amico M. 2008. *La commedia all'italiana. Il cinema comico in Italia dal 1945 al 1975*. Milano: Il Saggiatore.
- Davies M. 2019. *The Movie Corpus*. Available online at: <https://www.english-corpora.org/movies/>.
- Davies M. 2021. The TV and Movies corpora: Design, construction, and use. *International Journal of Corpus Linguistics*, 26(1), pp. 10-37.
- De Iacovo V., Colonna V., Romano A. 2020. La pausa. *Bollettino del laboratorio di fonetica sperimentale "Arturo Genre"*, 5, pp. 41-48.
- De Mauro T. 1986. La lingua e i dialetti della commedia. In Napolitano R. (ed.), *Commedia all'italiana. Angolazioni e controcampi*. Roma: Gangemi Editore, pp. 161-177.
- De Mauro T., Mancini F. 1993. *Lessico di frequenza dell'italiano parlato*. Milano: Etas libri.
- Delabastita D. 1993. *There's a Double Tongue: An Investigation into the Translation of Shakespeare's Wordplay, with Special Reference to Hamlet*. Amsterdam: Rodopi.
- Di Giovanni E. 2018. Audio description and reception-centred research. In Di Giovanni E., Gambier Y. (eds.), *Reception Studies and Audiovisual Translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 225-250.
- Di Giovanni E., Romero-Fresco P. 2019. Are we all together across languages? An eye tracking study of original and dubbed films. In Ranzato I., Zanotti S. (eds.), *Reassessing Dubbing: Historical Approaches and Current Trends*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 125-144. DOI: <https://doi.org/10.1075/btl.148.06di>.

- Di Napoli J. 2020. Filled pauses and prolongations in Roman Italian task-oriented dialogue. *Proceedings of the Laughter and Other Non-Verbal Vocalisations Workshop*. Bielefeld, Germany, pp. 24-27.
- Díaz Cintas J. 2009. *New Trends in Audiovisual Translation*. Bristol: Multilingual Matters. DOI: <https://doi.org/10.21832/9781847691552>.
- Díaz Cintas J., Remael A. 2007. *Audiovisual Translation: Subtitling*. Manchester: St. Jerome Publishing.
- Dynel M. 2011a. “You talking to me?” The viewer as a ratified listener to film discourse. *Journal of Pragmatics*, 43(6), pp. 1628-1644. DOI: <https://doi.org/10.1016/j.pragma.2010.11.016>.
- Dynel M. 2011b. Revisiting Goffman’s postulates on participant statuses in verbal interaction. *Language and Linguistics Compass*, 5(7), pp. 454-465. DOI: <https://doi.org/10.1111/j.1749-818X.2011.00286.x>.
- Dynel M. 2014. On the part of ratified participants: Ratified listeners in multi-party interaction. *Brno Studies in English*, 40(1), pp. 27-44.
- Dynel M. 2017. Participation as audience design. In Hoffmann C.R., Bublitz W. (eds.), *Pragmatics of Social Media*. Berlin/Boston: De Gruyter Mouton, pp. 61-82.
- Erard M. 2007. *Um... Slips, Stumbles, and Verbal Blunders, and What They Mean*. New York: Pantheon Books.
- Ferri A.J. 2007. *Willing Suspension of Disbelief: Poetic Faith in Film*. Lanham/Plymouth: Lexington Books.
- Fischer K. 2000. Discourse particles, turn-taking, and the semantics–pragmatics interface. *Revue de Sémantique et Pragmatique*, 8, pp. 111-137.
- Forchini P. 2012. *Movie Language Revisited: Evidence from Multi-Dimensional Analysis and Corpora*. Bern: Peter Lang.
- Forchini P. 2013. A diachronic study of familiarizers (‘man’, ‘guys’, ‘buddy’, ‘dude’) in movie language. *Perspectives: Studies in Translation Theory and Practice*, 21(4), pp. 504-525. DOI: <https://doi.org/10.1080/0907676X.2013.831923>.
- Forchini P. 2021. *The American Movie Corpus: A Tool for the Development of Spoken Lexico Grammatical Competence*. Milano: EDUCatt.
- Forchini P., Palmero Aproso A. 2025. *AMC-Xplorer (web-based interface)*. Online: <https://amc-xplorer.americanmoviecorpus.net>.
- Formentelli M., Ghia E. 2021. ‘What the hell’s going on?’ A diachronic perspective on intensifying expletives in original and dubbed film dialogue. *Textus*, 1, pp. 47-73.

- Fox Tree J.E. 1995. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, 34(6), pp. 709-738. DOI: <https://doi.org/10.1006/jmla.1995.1032>.
- Fox Tree J.E. 2001. Listeners' uses of um and uh in speech comprehension. *Memory and Cognition*, 29(2), pp. 320-326.
- Fox Tree J.E. 2006. Placing like in telling stories. *Discourse Studies*, 8(6), pp. 723-743. DOI: <https://doi.org/10.1177/1461445606069287>.
- Fox Tree J.E. 2007. Folk notions of um and uh, you know, and like. *Text and Talk*, 27(3), pp. 297-314.
- Freddi M. 2009. The phraseology of contemporary film speech: Formulaic language and translation. In Freddi M., Pavesi M. (eds.), *Analysing Audiovisual Dialogue: Linguistic and Translational Insights*. Bologna: Clueb, pp. 101-123.
- Freddi M. 2011. A phraseological approach to film dialogue: Film stylistics revisited. *Yearbook of Phraseology*, 2(1), pp. 137-162. DOI: <https://doi.org/10.1515/9783110236200.137>.
- Freddi M. 2013. Constructing a corpus of translated films: A corpus view of dubbing. *Studies in Translation Theory and Practice*, 21(4), pp. 491-503.
- Freddi M., Pavesi M. 2009. *Analysing Audiovisual Dialogue: Linguistic and Translational Insights*. Bologna: Clueb.
- Galiano L. 2024. Pragmatic markers in English and Italian film dialogue: Distribution and translation. *Pragmatics*, 34(4), pp. 501-533.
- Gallo E., Schettino L., Cutugno F. 2024. About thee hesitant voice: Acoustic and functional analysis of prolongations in Italian spontaneous speech. *Proceedings of Laughter and Other Non-Verbal Vocalisations Workshop, ISCA Archive*. Belfast, United Kingdom, pp. 4-6. DOI: <https://doi.org/10.21437/LW.2024-1>.
- Gennari D.T. 2009. *Post-War Italian Cinema: America Intervention, Vatican Interests*. London/New York: Routledge.
- Ghia E. 2014. "That is the question": Direct interrogatives in English film dialogue and dubbed Italian. In Pavesi M., Formentelli M., Ghia E. (eds.), *The Languages of Dubbing: Mainstream Audiovisual Translation in Italy*. Bern: Peter Lang, pp. 57-88.
- Ghia E. 2019. (Dis)aligning across different linguacultures: Pragmatic questions from original to dubbed film dialogue. *Multilingua*, 38(5), pp. 583-600. DOI: <https://doi.org/10.1515/multi-2018-0120>.

- Giannini A. 2003a. Prolungamenti vocalici e vocalizzazioni. In Così P., Caldognetto Magno E., Zamboni A. (eds.), *Voce, canto, parlato. Studi in onore di Franco Ferrero*. Padova: Unipress, pp. 163-172.
- Giannini A. 2003b. Hesitation phenomena in spontaneous Italian. *Proceeding of the 15th International Congress of Phonetic Sciences*. Barcelona, Spain, pp. 2653-2656.
- Gillings M. 2022. How to use corpus linguistics in forensic linguistics. In O’Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 589-601.
- Gilquin G. 2008. Hesitation markers among EFL learners: Pragmatic deficiency or difference?. In Romero-Trillo J. (ed.), *Pragmatics and Corpus Linguistics: A Mutualistic Entente*. Berlin/New York: De Gruyter Mouton, pp. 119-149.
- Goffman E. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York: The Free Press of Glencoe.
- Goffman E. 1967. *Interaction Ritual: Essays on Face-to-Face Behavior*. New York: Anchor Books.
- Goffman E. 1981. *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Goldman-Eisler F. 1954. On the variability of the speed of talking and on its relation to the length of utterances in conversations. *British Journal of Psychology*, 45(2), pp. 94-107.
- Goldman-Eisler F. 1957. Speech production and language statistics. *Nature*, 180(4600), pp. 1497- 1497.
- Goldman-Eisler F. 1958. Speech analysis and mental processes. *Language and Speech*, 1(1), pp 59-75.
- Goldman-Eisler F. 1961. A comparative study of two hesitation phenomena. *Language and Speech*, 4(1), pp. 18-26.
- Goldman-Eisler F. 1968. *Psycholinguistics: Experiments in Spontaneous Speech*. London: Academic Press.
- Goodwin C. 1987. Forgetfulness as an interactive resource. *Social Psychology Quarterly*, 50(2), pp. 115-131.
- Gray B. 2022. What can a corpus tell us about registers and genres?. In O’Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 235-249.

