

# Studies on Chinese Language and Linguistics in Italy

Edited by  
Serena Zuccheri

Studi Interdisciplinari su Traduzione, Lingue e Culture

**Bologna**  
University Press

Studi Interdisciplinari su Traduzione, Lingue e Culture

45

## Studi Interdisciplinari su Traduzione, Lingue e Culture

Collana a cura del Dipartimento di Interpretazione e Traduzione (DIT) dell'Alma Mater Studiorum – Università di Bologna, sede di Forlì.

La Collana, fondata nel 2004, raccoglie le pubblicazioni scientifiche dei suoi afferenti e degli studiosi che operano in ambiti affini a livello nazionale e internazionale.

A partire da una riflessione generale sul tradurre come luogo di incontro e scontro tra lingue e culture, la Collana si propone di diffondere e rendere disponibili, a livello cartaceo e/o su supporto elettronico, i risultati della ricerca in molteplici aree, come la linguistica teorica e applicata, la linguistica dei *corpora*, la terminologia, la traduzione, l'interpretazione, gli studi letterari e di genere, il teatro, gli studi culturali e sull'umorismo.

Le pubblicazioni della Collana sono approvate dal Dipartimento, sentito il motivato parere di almeno due esperti qualificati esterni.

Il/la responsabile della Collana è il/la Direttore/rice del DIT, cui si affianca un comitato scientifico internazionale che varia in relazione alle tematiche trattate.

# Studies on Chinese Language and Linguistics in Italy

Edited by Serena Zuccheri

**Bologna**  
University Press

Progetto Open Access Consorzio Alphabet

Il volume beneficia di un contributo alla pubblicazione da parte del Dipartimento di Interpretazione e Traduzione dell'Alma Mater Studiorum - Università di Bologna

Fondazione Bologna University Press  
Via Saragozza 10 – 40123 Bologna  
tel. (+39) 051 232 882  
fax (+39) 051 221 019

[www.buonline.com](http://www.buonline.com)  
email: [info@buonline.com](mailto:info@buonline.com)

Quest'opera è pubblicata sotto licenza CC-BY-4.0

ISSN: 2283-8910  
ISBN: 979-12-5477-177-8  
ISBN online: 979-12-5477-178-5  
DOI 10.30682/sitec45

Grafica: Alessio Bonizzato  
Impaginazione: Sara Celia

Prima edizione: aprile 2023

# Table of contents

- 7 Preface  
*Serena Zuccheri*
- 9 On Prefixation in Modern Chinese  
*Giorgio Francesco Arcodia*
- 39 Mirative *nǐ kàn* 你看: An Analysis at Syntax/Pragmatics Interface  
*Linda Badan and Yuan Huahung*
- 71 Complex Deadjectival Verbs Based on Open Scale Adjectives in Mandarin Chinese: A Comparison Between *jiā* 加+Adj. and *nòng* 弄+Adj. Verbs  
*Bianca Basciano*
- 95 Chinese Lexicography and the Critical Analysis of Lexicographical Discourse  
*Chiara Bertulesi*
- 117 Discourse Functions of *ránhòu* 然后 and Overtness Requirement for Subjects: A Corpus-Driven Formal Account  
*Marco Casentini and Sergio Conti*
- 147 Investigation on Some Italian Artistic Terms Entered in Chinese: A Diachronic and Synchronic Perspective  
*Feng Lisi*
- 167 Input-Based and Output-Based Instructions: Teaching Activities for Interrogative Constructions in Chinese as a Foreign Language Textbooks for Italian Learners  
*Gloria Gabbianelli*

- 197 Designing and Compiling the Written Sub-Corpus of  
the Bimodal Italian Learner Corpus of Chinese (BILCC):  
Methodological Issues  
*Alessia Iurato*
- 229 Pragmatic Markers and the Right Periphery in Mandarin  
Chinese: A Systematic Review of Types, Functions and  
Co-Occurrence  
*Carmen Lepadat*
- 261 Cantonese Tones and Tone Marks: How Past Studies Can  
Help Present Learners  
*Luisa M. Paternicò*
- 289 A Two-Tiered Analysis of Chinese Political Discourse:  
The Case of Xi Jinping's Commemorative Speech for the  
Centennial of the CCP  
*Carlotta Sparvoli and Chiara Romagnoli*
- 325 The Syntax of *de* (的)-Omission in Post-Numeral Positions  
*Sun Yangyu*
- 347 The Iconic Cognitive Principles of Mandarin Chinese Word  
Order: Pedagogical and Learning Perspectives  
*Tommaso Tucci*
- 373 The Identification and Communication of Expressions of  
Anger in Italian and Chinese Using Emotional Script  
*Valeria Varriano and Serena Zuccheri*
- 405 Contributors