- Greco G.M. 2019. Accessibility studies: Abuses, misuses and the method of poetic design. In Stephanidis C., (ed.), *HCI International 2019 - Late Breaking Papers* (Vol. 11786). Cham: Springer International Publishing, pp. 15-27.
- Gregory M. 1967. Aspects of varieties differentiation. *Journal of Linguistics*, 3(2), pp. 177-198. DOI: <https://doi.org/10.1017/S0022226700016601>.
- Gregory M., Carroll S. 1978. *Language and Situation. Language Varieties and their Social Contexts*. London/New York: Routledge.
- Grice H.P. 1975. Logic and conversation. In Cole P., Morgan J. (eds.), *Syntax and Semantics. Vol. 3, Speech Acts*. New York: Academic Press, pp. 41-58.
- Gries S.T. 2009. *Statistics for Linguistics with R: A Practical Introduction*. Berlin/Boston: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110216042>.
- Halliday M.A.K. 2003. On the “architecture” of human language. In Webster J.J. (ed.), *On Language and Linguistics*. London/New York: Continuum, pp. 1-32.
- Hart J.T. 1965. Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56(4) pp. 208-216. DOI: <https://doi.org/10.1037/h0022263>.
- Hawkins P. 1971. The syntactic location of hesitation pauses. *Language and Speech*, 14(3), pp. 277-288.
- Heiss C., Soffritti M. 2008. Forlì 1 – The Forlì Corpus of Screen Translation: Exploring microstructures. In Chiaro D., Heiss C., Bucaria C. (eds.), *Between Text and Image: Updating Research in Screen Translation Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 37-50.
- Heyd T. 2010. How you guys doin’? Staged orality and emerging plural address in the television series *Friends*. *American Speech*, 85(1), pp. 33-66. DOI: <https://doi.org/10.1215/00031283-2010-002>.
- Hieke A. 1981. A content-processing view of hesitation phenomena. *Language and Speech*, 24(2), pp. 147-160. DOI: <https://doi.org/10.1016/j.pragma.2020.11.004>.
- Hothorn T., Hornik K., Wien W., Zeileis A. 2015. Ctree: Conditional Inference Trees. Available online at: <https://cran.r-project.org/web/packages/partykit/vignettes/ctree.pdf>.
- Hothorn T., Hornik K., Zeileis A. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3), pp. 651-674. DOI: <https://doi.org/10.1198/106186006X133933>.
- Hothorn T., Zeileis A. 2015. Partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, 16(118), pp. 3905-3909. Available online at: <https://jmlr.org/papers/v16/hothorn15a.html>.

- Hutchby I., Wooffitt R. 2008. *Conversation Analysis* (2nd edition). Cambridge: Polity Press.
- Jackson H., Zé Amvela E. 2007. *Words, Meaning and Vocabulary: An Introduction to Modern English Lexicology* (2nd edition). London/New York: Continuum.
- Jakobson R. 1960. Closing statement: Linguistics and poetics. In Sebeok T. (ed.), *Style in Language*. Cambridge: Massachusetts Institute of Technology Press, pp. 350-377.
- Jiang L., Phutane M., Azenkot S. 2023. Beyond audio description: Exploring 360° video accessibility with blind and low vision users through collaborative creation. *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '23)*. Association for Computing Machinery, New York, USA, Article 50, pp. 1-17. DOI: <https://doi.org/10.1145/3597638.3608381>.
- Johnstone B. 2002. *Discourse Analysis*. Chichester: Wiley-Blackwell.
- Jones C. 2022. What are the basics of analysing a corpus?. In McCarthy M., O'Keeffe A. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 126-139.
- Jucker A.H. 2015a. Uh and um as planners in the Corpus of Historical American English. In Taavitsainen I., Kytö M., Claridge C., Smith J. (eds.), *Developments in English: Expanding Electronic Evidence*. Cambridge: Cambridge University Press, pp. 162-177.
- Jucker A.H. 2015b. Pragmatics of fiction: Literary uses of uh and um. *Journal of Pragmatics*, 86, pp. 63-67.
- Jucker A.H. 2021. Features of orality in the language of fiction: A corpus-based investigation. *Language and Literature*, 30(4), pp. 341-360.
- Jucker A.H., Landert D. 2023. The diachrony of im/politeness in American and British movies (1930-2019). *Journal of Pragmatics*, 209, pp. 123-141. DOI: <https://doi.org/10.1016/j.pragma.2023.02.020>.
- Kahane H.R., Pietrangeli A. 1959. *Structural Studies on Spanish Themes*. Urbana: University Illinois Press.
- Kallan R.A. 2009. Review of the book *Willing Suspension of Disbelief: Poetic Faith in Film*. *Journal of Film and Video*, 61(4), pp. 65-66. DOI: <https://doi.org/10.1353/jfv.0.0046>.
- Kjellmer G. 2003. Hesitation. In defence of ER and ERM. *English Studies*, 84(2), pp. 170-198.
- Kosmala L., Morgenstern A. 2018. Should 'uh' and 'um' be categorized as markers of disfluency? The use of fillers in a challenging conversational context. In Degand L.,

- Gilquin G., Meurant L., Simon A.C. (eds.), *Fluency and Disfluency across Languages and Language Varieties. Corpora and Language in Use*. Louvain-la-Neuve: Presses universitaires de Louvain, pp. 67-90.
- Kozloff S. 2000. *Overhearing Film Dialogue*. Berkeley: University of California Press.
- Lapadat J.C., Lindsay A.C. 1999. Transcription in research and practice: From standardization of technique to interpretative positionings. *Qualitative Inquiry*, 5(1), pp. 64-86. DOI: <https://doi.org/10.1177/107780049900500104>.
- Laserna C.M., Yi-Tai S., Pennebaker J.W. 2014. Um ... who like says you know: Filler word use as a function of age, gender, and personality. *Journal of Language and Social Psychology*, 33(3), pp. 328-338.
- Leech G. 2014. *The Pragmatics of Politeness*. New York: Oxford University Press.
- Levelt W.J.M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT Press.
- Levinson S.C. 2013. Action formation and ascription. In Sidnell J., Stivers T. (eds.), *The Handbook of Conversation Analysis*. Chichester: Wiley-Blackwell, pp. 103-130. DOI: <https://doi.org/10.1002/9781118325001.ch6>.
- Levshina N. 2015. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Levshina N. 2021. Conditional inference trees and random forests. In Paquot M., Gries T. (eds.), *A Practical Handbook of Corpus Linguistics*. Cham: Springer International Publishing, pp. 611-643.
- Lickley R.J. 1995. Missing disfluencies. *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, 4, pp. 192-195.
- Lickley R.J. 2015. Fluency and disfluency. In Redford M.A. (ed.), *The Handbook of Speech Production*. Hoboken, NJ: Wiley, pp. 445-474.
- Locher M., Jucker A., Landert D., Messerli T. 2023. *Fiction and Pragmatics*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781009091688>.
- Locher M.A., Jucker A.H. 2021. *The Pragmatics of Fiction: Literature, Stage and Screen Discourse*. Edinburgh: Edinburgh University Press.
- Lounsbury F.G. 1954. Transitional probability, linguistic structure, and systems of habit-family hierarchies. In Osgood C.E., Seheok T.A. (eds.), *Psycholinguistics: A Survey of Theory and Research Problems*. Baltimore: Waverly Press, pp. 93-101.