# DISCOURSE FUNCTIONS OF *RÁNHÒU* 然后 AND OVERTNESS REQUIREMENT FOR SUBJECTS: A CORPUS-DRIVEN FORMAL ACCOUNT\*

Marco Casentini

*Ca' Foscari University of Venice*

Sergio Conti

*Roma Tre University*

## 1. Introduction

Discourse markers (DMs) can be described as “sequentially dependent elements which bracket units of talk” (Schiffrin 1987: 31) and signal the relationship between two succeeding discourse segments (Fraser 1999). From a syntactic point of view, DMs are detachable from the sentence and usually appear in sentence initial position (Schiffrin 1987), without affecting its propositional meaning (Sankoff *et al.* 1997).

One of the main properties of DMs is their multifunctionality, and the range of functions that they can perform depends on the communicative context in which they occur (see, among others, Bazzanella 2006;

---

\* This paper stems from a close collaboration between the authors. For the concerns of Italian Academy, Marco Casentini is responsible for Sections 2.2, 2.3, 3.1, 5, and 6, whereas Sergio Conti is responsible for Sections 1, 2.1, 3.2, 3.3, and 4.



Fischer 2006). In addition, DMs also play an important interactional role, since they can be used by the Speaker to guide the Addressee toward a specific interpretation of the following proposition, as well as to convey the Speaker's attitude toward the Addressee and/or the content of the discourse (Fischer 2006).

In Chinese, DMs became a much-discussed topic especially after the 2000s (Xian and Li 2015). Since then, a large body of contributions has appeared, both addressing the issue in general terms (e.g., Feng 2019; Liu 2011; Xu 2015) as well as focusing on specific DMs (e.g., Dong 2007; Piccinini 2020).

The present study deals with *ránhòu* 然后 'then', one of the most frequent DMs in Chinese (see Liu 2011; Xu 2015). Traditionally described as a conjunction (CJ) expressing temporality (e.g., Lü 1999: 461), the discourse and pragmatic functions of *ránhòu* have been widely pointed out in the literature (e.g., see Su 1998; Wang and Huang 2006). In particular, *ránhòu* has been often described as a marker for Topic shift or resumption (e.g., Wang 2018). Based on these premises, the present study draws on corpus data to provide a syntactic explanation of the correlation between the functions of *ránhòu* and sentential Topics, particularly the co-occurrence of *ránhòu* with null or overt Subjects/Topics.

The paper is organized as follows: in Section 2 we briefly introduce the relevant literature on the main topics of the study, i.e., DM *ránhòu* (2.1), Topic (2.2), and Null Subject (2.3); in Section 3 we state our research goals (3.1) and outline the method adopted for data collection, annotation (3.2), and analysis (3.3); in Section 4 we report the statistical results of the analysis, which are interpreted and commented in Section 5; lastly, concluding remarks and possible suggestions for future research are reported in Section 6.

## 2. Literature review

### 2.1 Discourse functions of *ránhòu*

The non-temporal discourse functions of *ránhòu* have been described in several studies, both addressing Chinese DMs in general (e.g., Liu 2011; Tsai and Chu 2015) and *ránhòu* in particular (e.g., Wang 2016;

Wang and Huang 2006). Among the earliest attempts are the studies by Wang (1996), according to which the core function of *ránhòu* is to mark continuation – particularly the succession between prior and upcoming Topics in discourse – and Su (1998), who identified two main discourse functions for *ránhòu*, i.e., ideational and interactional. According to Su, the first category comprises sequentiality-marking uses (temporality, consequence, and lists following a logical hierarchy), while the functions in the second category include signalling condition or concession, marking Topic succession, and serving as a verbal filler that reflects conceptual planning operations.

More recent contributions differ from the source data as well as the granularity of the description. Wang and Huang (2006) analysed a corpus of tape-recorded radio programs and distinguished seven discourse functions for *ránhòu*. Adding on those identified by Su (1998), the authors proposed two more functions, i.e., resumptive opener and additive use. The former consists in resuming an old, digressed Topic, whereas the latter links together a series of related events or actions encompassed within a larger discourse frame, that is, ‘piling’ new information onto old and establishing cohesion.