- Lynch S. 2002. Filler words become regular practice. In *The Pendulum Online* (Elon University). Available online at: <https://newspapers.digitalnc.org/lccn/2015236760/2002-10-31/ed-1/seq-12.pdf>.
- Maclay H., Osgood C.E. 1959. Hesitation phenomena in spontaneous English speech. *WORD*, 15(1), pp. 19-44.
- Maffia M., Pellegrino E. 2016. Le emozioni nel parlato doppiato: uno studio cross-linguistico. In Bardel C., De Meo A. (eds.), *Parler les langues romanes/Parlare le lingue romanze/Hablar las lenguas romances/Falando línguas românicas*. Napoli: Il Torcoliere, pp. 111-126.
- Marzà A., Chaume F. 2009. The language of dubbing: Present facts and future perspectives. In Pavesi M., Freddi M. (eds.), *Analysing Audiovisual Dialogue: Linguistic and Translational Insights*. Bologna: Clueb, pp. 31-39.
- Matamala A. 2018. Voice-over: Practice, research and future prospects. In Pérez-González L. (ed.), *The Routledge Handbook of Audiovisual Translation*. London/New York: Routledge, pp. 64-81.
- Mattiello E. 2013. *Extra-grammatical Morphology in English. Abbreviations, Blends, Reduplicatives, and Related Phenomena*. Berlin/Boston: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110295399>.
- Mautner G. 2022. What can a corpus tell us about discourse?. In O’Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 250-262.
- Mazur I. 2023. Same film, different audio descriptions: Describing foreign films from a functional perspective. *Universal Access in the Information Society*, 23(2), pp. 511-521. DOI: <https://doi.org/10.1007/s10209-023-00998-9>.
- McIntyre D., Lugea J. 2015. The effects of deaf and hard-of-hearing subtitles on the characterisation process: A cognitive stylistic study of *The Wire*. *Perspectives: Studies in Translation Theory and Practice*, 23(1), pp. 62-88.
- Mereu Keating C. 2012. Censorial interferences in the dubbing of foreign films in fascist Italy: 1927-1943. *Meta*, 57(2), pp. 294-309. DOI: <https://doi.org/10.7202/1013946ar>.
- Mereu Keating C. 2016. *The Politics of Dubbing: Film Censorship and State Intervention in the Translation of Foreign Cinema in Fascist Italy*. Oxford: Peter Lang Verlag.
- Mereu Keating C. 2019a. A ‘delirium tremens’: Italian-language film versions and early dubbings by Paramount, MGM, and Fox (1930–3). In O’Sullivan C., Cornu J.F.

- (eds.), *The Translation of Films, 1900-1950*. Oxford: Oxford University Press, pp. 151-176. DOI: <https://doi.org/10.5871/bacad/9780197266434.003.0009>.
- Mereu Keating C. 2019b. Locked into dubbing: Retracing the origins, establishment and fortune of Italy's mainstream AVT practice. In Ranzato I., Zanotti S. (eds.), *Reassessing Dubbing: Historical Approaches and Current Trends*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 64-78. DOI: <https://doi.org/10.1075/btl.148.03kea>.
- Messerli T. 2017. Participation structure in fictional discourse: Authors, scriptwriters, audiences and characters. In Locher M., Jucker A. (ed.), *Pragmatics of Fiction*. Berlin/Boston: De Gruyter Mouton, pp. 25-54. DOI: <https://doi.org/10.1515/9783110431094-002>.
- Messina S. 2017. Quando lo schermo racconta: rapporti linguistici tra cinema e televisione. In Patota G., Rossi F. (eds.), *L'italiano al cinema, l'italiano nel cinema*. Firenze: goWare, Accademia della Crusca.
- Minutella V. 2015. 'It ain't ogre til it's ogre': The dubbing of *Shrek* into Italian. In Díaz Cintas J., Neves J. (eds.), *Audiovisual Translation: Taking Stock*. Newcastle Upon Tyne: Cambridge Scholars Publishing, pp. 140-158.
- Minutella V. 2016. British dialects in animated films: The case of *Gnomeo & Juliet* and its creative Italian dubbing. *Status Quaestionis*, 11, pp. 222-259.
- Minutella V. 2021a. *(Re)Creating Language Identities in Animated Films: Dubbing Linguistic Variation*. Cham: Palgrave Macmillan.
- Minutella V. 2021b. Wow! Ehi, amico. Lascia che ti spieghi... Okay? Già. The English element in dubbed Italian. The case of animated films. *Textus*, 34(1), pp. 75-100.
- Montgomery C. 2017. Double doublage: Vocal performance in the French-dubbed versions of Pixar's *Toy Story* and *Cars*. In Whittaker T., Wright S. (eds.), *Locating the Voice in Film: Critical Approaches and Global Practices*. New York: Oxford University Press, pp. 83-100.
- Morandini L., Morandini L., Morandini M. 2024. *Il Morandini 2024. Dizionario dei film e delle serie televisive*. Bologna: Zanichelli.
- Motta D. 2015. Diachronic models of TV dubbing on contemporary Italian. *ITALICA*, 92(4), pp. 959-978.
- Munday J. 2009. *The Routledge Companion to Translation Studies*. London/New York: Routledge.

- Musacchio M.T. 2004. The distribution of information in LSP translation. A corpus study of Italian. *Rivista internazionale di tecnica della traduzione/International Journal of Translation*, 8. Trieste: EUT Edizioni Università di Trieste, pp. 89-105.
- Myers L. 1973. The art of dubbing. *Filmmakers Newsletter*, 6(6), pp. 56-58.
- Nencioni G. 1983. *Di scritto e di parlato: discorsi linguistici*. Bologna: Zanichelli.
- Neves J. 2009. Interlingual subtitling for the deaf and hard-of-hearing. In Díaz Cintas J., Anderman G. (eds.), *Audiovisual Translation: Language Transfer on Screen*. Basingstoke/New York: Palgrave Macmillan, pp. 151-169.
- Neves J. 2019. Subtitling for deaf and hard of hearing audiences: Moving forward. In Pérez-González L. (ed.), *The Routledge Handbook of Audiovisual Translation Studies*. London/New York: Routledge, pp. 82-95.
- Newmark P. 1988. *A Textbook of Translation*. New York: Prentice Hall.
- Nida E.A. 1964. *Towards a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. Leiden: Brill.
- Norrick N.R. 2014. Interjections. In Rühlemann C., Aijmer K. (eds.), *Corpus Pragmatics: A Handbook*. Cambridge: Cambridge University Press, pp. 249-276.
- O'Connell D.C., Kowal S. 2004. The history of research on the filled pause as evidence of "The Written Language Bias in Linguistics" (Linell, 1982). *Journal of Psycholinguistic Research*, 33(6), pp. 459-474. DOI: <https://doi.org/10.1007/s10936-004-2666-6>.
- O'Connell D.C., Kowal S. 2005. Uh and um revisited: Are they interjections for signaling delay?. *Journal of Psycholinguistic Research*, 34(6), pp. 555-576.
- O'Keeffe A. 2006. *Investigating Media Discourse*. London/New York: Routledge.
- O'Sullivan C., Cornu J.F. 2019a. History of audiovisual translation. In Pérez-González L. (ed.), *The Routledge Handbook of Audiovisual Translation Studies*. London/New York: Routledge, pp. 15-30.
- O'Sullivan C., Cornu J.F. 2019b. *The Translation of Films, 1900-1950*. Oxford: Oxford University Press.
- Origlia A., Savy R., Poggi I., Cutugno F., Alfano I., D'Errico F., Vincze L., Cataldo V. 2018. An audiovisual corpus of guided tours in cultural sites: Data collection protocols in the CHROME project. *Proceedings of the 2018 AVI-CH Workshop on Advanced Visual Interfaces for Cultural Heritage*. Castiglione della Pescaia, Italy, pp. 1-4.
- Paolinelli M., Di Fortunato E. 2005. *Tradurre per il doppiaggio. La trasposizione linguistica dell'audiovisivo: teoria e pratica di un'arte imperfetta*. Milano: Hoepli.

- Pavesi M. 1996. L'allocuzione nel doppiaggio dall'inglese all'italiano. In Heiss C., Bollettieri Bosinelli R.M. (eds.), *Traduzione multimediale per il cinema, la televisione e la scena*. Bologna: Clueb, pp. 117-130.
- Pavesi M. 2005. *La traduzione filmica: aspetti del parlato doppiato dall'inglese all'italiano*. Roma: Carocci editore.
- Pavesi M. 2008. Spoken language in film dubbing: Target language norms, interference and translational routines. In Chiaro D., Heiss C., Bucaria C. (eds.), *Between text and image: Updating research in screen translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 79-99.
- Pavesi M. 2009a. Referring to third persons in dubbing: Is there a role for source language transfer?. In Freddi M., Pavesi M. (eds.), *Analysing Audiovisual Dialogue: Linguistic and Translational Insights*. Bologna: Clueb, pp. 125-141.
- Pavesi M. 2009b. Dubbing English into Italian: A closer look at the translation of spoken language. In Díaz Cintas J. (ed.), *New Trends in Audiovisual Translation*. Bristol: Multilingual Matters, pp. 197-209.
- Pavesi M. 2013. This and that in the language of film dubbing: A corpus-based analysis. *Meta*, 58(1), pp. 103-133. DOI: <https://doi.org/10.7202/1023812ar>.
- Pavesi M. 2014. The Pavia Corpus of Film Dialogue: A means to several ends. In Pavesi M., Formentelli M., Ghia E. (eds.), *The Languages of Dubbing: Mainstream Audiovisual Translation in Italy*. Bern: Peter Lang, pp. 29-55.
- Pavesi M. 2016. Formulaicity in and across film dialogue: Clefts as translational routines. *Across Languages and Culture*, 17(1), pp. 99-121.
- Pavesi M. 2018. Reappraising verbal language in audiovisual translation: From description to application. *Journal of Audiovisual Translation*, 1(1), pp. 101-121. DOI: <https://doi.org/10.47476/jat.v1i1.47>.
- Pavesi M. 2019a. Dubbing. In Baker M., Saldanha G. (eds.), *The Routledge Encyclopedia of Translation Studies* (third edition). London/New York: Routledge, pp. 156-161.
- Pavesi M. 2019b. Corpus-based audiovisual translation studies: Ample room for development. In Pérez-González L. (ed.), *The Routledge Handbook of Audiovisual Translation Studies*. London/New York: Routledge, pp. 315-333.
- Pavesi M. 2020, 'I shouldn't have let this happen'. Demonstratives in film dialogue and film representation. In Hoffmann C., Kirner-Ludwig M. (eds.), *Telecinematic Stylistics*. London: Bloomsbury Academic, pp. 19-38.