The study by Wang (2016) is based on two-hour videotaped natural conversations among four native speakers of Chinese. Wang’s analysis focused on three turn management functions of *ránhòu*, namely extension of the turn-constructive unit (TCU), turn continuation, and new turn initiations. While the last type is self-explanatory, TCU extensions and turn continuation differ in that the former consists of “added constituents that serve to complete the prior TCU in terms of syntax, semantics, or pragmatics” (Wang 2016: 149), whereas the latter occurs after a possible syntactic completion point of the turn.

In a following study, Wang (2018) offered a more fine-grained and comprehensive description of the discourse functions of *ránhòu*. Of the 514 corpus occurrences, 40% fulfilled an additive use, introducing additional information on to the current Topic and contributing to turn organization. The second and third functions in terms of the number of occurrences were consequential use (30%) and Topic-shifting (21%). In the former case, *ránhòu* marks the consequence of a prior proposition and is often accompanied by the adverb *jiù* 就, another consequential marker. In the latter case, Wang (2018) distinguished

between disjunctive and stepwise Topic shift, the second being more gradual – e.g., shifting from storytelling to comments or vice versa. A third type of Topic-shifting is ‘skip-connecting’, which is similar to Wang and Huang’s (2006) ‘resumptive opener’ use. Finally, less frequent functions were temporal use – i.e., conjunction (8%), and trail-off (1%).

## 2.2 Topic and Topics

As shown in Section 2.1, the discourse functions of *ránhòu* are closely related to Topic management. Thus, an analysis of the concept of Topic is in need. Topic is one of the basic categories that makes up the so-called Information Structure (IS), together with Focus and Givenness (Féry and Ishihara 2016). The notion of IS refers to the manner in which the speaker accommodates the information to be sent to the addressee, based on their (assumed) shared knowledge (Chafe 1976), also known as Common Ground (CG; Stalnaker 1974, 2002).

With regards to the position of Topics within the sentence, scholars agree that they usually appear in the left periphery, followed by a clause that predicates something about them, which can be called Comment (cf. Badan 2020; Li and Thompson 1976; Shi 2000). Consequently, according to the IS terminology, this type of structure is called Topic-Comment Structure.

In the relevant literature, different definitions can be found for the category of Topic. The present analysis is built on Krifka’s (2008: 265) definition, according to which Topics can be defined as “entity or set of entities under which the information expressed in the Comment constituent should be stored in the CG content”. That is to say, the notion of Topic we assumed is that of sentential Topic.

From a syntactic point of view, more than one Topic can be present in the same clause. However, they play different roles from an IS viewpoint. A first distinction concerns Hanging Topics (HTs) and Left Dislocated Topics (LDTs; for a definition, see Benincà and Poletto 2004). Contrary to the latter, HT is a bare Determinant Phrase (DP) that cannot be preceded by a Prepositional Phrase, and it can be resumed in the relevant sentence either by a tonic pronoun or a full referential DP (for further discussion, see Benincà 2001; Benincà *et al.* 1988; Benincà and Poletto 2004; Lopez 2016; Poletto and Bocci

2016, among others). In addition, according to Benincà and Poletto (2004), there can only be a single HT in each clause, while LDTs can be multiple and, when both an HT and an LDT are present, the former always precedes the latter.

Furthermore, according to Frascarelli and Hinterözl (2007), different types of LDTs exist, namely, Aboutness-Shift Topic (A-Topic), Contrastive Topic (C-Topic), and Given Topic (G-Topic). Specifically, an A-Topic can be associated to Givón's (1983: 8) "newly introduced, newly changed or newly returned to" entity, to which the Comment predicates something about (see also Lambrecht 1994). Thus, only one A-Topic can be realized in each clause.

A C-Topic is an element that creates oppositional pairs with respect to other Topics (Büring 1999; Frascarelli and Hinterözl 2007; Kuno 1976), each of which can be followed by a relevant Comment. Finally, a G-Topic is a Discourse-linked constituent used for Topic continuity, or to resume background information (cf. Givón 1983; Pesetsky 1987). When continuous, it can be null (i.e., silent). Due to their discourse properties, more than one G-Topic can be realized within the same clause.

According to Casentini (2022a) and Frascarelli and Hinterözl (2007), A-, C- and G-Topics are organized in a fixed hierarchy in Italian, German, and Mandarin Chinese. That is to say, when multiple LDTs are realized in the same clause, an A-Topic always precedes a C-Topic, which, in turn, always precedes G-Topics.

Most importantly for the aims of our study, the concept of Topic is strongly connected to the interpretation of Null Subjects (NSs), as observed in Frascarelli (2007). This connection will be further described in Section 2.3.

### 2.3 Null Subjects

The study of NSs is a major issue for linguistic research since the 70's. The original formulation of the NS parameter (Perlmutter 1971) captured the empirical observation that some languages (like Italian and Spanish) could leave definite, referential, pronominal Subject unexpressed, whereas other (like English) could not.

From a typological point of view, at least three types of NS languages can be distinguished, namely, consistent, partial, and radical NS lan-

guages. Specifically, in partial NS languages the possibility to leave the Subject phonologically null depends on specific restrictions, contrary to consistent and radical NS languages, which seem to realize NSs quite freely.

The difference between consistent and radical NS languages lies in the presence/absence of an inflectional verb system. As a matter of fact, the possibility for a given language to license a NS seems to be tightly linked either to the presence of a rich inflectional system or to its complete absence (see, among others, Biberauer *et al.* 2010; Rizzi 1986, 1997; Saito 2007). In these respects, Chinese is considered a radical NS language, since it does not realize any verbal inflection and can leave the Subject, as well as the object, unexpressed.

As for NS interpretation, in recent works (e.g., Casentini 2022a; Frascarelli and Casentini 2019) it is argued that a NS is interpreted as co-referent with the relevant A-Topic which, in a multi-causal domain, can be realized with (silent) low copies (i.e., G-Topics) if continuous. Additionally, according to Casentini (2022a) and Frascarelli and Casentini (2019), what is generally considered to be a Subject can actually be regarded as Topic, at least in unmarked structures. An overt full DP Subject can thus be interpreted as an A-Topic, allowing for the interpretation of the following NSs. These claims are crucial for the analysis conducted in the present paper, since the terminology adopted is built on these recent works. Hence, we will use the term ‘Topic’ to refer to the entity the Comment is about, which can coincide with the overt grammatical Subject. This distinction has also important implicatures from the syntactic point of view.

As a matter of fact, clauses universally consist of three major phrases (from Chomsky 1981 onwards), namely, according to Generative tenets, the Verb Phrase (VP), the Inflectional Phrase (IP), usually referred to as TP in the English literature (from Tense Phrase), and the Complementizer Phrase (CP, also C-Domain). Specifically, the VP is the layer in which theta assignment takes place; the IP/TP layer is responsible for the licensing of formal features such as case and agreement; whereas the CP is the layer where illocutionary Force is encoded and discourse-related categories (such as Topic or Focus) as well as different Operator-like elements (wh-constituents, relative pronouns, quantifiers, etc.; see Rizzi 1997) are hosted. Thus, the term ‘Subject’ refers to the constituent sitting in the specifier (Spec) of TP (in languages with-

out inflection), whereas Topics sit in the Spec of specific phrases in the C-Domain, as it is shown in the following structure in 1:

- 1) [<sub>ShiftP</sub> A-Topic [<sub>ContrP</sub> C-Topic [<sub>FamP</sub> G-Topic [<sub>TP</sub> Subject...]

As stated above, the Topic can coincide with the Subject in unmarked clauses. According to the evidence provided in Casentini (2022a) and Frascarelli and Casentini (2019), the relevant ‘Subject’ is an A-Topic co-referent with a lower silent copy in Spec,TP, as it is shown in example 2 (adapted from Casentini 2022a: 77)<sup>1</sup>:

- 2) 张三很高兴。  
 [<sub>ShiftP</sub> *Zhāngsān* [<sub>TP</sub> <*Zhāngsān*> *hěn* *gāoxìng.*]]  
 Zhangsan very happy  
 ‘Zhangsan is happy.’

### 3. Method

#### 3.1 Research goals and hypothesis

Considered that DM *ránhòu* seems to have different functions related to Topic management, and given the overlapping notions of Subject and Topic, the present analysis aims to investigate whether there is any correlation between the use of *ránhòu* and the following Subject in the utterance. In the light of the existing literature, our hypothesis is that a strong correlation between DM *ránhòu* and overt Subject/Topic exists.

#### 3.2 Sample and data coding

The sample for this study consisted of 7 telephone conversation (roughly 2.5 hours) between Chinese native speakers living in the US. The conversations were randomly selected from the oral corpus *CALL-FRIEND Mandarin Chinese – Mainland Dialect* (Canavan *et al.* 2018), while the extraction of the concordances containing *ránhòu* was conducted using AntConc (ver. 3.5.9 for Macintosh; Anthony 2019), a freeware corpus analysis toolkit for concordancing and text analysis.