- Pavesi M. 2022. Corpora and the language of films: Exploring dialogue in English and Italian. In O’Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (2nd edition). London/New York: Routledge, pp. 547-561.
- Pavesi M., Perego E. 2006. Profiling film translators in Italy: A preliminary analysis. *The Journal of Specialised Translation*, 6, pp. 99-144. DOI: <https://doi.org/10.26034/cm.jostrans.2006.741>.
- Pavesi M., Zanotti S. 2023. Translational routines at the crossroads of corpus studies and historical research: The case of yeah > già in dubbing. In Faraci D., Iamartino G., Lopriore L., Nied M., Zanotti S. (eds.), ‘*When I use a word, it means just what I choose it to mean — neither more nor less*’. *Studies in Honour of Stefania Nuccorini*. Roma: Roma TrE-Press, pp. 165-187.
- Pedersen J. 2008. High felicity: A speech act approach to quality assessment in subtitling. In Chiaro D., Heiss C., Bucaria C. (eds.), *Between Text and Image: Updating Research in Screen Translation Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 101-115.
- Pedersen J. 2011. *Subtitling Norms for Television: An Exploration Focussing on Extralinguistic Cultural References*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Perego E. 2019. Audio description: Evolving recommendations for usable, effective and enjoyable practices. In Pérez-González L. (ed.), *The Routledge Handbook of Audiovisual Translation*. London/New York: Routledge, pp. 114-129.
- Pérez-González L. 2014. *Audiovisual Translation: Theories, Methods and Issues*. London/New York: Routledge.
- Piazza R., Bednarek M., Rossi F. 2011. *Telecinematic Discourse: Approaches to the Language of Films and Television Series*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Plag I., Braun M., Lappe S., Schramm S. 2007. *Introduction to English Linguistics*. Berlin/München/Boston: De Gruyter Mouton.
- Posit team. 2025. *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA. Online: <http://www.posit.co/>.
- Poyatos F. 2002a. *Nonverbal Communication across Disciplines, Volume I: Culture, Sensory Interaction, Speech, Conversation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Poyatos F. 2002b. *Nonverbal Communication across Disciplines, Volume II: Paralanguage, Kinesics, Silence, Personal and Environmental Interaction*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Quaglio P. 2009. *Television Dialogue: The Sitcom Friends vs. Natural Conversation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Ranzato I. 2010. *La traduzione audiovisiva: analisi degli elementi culturospecifici*. Roma: Bulzoni.
- Ranzato I., Zabalbeascoa P. 2022. The portrayal of real-life people in audiovisual translation. *Między Oryginałem a Przekładem*, 28(1), pp. 11-29.
- Ranzato I., Zanotti S. 2019. The dubbing revolution. In Ranzato I., Zanotti S. (eds.), *Reassessing Dubbing: Historical Approaches and Current Trends*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 1-14.
- Rayson P., Leech G., Hodges M. 1997. Social differentiation in the use of English vocabulary: Some analyses of the conversational component of the British National Corpus. *International Journal of Corpus Linguistics*, 2(1), pp. 133-152.
- Richardson K. 2010. *Television Dramatic Dialogue: A Sociolinguistic Study*. New York: Oxford University Press.
- Rietveld T., Hout R. 2005. *Statistics in Language Research: Analysis of Variance*. Berlin/New York: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110877809>.
- Rochester S.R. 1973. The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2(1), pp. 51-81.
- Romero-Fresco P. 2007. Synching and swimming naturally on the side – the translation of hesitation in dubbing. *Linguistica Antverpiensia*, 6, pp. 185-202.
- Romero-Fresco P. 2009a. Naturalness in the Spanish dubbing language: A case of not-so-close *Friends*. *Meta*, 54(1), pp. 49-72.
- Romero-Fresco P. 2009b. The fictional and translational dimensions of the language used in dubbing. In Freddi M., Pavesi M. (eds.), *Analysing Audiovisual Dialogue: Linguistic and Translational Insights*. Bologna: Clueb, pp. 41-47. DOI: <https://doi.org/10.1400/133802>.
- Romero-Fresco P. 2012. Dubbing dialogues... Naturally: A pragmatic approach to the translation of transition markers in dubbing. *MonTI: Monografías de traducción e interpretación*, 4, pp. 181-205. DOI: <https://doi.org/10.6035/MonTI.2012.4.8>.

- Romero-Fresco P. 2013. Accessible filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking. *The Journal of Specialised Translation*, 20, pp. 201-223. DOI: <https://doi.org/10.26034/cm.jostrans.2013.410>.
- Romero-Fresco P. 2015. *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe: UK, Spain, Italy, Poland, Denmark, France and Germany*. Bern: Peter Lang.
- Romero-Fresco P. 2019. *Accessible Filmmaking: Integrating Translation and Accessibility into the Filmmaking Process*. London/New York: Routledge.
- Romero-Fresco P. 2020. The dubbing effect: An eye-tracking study on how viewers make dubbing work. *The Journal of Specialised Translation*, 33(3), pp. 17-40. DOI: <https://doi.org/10.26034/cm.jostrans.2020.547>.
- Rossi F. 1999. *Le parole dello schermo. Analisi linguistica del parlato di sei film dal 1948 al 1957*. Roma: Bulzoni.
- Rossi F. 2006. *Il linguaggio cinematografico*. Roma: Aracne.
- Rossi F. 2011. Discourse analysis of film dialogues: Italian comedy between linguistic realism and pragmatic non-realism. In Piazza R., Bednarek M., Rossi F. (eds.), *Telecinematic Discourse: Approaches to the Language of Films and Television Series*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 21-46. DOI: <https://doi.org/10.1075/pbns.211.04ros>.
- Rühlemann C., Aijmer K. 2014. Corpus pragmatics: Laying the foundations. In Aijmer K., Rühlemann C. (eds.), *Corpus Pragmatics: A Handbook*. Cambridge: Cambridge University Press, pp. 1-26.
- Rühlemann C., Hilpert M. 2017. Colloquialization in journalistic writing: The case of inserts with a focus on well. *Journal of Historical Pragmatics*, 18(1), pp. 104-135.
- Ryan J., Granville S. 2020. The suitability of film for modelling the pragmatics of interaction: Exploring authenticity. *System*, 89, 102186-13. DOI: <https://doi.org/10.1016/j.system.2019.102186>.
- Sacks H., Schegloff E.A., Jefferson G. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), pp. 696-735. DOI: <https://doi.org/10.2307/412243>.
- Salway A. 2007. A corpus-based analysis of audio description. In Díaz Cintas J., Orero P., Remael A. (eds.), *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*. Amsterdam: Rodopi, pp. 151-174.
- Sánchez Mompeán S. 2015. Dubbing animation into Spanish: Behind the voices of animated characters. *The Journal of Specialised Translation*, 23, pp. 270-291.

- Sánchez-Mompeán S. 2017. *The Rendition of English Intonation in Spanish Dubbing*. Unpublished doctoral dissertation. Murcia: Universidad de Murcia.
- Sánchez-Mompeán S. 2019. Dubbing attitudes through tonal patterns: When tones speak louder than words. *VIAL*, 16, pp. 135-155.
- Sánchez-Mompeán S. 2020a. *The Prosody of Dubbed Speech: Beyond the Character's Words*. Cham: Palgrave Macmillan. DOI: <https://doi.org/10.1007/978-3-030-35521-0>.
- Sánchez-Mompeán S. 2020b. Prefabricated orality at tone level: Bringing dubbing intonation into the spotlight. *Perspectives: Studies in Translation Theory and Practice*, 28(2), pp. 284-299. DOI: <https://doi.org/10.1080/0907676X.2019.1616788>.
- Sánchez-Mompeán S. 2021. Netflix likes it dubbed: Taking on the challenge of dubbing into English. *Language & Communication*, 80, pp. 180-190. DOI : <https://doi.org/10.1016/j.langcom.2021.07.001>.
- Sánchez-Mompeán S. 2023a. Engaging English audiences in the dubbing experience: A matter of quality or habituation?. *Íkala*, 28(2), pp. 1-18. DOI: <https://doi.org/10.17533/UDEA.IKALA.V28N2A13>.
- Sánchez-Mompeán S. 2023b. The melody of Spanish dubbed dialogue. *Translation Spaces*, 12(2), pp. 326-347.
- Sánchez-Mompeán S., Zanotti S. 2025. Riding the streaming wave: Broadening global access through audiovisual translation and media accessibility. *Translation & Interpreting*, 17(2), pp. 1-15. DOI: <https://doi.org/10.12807/>.
- Savino M., Refice M. 2000. Acoustic cues for classifying communicative intentions in dialogue systems. In Sojka P., Kopeček I., Pala K. (eds.), *Text, Speech and Dialogue. TSD 2000. Lecture Notes in Computer Science, 1902*. Berlin: Springer International Publishing, pp. 421-426.
- Schegloff E.A. 2007. *Sequence Organization in Interaction: A Primer Conversation Analysis*. Cambridge: Cambridge University Press.
- Schegloff E.A. 2010. Some other “uh(m)”s. *Discourse Processes*, 47(2), pp. 130-174.
- Schegloff E.A., Sacks H. 1973. Opening up closings. *Semiotica*, 8, pp. 289-327. DOI: <http://dx.doi.org/10.1515/semi.1973.8.4.289>.
- Schettino L., Betz S., Cutugno F., Wagner P. 2021a. Pitch and functional characterization of hesitation phenomena in Italian discourse. *Proceedings of the 4th Phonetics and Phonology in Europe*. Barcelona, Spain, pp. 281-283.
- Schettino L., Betz S., Cutugno F., Wagner P. 2021b. Hesitations and individual variability in Italian tourist guides' speech. In Bernardasci C., Dipino D., Garassino D.,

- Negrinelli S., Pellegrino E., Schmid S. (eds.), *Speaker Individuality in Phonetics and Speech Sciences: Speech Technology and Forensic Applications*, *Studi AISV* 8. Milano: Officinaventuno, 8, pp. 243-262.
- Schettino L., Betz S., Wagner P. 2021c. Hesitations distribution in Italian discourse. *Proceedings of DiSS2021, The 10th Workshop on Disfluency in Spontaneous Speech*. Paris, France, pp. 29-34.
- Schettino L., Eklund R. 2023. Prolongation in Italian. *Proceedings of the Disfluency in Spontaneous Speech (DiSS) Workshop 2023*, pp. 81-85. DOI: <https://doi.org/10.21437/DiSS.2023-17>.
- Schettino L., Origlia A., Cutugno F. 2024. Though this be hesitant, yet there is method in 't: Effects of disfluency patterns in neural speech synthesis for cultural heritage presentations. *Computer Speech and Language*, 85(3), 101585. DOI: <https://doi.org/10.1016/j.csl.2023.101585>.
- Schettino L., Origlia A., Matrone G. 2022. Modeling hesitations. Speech synthesis applications and evaluation. In Orrico R., Schettino L. (eds.), *The position of the speaker in interaction: attitudes, intentions, and emotions in verbal communication*, *Studi AISV* 9. Milano: Officinaventuno, pp. 191-206. DOI: <https://doi.org/10.17469/O2109AISV000010>.
- Schiffrin D. 1987. *Discourse Markers*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511611841>.
- Searle J.R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Shriberg E., Lickley R. 1993. Intonation of clause-internal filled pauses. *Phonetica*, 50(3), pp. 172-179. DOI: <https://doi.org/10.1159/000261937>.
- Sinclair J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair J., Jones S., Daley R., Krishnamurthy R. 2004. *English Collocation Studies: The OSTI Report*. London/New York: Continuum.
- Sisto A. 2017. The practice of dubbing and the evolution of the soundtrack in Italian cinema: A schizophonic take. In Burke F. (ed.), *A Companion to Italian Cinema*. Chichester: Wiley-Blackwell, pp. 393-407.
- Smith V.L., Clark H.H. 1993. On the course of answering questions. *Journal of Memory and Language*, 32(1), pp. 25-38. DOI: <https://doi.org/10.1006/jmla.1993.1002>.
- Sorlin S. 2015. Breaking the fourth wall: The pragmatic functions of the second person pronoun in House of Cards. In Gardelle L., Sorlin S. (eds.), *The Pragmatics of*

- Personal Pronouns*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 125-146. DOI: <https://doi.org/10.1075/slcs.171.07sor>.
- Spiteri Miggiani G. 2019. *Dialogue Writing for Dubbing: An Insider's Perspective*. Cham: Palgrave Macmillan.
- Spiteri Miggiani G. 2021. Exploring applied strategies for English-language dubbing. *Journal of Audiovisual Translation*, 4(1), pp. 137-156. DOI: <https://doi.org/10.47476/jat.v4i1.2021.166>.
- Spiteri Miggiani G. 2025. Translators and adapters in training: Addressing quality issues in dubbing scripts. In Poveda-Balbuena M.L., Marrahi-Gómez V., Mangiron C. (eds.), *Traducción y lingüística de corpus: Avances en la era digital*. Valencia: Tirant Lo Blanch, pp. 71-98.
- Staley L., Jucker A.H. 2021. “The uh deconstructed pumpkin pie”: The use of uh and um in Los Angeles restaurant server talk. *Journal of Pragmatics*, 172, pp. 21-34.
- Stenström A. 1990. Pauses in monologue and dialogue. In Svartvik J. (ed.), *The London-Lund Corpus of Spoken English: Description and Research*. Lund: Lund University Press, pp. 211-252.
- Stenström A. 2012. Pauses and hesitations. In Andersen G., Aijmer K. (eds.), *Pragmatics of Society*. Berlin/Boston: De Gruyter Mouton, pp. 537-567.
- Stenzl C. 1983. *Simultaneous Interpretation: Groundwork Towards a Comprehensive Model*. Unpublished MA dissertation. London: Birkbeck College, University of London.
- Strauss S. 2002. This, that, and it in spoken American English: A demonstrative system of gradient focus. *Language Sciences*, 24(2), pp. 131-152. DOI: [https://doi.org/10.1016/S0388-0001\(01\)00012-2](https://doi.org/10.1016/S0388-0001(01)00012-2).
- Styler W. 2023. *Using Praat for Linguistic Research* (Document Version 1.9.2). Available online at: <https://wstyler.ucsd.edu/praat/>.
- Tamayo A., Chaume F. 2017. Subtitling for d/deaf and hard-of-hearing children: Current practices and new possibilities to enhance language development. *Brain Sciences*, 7(7), 75. DOI: <https://doi.org/10.3390/brainsci7070075>.
- Tannen D. 1985. Silence: Anything but. In Tannen D., Saville-Troike M. (eds.), *Perspectives on Silence*. Norwood, NJ: Ablex, pp. 93-111.
- Taylor C. 2004. The language of film: Corpora and statistics in the search for authenticity. *Notting Hill* (1998): A case study. *Miscelánea: A Journal of English and American Studies*, 30, pp. 71-86.