<sup>1</sup> Silent constituents are signaled by angle brackets.

The data thus obtained were coded according to different criteria, corresponding to the variables under analysis. First, all the occurrences of *ránhòu* were coded based on their use as a CJ or DM. More specifically, we coded *ránhòu* as CJ when it was clearly used to express temporality, i.e., “to mark an interclausal temporal relationship between adjacent clauses” (Wang and Huang 2006: 995). In example 3, for instance, the events ‘buying a computer’ and ‘start studying it’ are chronologically ordered, as also suggested by the use of *xiān* 先 ‘first’ in the first sentence<sup>2</sup>:

- 3) 你先家里买一个，然后慢慢，慢慢开始学就好了，很容易啊。

*Nǐ xiān jiā-lǐ mǎi yí-gè, ránhòu mànman, mànman*  
 2SG first house-in buy one-CL then slowly slowly  
*kāishǐ xué jiù hǎo le, hěn róngyì a.*  
 start study than good FP very easy FP  
 ‘First you buy one (computer) for your house, then slowly, slowly start studying it and that’s it, it’s easy.’

By contrast, all the non-temporal uses of *ránhòu* were coded as DM. In example 4, the speaker was listing the reasons why they are not satisfied with the school they applied and used *ránhòu* to mark sequentiality:

- 4) 离那学校特别远。然后我想化学一样难嘛。

*Lí nà xuéxiào tèbié yuǎn. Ránhòu wǒ xiǎng*  
 From that school particularly far then 1SG think  
*huàxué yíyàng nán ma.*  
 chemistry equally difficult FP  
 ‘It’s very far from that school. Then I think that chemistry is equally difficult.’

<sup>2</sup> Interlinear glosses generally follow the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>, visited 2023/02/20), adapted to Chinese based on Li and Thompson’s (1981) notations. The abbreviation FP stands for ‘final particle’.

Second, the relevant Subject/Topic following *ránhòu* was coded as either ‘null’ (*tā* 他 ‘he’ in example 5) or ‘overt’ (*tā* 它 ‘it’ in example 6). Cases where it was not possible to determine the Subject – for example, when *ránhòu* occurred in turn-final position – were all grouped under the label ‘other’. This category also includes expletive Subjects (e.g., *měi rén yī kuài dì, ránhòu Ø hěn duō cài a* 每人一块地, 然后很多菜啊 ‘a piece of land per person, then [there were] so many vegetables’), which, being null by definition (Law and Ndayiragije 2017), do not contribute to answering our research questions.

5) 他那就在附近一个湖, 然后就跟别人去。

*Tā nà jiù zài fùjìn yí-gè hú, ránhòu*  
3SG.M that just at nearby one-CL lake then

<*tā*> *jiù gēn biérén qù.*

3SG.M just with other-people go

‘He just (went fishing) by a lake nearby, then (he) just went with other people.’

6) 然后它还给你十块钱。

*Ránhòu tā hái gěi nǐ shí-kuài qián.*

Then 3SG.N also give 2SG ten-CL money

‘Then it [the CALLFRIEND program] also gives you ten dollars.’

Third, the sentence Topic following *ránhòu* was coded as ‘different’ (thus construed as either A- or C-Topics; example 7) or ‘same’ (i.e., thus interpreted as a G-Topic; example 8) in relation to the Topic of the clause immediately preceding *ránhòu*.

7) 所以这一个星期不算, 然后再一个星期不算

*Suǒyǐ [zhè yí-gè xīngqī]<sub>i</sub> bú suàn, ránhòu [zài*

So this one-CL week NEG count then next

*yí-gè xīngqī]<sub>j</sub> bú suàn*

one-CL week NEG count

‘So this week doesn’t count, then again another week doesn’t count’



- 8) 八月中旬我会过去上班，然后我再回来答辩就对了。

*Bāyuè*            *zhōngxún*            [*wǒ*]<sub>i</sub>    *huì*    *guòqu*  
August            middle            1SG    will    over-go

*shàng-bān,*    *ránhòu*            [*wǒ*]<sub>j=i</sub>  
attend-duty    then            1SG

*zài*            *huílai*    *dábiàn*    *jiù*    *duì*    *le*.  
again            return    discuss    just    right    FP

‘In the middle of August, I will go over there to work, then I’ll just come back again to discuss (my thesis) and that’s it.’