- Taylor C. 2008. Predictability in film language: Corpus-assisted research. In Torsello C.T, Ackerley K., Castello E. (eds.), *Corpora for University Language Teachers*. Bern: Peter Lang, pp. 167-181.
- Taylor C. 2015. The language of AD. In Remael A., Reviere N., Vercauteren G., (eds.), *Pictures Painted in Words: ADLAB Audio Description Guidelines*. Trieste: EUT, pp. 48-51.
- Tilley S.A. 2003. Challenging research practices: Turning a critical lens on the work of transcription. *Qualitative Inquiry*, 9(5), pp. 750-773. DOI: <https://doi.org/10.1177/1077800403255296>.
- Tissi B. 2000. Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. *The Interpreters' Newsletter*, 10, pp. 103-127.
- Tognini-Bonelli E. 2001. *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Tognini-Bonelli E. 2010. Theoretical overview of the evolution of corpus linguistics. In O'Keeffe A., McCarthy M. (eds.), *The Routledge Handbook of Corpus Linguistics* (1st edition). London/New York: Routledge, pp. 14-28.
- Tonetti Tübben I., Landert D. 2022. Uh and um as pragmatic markers in dialogues: A contrastive perspective on the functions of planners in fiction and conversation. *Contrastive Pragmatics*, 4(2), pp. 350-381. DOI: <https://doi.org/10.1163/26660393-bja10049>.
- Tottie G. 2011. Uh and um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics*, 16(2), pp. 173-197.
- Tottie G. 2014a. Turn management and the fillers uh and um. In Aijmer K., Rühlemann C. (eds.), *Corpus Pragmatics: A Handbook*. Cambridge: Cambridge University Press, pp. 381-407.
- Tottie G. 2014b. On the use of uh and um in American English. *Functions of Language*, 21(1), pp. 6-29.
- Tottie G. 2015. Uh and um in British and American English: Are they words? Evidence from co-occurrence with pauses. In Dion N., Lapierre A., Torres Cacoullos R. (eds.), *Linguistic Variation: Confronting Fact and Theory*. London/New York: Routledge, pp. 38-54.
- Tottie G. 2017. From pause to word: Uh, um and er in written American English. *English Language and Linguistics*, 23(1), pp. 105-130.
- Tottie G. 2020. Word-search as word-formation? The case of uh and um. In Núñez-Pertejo P., López-Couso M.J., Méndez-Naya B., Pérez-Guerra J. (eds.), *Crossing Linguistic*

- Boundaries: Systemic, Synchronic and Diachronic Variation in English*. London: Bloomsbury Academic, pp. 29-42.
- Toury G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Valdeón R.A. 2008. Inserts in modern script-writing and their translation into Spanish. In Chiaro D., Heiss C., Bucaria C. (eds.), *Between Text and Image: Updating Research in Screen Translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp.117-132.
- Valdeón R.A. 2009. Imitating the conversational mode in audiovisual fiction: Performance phenomena and non-clausal units. In Amador C., Nunes A. (eds.), *The Representation of the Spoken Mode in Fiction*. Lewiston, NY: Edwin Mellen, pp. 197-221.
- Valdeón R.A. 2011. Dysfluencies in simulated English dialogue and their neutralization in dubbed Spanish. *Perspectives: Studies in Translation Theory and Practice*, 19(3), pp. 221-232.
- Valdeón R.A. 2015. The (ab)use of taboo lexis in audiovisual translation: Raising awareness of pragmatic variation in English-Spanish. *Intercultural Pragmatics*, 12(3), pp. 363-385.
- Valdeón R.A. 2020. Swearing and the vulgarization hypothesis in Spanish audiovisual translation. *Journal of Pragmatics*, 155, pp. 261-272.
- Valdeón R.A. 2022. Latest trends in audiovisual translation. *Perspectives: Studies in Translation Theory and Practice*, 30(3), pp. 369-381. DOI: <https://doi.org/10.1080/0907676X.2022.2069226>.
- Valentini C. 2008. Forlì 1 – The Forlì Corpus of Screen Translation: Exploring macrostructures. In Chiaro D., Heiss C., Bucaria C. (eds.), *Between Text and Image: Updating Research in Screen Translation Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 37-50.
- Valentini C. 2013. Phrasal verbs in Italian dubbed dialogues: A multimedia corpus-based study. *Perspectives: Studies in Translation Theory and Practice*, 21(4), pp. 543-562. DOI: <https://doi.org/10.1080/0907676X.2013.831919>.
- Valentini C., Linardi S. 2009. Forlì 1: A multimedia database for AVT research. In Marrano M.G., Nadiani G., Rundle C. (eds.), *inTralinea Special Issue: The Translation of Dialects in Multimedia*. Available online at:

https://www.intralinea.org/specials/article/Forlitt_1_A_multimedia_database_for_AVT_research.

- Veirano Pinto M. 2013. *A linguagem dos filmes norte-americanos ao longo dos anos: Uma abordagem multidimensional* [The Language of North American Movies over the Years: A Multidimensional Study]. Unpublished doctoral dissertation. São Paulo, Brazil: Catholic University of São Paulo.
- Veirano Pinto M. 2014. Dimensions of variation in North American movies. In Berber S., Veirano Pinto M. (eds.), *Multi-Dimensional Analysis, 25 Years on: A tribute to Douglas Biber*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 109-147.
- Vignozzi G. 2021. Disney animated films vis-à-vis their live-action remakes: A study on the representation and the Italian dubbing of vocative forms. In Sileo A. (ed.), *Audiovisual Translation as Trans-Creation A Collection of Essays*. Roma: UniversItalia, pp. 67-101.
- Villar G., Castillo P. 2017. The presence of ‘um’ as a marker of truthfulness in the speech of TV personalities. *Psychiatry, Psychology and Law*, 24(4), pp. 549-560.
- Vinay J.P., Darbelnet J. 1995. *Comparative stylistics of French and English: A Methodology for Translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Voghera M. 2017. *Dal parlato alla grammatica: costruzione e forma dei testi spontanei*. Roma: Carocci.
- Warren B. 2005. A model of idiomaticity. *Nordic Journal of English Studies*, 4(1), pp. 35-54. DOI: <https://doi.org/10.35360/njes.172>.
- Watts J. 2003. *Politeness*. Cambridge: Cambridge University Press.
- Werner V. 2021. A diachronic perspective on telecinematic language. *International Journal of Corpus Linguistics*, 26(1), pp. 38-70.
- Whitman-Linsen C. 1992. *Through the Dubbing Glass: The Synchronization of American Motion Pictures into German, French and Spanish*. Bern: Peter Lang.
- Zabalbeascoa P. 1997. Dubbing and the non-verbal dimension of translation. In Poyatos F. (ed.), *Nonverbal Communication and Translation: New perspectives and Challenges in Literature, Interpretation and the Media*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 327-342.
- Zabalbeascoa P. 2012. Translating dialogues in audiovisual fiction. In Brumme J., Espunya A. (eds.), *The Translation of Fictive Dialogue*. Leiden: Brill, pp. 63-78.

- Zago R. 2016. *From Originals to Remakes: Colloquiality in English Film Dialogue over Time*. Acireale/Roma: Bonanno Editore.
- Zago R. 2019. English films vs Italian films: A comparative analysis via the *Pavia Corpus of Film Dialogue* and the *WordSmith Tools*. In Ranzato I., Zanotti S. (eds.), *Reassessing Dubbing: Historical Approaches and Current Trends*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 229-241.
- Zago R. 2021. Film discourse. In Friginal E., Hardy J.A. (eds.), *The Routledge Handbook of Corpus Approaches to Discourse Analysis*. London/New York: Routledge, pp. 168-182.
- Zamora P., Pavesi M. 2021. El lenguaje malsonante y su traducción: comparación entre los doblajes de filmes anglófonos al italiano y al español. *Quaderns: Revista de Traducció*, 28, pp. 175-195.
- Zanotti S. 2014a. Translation and transcreation in the dubbing process: A genetic approach. *CULTUS*, 7, pp. 107-132.
- Zanotti S. 2014b. “It feels like bits of me are crumbling or something”: General extenders in original and dubbed television dialogue. In Pavesi M., Formentelli M., Ghia E. (eds.), *The Languages of Dubbing: Mainstream Audiovisual Translation in Italy*. Bern: Peter Lang, pp. 113-140.
- Zanotti S. 2014c. Observing translation norms in dubbed audiovisuals: The case of vague language expressions. In Molino A., Zanotti S. (eds.), *Observing Norms, Observing Usage: Lexis in Dictionaries and in the Media*. Bern: Peter Lang, pp. 351-371.
- Zanotti S. 2015. Analysing redubs: Motives, agents, and audience response. In Baños Piñero R., Díaz Cintas J. (eds.), *Audiovisual Translation in a Global Context: Mapping an Ever-changing Landscape*. London: Palgrave Macmillan, pp. 110-139.
- Zanotti S. 2018. Archival resources and uncertainties in film retranslation research. *STATUS QUAESTIONIS*, 15, pp. 52-76. DOI: <https://doi.org/10.13133/2239-1983/14571>.
- Zanotti S. 2019. Investigating the genesis of translated films: A view from the Stanley Kubrick Archive. *PERSPECTIVES*, 27(2), pp. 201-217. DOI: <https://doi.org/10.1080/0907676X.2018.1490784>.
- Zanotti S. 2022. Audiovisual translation. In Malmkjær K. (ed.), *The Cambridge Handbook of Translation*. Cambridge: Cambridge University Press, pp. 440-460.
- Zanotti S. 2025. Film translation practices in the silent film world. *Journal of Audiovisual Translation*, 8(2), pp. 1-21. DOI: <https://doi.org/10.47476/jat.v8i8.2025.372>.