In cases of matrix clauses followed by reported/direct speech, as well as hypothetical, temporal, or other types of subordinate clauses, we considered the Topic of the main clause. In example 9, for instance, *ránhòu* is followed by the temporal clause ‘[when] the new boss arrived’, but the Topic of the main sentence is a null constituent, co-referent with *hěn duō rén* 很多人 ‘many people’ in the preceding clause.

- 9) 所以很多人都很高兴，然后新老板来都拍马屁。

*Suǒyǐ* [*hěn duō rén*]<sub>i</sub>    *dōu*    *hěn*    *gāoxìng,*    *ránhòu*  
So    very    many    people    all    very    happy    then

*xīn lǎobǎn lái*    *lái*    <*tāmen*><sub>j=i</sub>    *dōu*    *pāi*    *mǎpì*.  
new boss    come    3PL    all    beat    horse-butt

‘So many people were happy, and then, when the new boss arrived, they were all licking boots.’

The Topic was coded as ‘same’ also in cases in which two or more different terms were used to refer to the same external referent. In example 10, for instance, the Topic of the two turns uttered by different interlocutors consists of the null logophoric pronouns *nǐ* 你 ‘you’ and *wǒ* 我 ‘I’, respectively; however, these Topics both refer to the same person (speaker B). Lastly, like in the case of Subject coding, sentence Topics that could not be determined were labelled as ‘other’.

10) A: 然后干嘛呢?

**Ránhòu** <nǐ><sub>i</sub> gàn má ne?  
 Then 2SG do what FP  
 ‘Then what do you do?’

B: 然后准备上学。

**Ránhòu** <wǒ><sub>j=i</sub> zhǔnbèi shàng-xué.  
 Then 1SG prepare attend-school  
 ‘Then I intend going to school.’

The last type of coding was related to the different discourse functions of *ránhòu*. The functions we identified based on the related literature as well as a preliminary analysis of the actual data are summarized in Table 1. As shown, the discourse functions of *ránhòu* are related to Topic and turn management, as well as the conceptual or logical organization of ideas (a more detailed description of the discourse functions of *ránhòu* can be found in Conti and Carella 2022).

| Function              | Description                           | Example   |
|-----------------------|---------------------------------------|---|
| Topic shift (TS)      | Marking a shift in the sentence Topic | 所以很大一个厅，然后厨房卧室都啊<br><i>Suǒyǐ hěn dà yí gè tīng, ránhòu chúfáng wòshì dōu a</i><br>‘So a big hall, then the kitchen, the bedroom [are] all ah’   |
| Topic resumption (TR) | Resuming an old, digressed Topic      | 对，我着急走，今天，因为 [...] 然后我着急走<br><i>Duì, wǒ zháojí zǒu, jīntiān, yīnwèi [...] ránhòu wǒ zháojí zǒu</i><br>‘Right, I had to hurry up leaving, today, because [...] then I had to hurry up leaving’ |

|                        |   |  |
|------------------------|---|--|
| Topic development (TD) | Marking successive development of a Topic introduced in the comment of the preceding proposition as a focus   | 是魏的一个朋友，然后他就要住在 Kensington<br><i>Shì Wèi de yí gè péngyou, ránhòu tā zhùzài Kensington</i><br>'He's a friend of Wei's, then he just lives at Kensington'   |
| Topic continuity (TC)  | Maintaining the same sentential Topic   | 就觉着特累哈，然后我，那天就在想 [...]<br><i>Jiù juézhè tè lèi ha, ránhòu wǒ, nà tiān jiù zài xiǎng</i><br>'[I] just felt so tired, ah, then I, that day was just thinking [...]   |
| Turn holding (TH)      | Signalling the speakers' intention to maintain the turn (e.g., while the interlocutor is obviously trying to cut in, at the possible end of a turn construction unit, etc.) | A: 他只有一岁多一些吧<br><i>Tā zhǐ yǒu yí suì duō yìxiē ba</i><br>'He's just one year or so, doesn't he'<br>B: 为啥呢? 嗯<br><i>Wèishá ne? En</i><br>'Why? Uh'<br>A: 然后，吃什么 [...]<br><i>Ránhòu, chī shénme</i><br>'Then, whatever he eats [...] |
| Verbal filler (VF)     | Pause marker for lexical retrieval or local syntactic planning  | 然后呢，那个，那什么，就那天小剛回来下班儿么<br><i>Ránhòu ne, nà gè, nà shénme, jiù nàtiān Xiǎogāng huílái xiàbānr me</i><br>'Then, that, that what, that is, that day Xiaogang came back from work, didn't he'  |