Sitography

- Anthony L. 2025. AntConc. In *Laurence Anthony's Website*. Online: <https://www.laurenceanthony.net/software/antconc/#section> (2.12.2025).
- Boersma P., Weenink D. 2025. *Praat: Doing Phonetics by Computer*. Online: <http://www.praat.org/> (1.12.2025).
- DeCoste F.C. 2002. *The Verdict*: “Who the hell do you think you are?”. In *Picturing Justice, the On-Line Journal of Law and Popular Culture*. Online: https://cap-press.com/sites/pj/verdict_decoste.htm?srsltid=AfmBOorQ-4X1Acx4VgUnAkuwfd1YR42xSvQmvvO5yfmXNATDEl-hElxh (20.09.2025).
- Lumet S. 2009. Sidney Lumet on *The Verdict*. In *American Film Institute YouTube Channel*. Online: <https://www.youtube.com/watch?v=QaPsVydfWvY> (20.09.2025).
- Oxford University Press. (n.d.). Hesitation, n., 1.a, 2.a. In *Oxford English dictionary*. Online: <https://doi.org/10.1093/OED/3545649217> (12.11.2025).
- Pavlik K. 2017. Actors and Movies with the Most Dialogue. In *Kaylin Pavlik: pop culture + data science*. Online: <https://www.kaylinpavlik.com/long-winded-actors-and-movies-with-the-most-dialogue/> (20.11.2025).
- Roxborough S. 2019. Netflix’s Global Reach Sparks Dubbing Revolution: “The Public Demands it”. In *The Hollywood Reporter*, August 13. Online: <https://www.hollywoodreporter.com/tv/tv-news/netflix-s-global-reach-sparks-dubbing-revolution-public-demands-it-1229761/> (28.12.2025).
- Scott S. 2021. *The Apartment*: A Masterpiece, Wilderwise. In *Medium: Read and write stories*. Online: <https://scottsm589.medium.com/the-apartment-a-masterpiece-wilderwise-102fb32f2b61> (30.08.2025).
- Sokoli S. 2024. Dubtitles: The “hidden” subtitles on Netflix. In *Notion*. Online: <https://ssokoli.notion.site/Dubtitles-the-hidden-subtitles-on-Netflix-1f8ba1130d2543bcaa82deee0b133e05> (18/04/2026).

Filmography

- Allen W. (Director). (1977). *Annie Hall*. Rollins-Joffe Productions. DVD.
- Beatty W., Henry B. (Directors). (1978). *Heaven Can Wait*. Paramount Pictures. Available from Amazon Prime Video.
- Benton R. (Director). (1979). *Kramer vs. Kramer*. Stanley Jaffe Productions. Available from Amazon Prime Video.
- Beresford B. (Director). (1989). *Driving Miss Daisy*. The Zanuck Company. DVD.
- Cholodenko L. (Director). (2010). *The Kids Are All Right*. Gilbert Films, Mandalay Vision, Antidote Films. DVD.
- Coppola F.F. (Director). (1967). *The Conversation*. The Directors Company. Available from Amazon Prime Video.
- Crowe C. (Director). (1996). *Jerry Maguire*. TriStar Pictures, Gracie Films, Vinyl Films Available from Amazon Prime Video.
- Demme J. (Director). (1991). *The Silence of the Lambs*. Strong Heart Productions. Available from Amazon Prime Video.
- Field T. (Director). (2001). *In the Bedroom*. Good Machine, Standard Film Company, GreeneStreet Films. DVD.
- Ford J., LeRoy M., Logan J. (Directors). (1955). *Mister Roberts*. Orange Productions. DVD.
- Gilroy T. (Director). (2007). *Michael Clayton*. Castle Rock Entertainment, Samuels Media, Mirage Enterprises, Section Eight Productions. Available from Amazon Prime Video.
- Kramer S. (Director). (1967). *Guess Who's Coming to Dinner*. Columbia Pictures. Available from Amazon Prime Video.
- Lonergan K. (Director). (2016). *Manchester by the Sea*. Pearl Street Films, K Period Media, The Media Farm, The A/Middleton Project, B Story Pictures. DVD.
- Lumet S. (Director). (1957). *Twelve Angry Men*. Orion-Nova Productions. Available from Amazon Prime Video.
- Lumet S. (Director). (1982). *The Verdict*. 20th Century Fox, Zanuck-Brown Company. DVD.
- Mann D. (Director). (1958). *Separate Tables*. Hecht-Hill-Lancaster Productions, Clifton Productions. Available from Amazon Prime Video.
- Mendes S. (Director). (1999). *American Beauty*. DreamWorks Pictures, Jinks/Cohen Company. Available from Netflix.

Minnelli V. (Director). (1950). *Father of the Bride*. Metro-Goldwyn-Mayer. DVD.

Nichols M. (Director). (1967). *The Graduate*. Lawrence Turman Productions. Available from Amazon Prime Video.

Parker A. (Director). (1988). *Mississippi Burning*. Orion Pictures. Available from Amazon Prime Video.

Payne A. (Director). (2004). *Sideways*. Michael London Productions. Available from Disney+.

Phillips T. (Director). (2019). *Joker*. Warner Bros. Pictures, DC Films, Village Roadshow Pictures, Bron Creative, Joint Effort. DVD.

Pollack S. (Director). (1982). *Tootsie*. Mirage Enterprises. Available from Amazon Prime Video.

Reitman J. (Director). (2007). *Juno*. Mandate Pictures, Mr. Mudd. Available from Disney+.

Rossen R. (Director). (1961). *The Hustler*. Rossen Enterprises. Available from Amazon Prime Video.

Russell D.O. (Director). (2012). *Silver Linings Playbook*. The Weinstein Company. Available from Amazon Prime Video.

Wilder B. (Director). (1960). *The Apartment*. The Mirisch Company. Available from Amazon Prime Video.

Zemeckis R. (Director). (1994). *Forrest Gump*. The Tisch Company. DVD.

APPENDIX A

List of Conventions in the Prosodic Transcription (based on Bonsignori 2009b)

=	latching
Hoke to Boolie	identification of the interlocutor
+some*	overlapping speech
((song playing))	comments
((change of scene))	
((whispering))	comments preceding the whole utterance
((laughing)) +some*	comments preceding specific words or parts of an utterance
((offscreen))	characters not visible on-screen
((onscreen))	characters visible on-screen
voice-over/ v.o.	off-screen voice-over
(XXX)	unintelligible speech
CAPITALS	emphasis, stress, loud voice tone, screaming, first person singular pronoun I, proper names, countries, cities, days of the week
Man/Woman	unidentified or unknown character
Lawyer	secondary or unknown characters identified by their jobs
Man 2, Man 3	secondary characters identified through the same label are numbered when they appear in the same scene
,	intonation break
.	brief pause (1 second)
..	medium pause (from 1 to 2 seconds)
...	long pause (3 or more seconds)
-	repeats
--	suspended speech, interrupted speech, self-correction, false start
:	prolonged sound
%	glottal sound
?	interrogative tone
!	exclamatory tone

APPENDIX B

Normalized Frequencies of Hesitations per Decade (occurrences normalized per 1,000 words)

Decade	Number of Words in the English Transcription	Raw Frequencies of Hesitations	Normalized Frequencies
1950s	55,071	280	5.08434566
1960s	48,631	221	4.54442639
1970s	50,770	437	8.60744534
1980s	49,875	267	5.35338346
1990s	46,832	146	3.11752648
2000s	45,274	367	8.10619782
2010s	50,943	283	5.55522839

APPENDIX C

Hesitations in English Originals (occurrences normalized per 1,000 words)