|                            |   |  |
|----------------------------|---|--|
| Trail-off (TO)             | Marking the closure of a turn and inviting the interlocutor's response  | A: 然后明天想去那儿, 去海滩, 然后.....<br><i>Ránhòu míngtiān xiǎng qù nàr, qù hǎitān, ránhòu...</i><br>'Then tomorrow I want to go there, to the sea, then...'<br>B: 去海滩<br><i>Qù hǎitān</i><br>'Go to the beach'                         |
| Conceptual sequencing (CS) | Organizing ideas in discourse time, marking the connection between successive sentences to establish cohesion | 他说离学校又近, 然后旁边就是都是 shopping mall 嘛<br><i>Tā shuō lí xuéxiào yòu jìn, ránhòu pángbiān jiù shì dōu shì shopping mall ma</i><br>'He said that it's close to the school, and then nearby it's all [surrounded by] shopping malls' |
| Consequential use (CU)     | Marking the consequence of a prior proposition  | 大家如果都考得不好, 然后它就要乘多少<br><i>Dàjiā rúguǒ kǎo de bù hǎo, ránhòu tā jiù chéng duōshǎo</i><br>'If everyone doesn't do well at the exam, then it will just multiply accordingly'  |
| Other                      | Unidentifiable functions or functions with just one occurrence  | 中国、印度, 然后墨西哥<br><i>Zhōngguó, Yìndù, ránhòu Mòxīgē</i><br>'China, India, and Mexico' (listing use)  |

Table 1. Discourse functions of *ránhòu*.

### 3.3 Analysis: variables and tests

Following the coding scheme described in Section 3.2, two independent raters coded the data based on both the recordings and the transcriptions of the interactions. Interrater agreement was assessed by means of Cohen's Kappa ( $\kappa$ ), which is traditionally used for nominal variables (Brezina 2018: 90). Values of  $\kappa$  suggest a very good agreement for Subject overttness ( $\kappa = 0.81$ ,  $p = .00 < .01$ ,  $z = 15.6$ ) and DM/

CJ use ( $\kappa = 0.83, p = .00 < .01, z = 11.6$ ), and a substantial agreement for Topic type ( $\kappa = 0.77, p = .00 < .01, z = 14.9$ ) and discourse functions ( $\kappa = 0.76, p = .00 < .01, z = 26.2$ ). A subsequent review of the discrepancies between the two raters found no systematic pattern of disagreement. Considered the nature of judgment variables, our coding scheme was deemed reliable in identifying the relevant variables.

The independent (explanatory) variables for this study are *ránhòu* use (CJ/DM), Topic (same or different), and the DM functions of *ránhòu*; the dependent (outcome) variable is Subject type (null or overt). To conduct significance testing we used the chi-square ( $\chi^2$ ) test for univariate and bivariate analysis, and logistic regression for multivariate analysis. Both these tests are commonly used in corpus linguistics when dealing with nominal variables (Brezina 2018; Stefanowitsch 2020). A chi-square test could not be performed to assess the relationship between DM functions of *ránhòu* and the outcome variable, as more than 20% of expected frequencies were smaller than 5. Following Brezina (2018: 113), we performed a log-likelihood ratio (*G*-test) test instead. All statistical tests were conducted in R (R Core Team, 2020) and its integrated development environment RStudio (RStudio Team 2021). The reported effect size is Cramer's *V* for overall effect and probability ratio (PR) for individual effects (for the interpretation of *V* and PR values, see Brezina 2018: 115-116).

#### 4. Results

The observed and expected frequency of *ránhòu* in our data are reported in Table 2. As shown, *ránhòu* was used as a DM in 157 cases (81.8%), while in the remaining 35 cases (18.2%) it was used as a CJ. The results of the univariate chi-square test demonstrate that this difference is highly significant ( $p < .001$ ):  $\chi^2(1) = 77.52, p = .00001$ .

|       | Frequency |          |          |
|-------|-----------|----------|----------|
| Use   | Observed  | Expected | $\chi^2$ |
| DM    | 157       | 96       | 38.76    |
| CJ    | 35        | 96       | 38.76    |
| Total | 192       |          | 77.52    |

Table 2. Occurrences of *ránhòu*.