Year	Film	Genre	Number of Words	Hesitation Tokens	Normalized Frequencies
1950	<i>Father of the Bride</i>	Comedy	13,409	100	7.45767768
1955	<i>Mister Roberts</i>	Comedy	14,186	52	3.66558579
1957	<i>Twelve Angry Men</i>	Non-comedy	15,189	94	6.18868918
1958	<i>Separate Tables</i>	Non-comedy	12,287	34	2.76715227
1960	<i>The Apartment</i>	Comedy	14,634	60	4.100041
1961	<i>The Hustler</i>	Non-comedy	10,040	41	4.08366534
1967	<i>Guess Who's Coming to Dinner</i>	Comedy	14,826	89	6.00296776
1967	<i>The Graduate</i>	Non-comedy	9,131	31	3.39502793
1974	<i>The Conversation</i>	Non-comedy	8,835	70	7.92303339
1977	<i>Annie Hall</i>	Comedy	16,210	132	8.14312153
1978	<i>Heaven Can Wait</i>	Comedy	12,731	131	10.2898437
1979	<i>Kramer vs. Kramer</i>	Non-comedy	12,994	104	8.00369401
1982	<i>Tootsie</i>	Comedy	17,387	91	5.23379536
1982	<i>The Verdict</i>	Non-comedy	12,357	129	10.439428
1988	<i>Mississippi Burning</i>	Non-comedy	10,433	27	2.58794211
1989	<i>Driving Miss Daisy</i>	Comedy	9,698	20	2.060228088
1991	<i>The Silence of the Lambs</i>	Non-comedy	9,445	56	5.929063
1994	<i>Forrest Gump</i>	Comedy	12,014	11	0.915598468
1996	<i>Jerry Maguire</i>	Comedy	16,008	60	3.74812594
1999	<i>American Beauty</i>	Non-comedy	9,365	19	2.02883075
2001	<i>In the Bedroom</i>	Non-comedy	9,328	95	10.1843911
2004	<i>Sideways</i>	Comedy	12,600	132	10.4761905
2007	<i>Michael Clayton</i>	Non-comedy	13,060	53	4.05819296
2007	<i>Juno</i>	Comedy	10,286	87	8.45809839
2010	<i>The Kids Are All Right</i>	Comedy	11,035	130	11.7806978
2012	<i>Silver Linings Playbook</i>	Comedy	17,370	30	1.72711572
2016	<i>Manchester by the Sea</i>	Non-comedy	13,992	105	7.50428816
2019	<i>Joker</i>	Non-comedy	8,546	18	2.10624854

APPENDIX D

Normalized Frequencies of Hesitations in Italian Dubbing per Decade
(occurrences normalized per 1,000 words)

Decade	Number of Words in the Italian Transcription	Raw Frequencies of Hesitations	Normalized Frequencies
1950s	47,890	160	3.34098977
1960s	43,404	218	5.02257856
1970s	46,909	260	5.5426464
1980s	46,010	195	4.23820908
1990s	44,709	95	2.12485182
2000s	42,035	207	4.92446771
2010s	46,435	207	4.4578443

APPENDIX E

Hesitations in Italian Dubbing (occurrences normalized per 1,000 words)

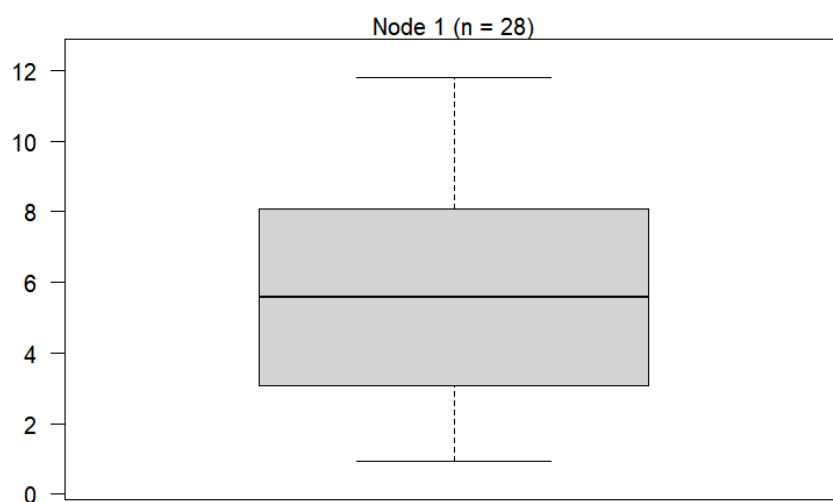
Year	Film	Genre	Number of Words	Hesitation Tokens	Normalized Frequencies
1950	<i>Father of the Bride</i>	Comedy	11,668	57	4.88515598
1955	<i>Mister Roberts</i>	Comedy	12,536	30	2.39310785
1957	<i>Twelve Angry Men</i>	Non-comedy	12,822	44	3.43160193
1958	<i>Separate Tables</i>	Non-comedy	10,864	29	2.66936672
1960	<i>The Apartment</i>	Comedy	13,124	100	7.61962816
1961	<i>The Hustler</i>	Non-comedy	8,545	23	2.69163253
1967	<i>Guess Who's Coming to Dinner</i>	Comedy	13,726	73	5.31837389
1967	<i>The Graduate</i>	Non-comedy	8,009	22	2.74690973
1974	<i>The Conversation</i>	Non-comedy	8,534	18	2.10921022
1977	<i>Annie Hall</i>	Comedy	14,883	73	4.90492508
1978	<i>Heaven Can Wait</i>	Comedy	11,983	127	10.5983477
1979	<i>Kramer vs. Kramer</i>	Non-comedy	11,509	42	3.64931793
1982	<i>Tootsie</i>	Comedy	15,495	84	5.42110358
1982	<i>The Verdict</i>	Non-comedy	11,102	69	5.31435777
1988	<i>Mississippi Burning</i>	Non-comedy	9,862	16	1.62238897
1989	<i>Driving Miss Daisy</i>	Comedy	9,551	26	2.72222804
1991	<i>The Silence of the Lambs</i>	Non-comedy	8,961	22	2.45508314
1994	<i>Forrest Gump</i>	Comedy	11,434	10	0.874584572
1996	<i>Jerry Maguire</i>	Comedy	15,610	50	3.20307495
1999	<i>American Beauty</i>	Non-comedy	8,704	13	1.49356618
2001	<i>In the Bedroom</i>	Non-comedy	9,169	61	6.652852
2004	<i>Sideways</i>	Comedy	11,484	67	5.83420411
2007	<i>Michael Clayton</i>	Non-comedy	12,032	31	2.57646277
2007	<i>Juno</i>	Comedy	9,350	48	5.13368984
2010	<i>The Kids Are All Right</i>	Comedy	10,344	91	8.79737046
2012	<i>Silver Linings Playbook</i>	Comedy	15,880	37	2.32997481
2016	<i>Manchester by the Sea</i>	Non-comedy	12,325	60	4.86815416
2019	<i>Joker</i>	Non-comedy	7,886	19	2.409333

APPENDIX F

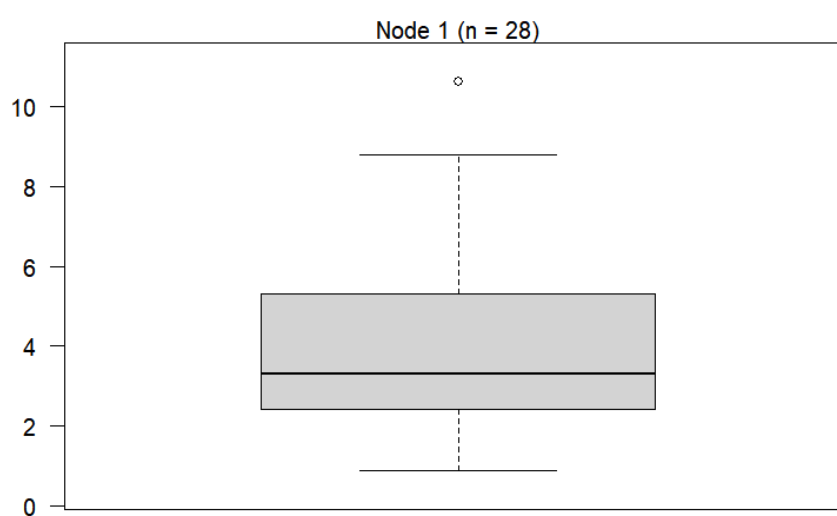
Conditional Inference Tree (CTree)

Retrieved using the code: `ctree (FreqHes ~ Decade + Genre)`

1) Anglophone Film Speech



2) Italian Dubbing



APPENDIX G

Analysis of Variance (ANOVA)

Retrieved using the code: aov (FreqHes ~ Decade + Genre + Decade:Genre)

1) Anglophone Film Speech

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Decade	6	96.35	16.058	1.490	0.268
Genre	2	10.02	5.009	0.465	0.640
Decade:Genre	8	46.90	5.863	0.544	0.802
Residuals	11	118.56	10.778		

2) Italian Dubbing

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Decade	6	31.74	5.290	0.962	0.4924
Genre	2	32.88	16.442	2.990	0.0919.
Decade:Genre	8	19.15	2.394	0.435	0.8763
Residuals	11	60.50	5.500		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1