For the subsequent analysis, we excluded all the cases in which the variables ‘Topic’ and ‘Subject’ were coded as ‘other’, as they were deemed immaterial for the purposes of this study. These include two instances of *ránhòu* functioning as a TO marking the closure of the turn, as in this case it cannot be followed by any Subject/Topic. The remaining data consist of 142 occurrences of *ránhòu*. The observed frequencies (with percentages) of all the variables are reported in Table 3.

| Variable          | Value     | Frequency | %    |
|-------------------|-----------|-----------|------|
| <i>Ránhòu</i> use | CJ        | 32        | 22.5 |
|                   | DM        | 110       | 77.5 |
| Topic             | Same      | 90        | 63.4 |
|                   | Different | 52        | 36.6 |
| Subject           | Null      | 54        | 38.0 |
|                   | Overt     | 88        | 62.0 |
| DM type           | TS        | 22        | 20.0 |
|                   | TR        | 4         | 3.6  |
|                   | TD        | 2         | 1.8  |
|                   | TC        | 14        | 12.7 |
|                   | TH        | 15        | 13.6 |
|                   | VF        | 3         | 2.7  |
|                   | CS        | 41        | 37.3 |
|                   | CU        | 4         | 3.6  |
|                   | Other     | 5         | 4.5  |

Table 3. Overall observed frequencies.

The results of the chi-square test assessing the association between *ránhòu* use (DM/CJ) and Subject type are highly significant ( $p < .001$ ), with a medium effect size:  $\chi^2(1) = 32.75$ ,  $p = .0001$ ,  $V = 0.48$ , 95% CI [0.316, 0.645]. The data in the cross-tabulation table (Table 4) show a clear preference for overt Subjects following DM *ránhòu* and null Subjects following CJ *ránhòu*. More precisely, overt Subjects are 3.98 times more likely to occur after DM *ránhòu* and null Subjects are 3.19 times more likely to occur after CJ *ránhòu*.

|                   |             | Subject           |                   |       |
|-------------------|-------------|-------------------|-------------------|-------|
| <i>Ránhòu</i> use | Frequency   | Overt             | Null              | Total |
| CJ                | Observed    | 6                 | 26                | 32    |
|                   | Expected    | 19.83             | 12.17             |       |
|                   | $\chi^2$    | 9.65              | 15.72             |       |
|                   | PR [95% CI] | 0.25 [0.12, 0.52] | 3.19 [2.23, 4.58] |       |
| DM                | Observed    | 82                | 28                | 110   |
|                   | Expected    | 68.17             | 41.83             |       |
|                   | $\chi^2$    | 2.81              | 4.57              |       |
|                   | PR [95% CI] | 3.98 [1.92, 8.25] | 0.31 [0.22, 0.45] |       |
| Total             |             | 88                | 54                | 142   |

Table 4. Cross-tabulation: Subject overt/ness by *ránhòu* use.

The relationship between Topic (same/different) and Subject (null/overt) are also highly significant ( $p < .001$ ) with medium effect size:  $\chi^2(1) = 17.85$ ,  $p = .0001$ ,  $V = 0.36$ , 95% CI [0.19, 0.519]. Overall, as predictable, the Subject tends to be null if the Topic in the two clauses connected by *ránhòu* is the same, while it tends to be overt otherwise. However, note that, in our data, overt Subjects are only 1.73 times more likely to occur with different Topics than same Topics, whereas null Subjects are 3.32 times more preferred with same Topics (see Table 5).

|           |             | Subject           |                   |       |
|-----------|-------------|-------------------|-------------------|-------|
| Topic     | Frequency   | Overt             | Null              | Total |
| Same      | Observed    | 44                | 46                | 90    |
|           | Expected    | 55.78             | 34.23             |       |
|           | $\chi^2$    | 2.49              | 4.05              |       |
|           | PR [95% CI] | 0.58 [0.45, 0.74] | 3.32 [1.70, 6.48] |       |
| Different | Observed    | 44                | 8                 | 52    |
|           | Expected    | 32.23             | 19.75             |       |
|           | $\chi^2$    | 4.30              | 7.01              |       |
|           | PR [95% CI] | 1.73 [1.36, 2.20] | 0.30 [0.15, 0.59] |       |
| Total     |             | 88                | 54                | 142   |

Table 5. Cross-tabulation: Subject overt/ness by Topic type.

The ability of *ránhòu* use (DM/CJ) and Topic (same/different) to predict Subject overt/ness was assessed by means of a logistic regression.

The cross-tabulated data are reported in Table 6. As baseline values for the model, we set *ránhòu* use = CJ and Topic = ‘same’ for the predictors and Subject = ‘null’ for the outcome variable.

|                   |           | Subject |      |       |
|-------------------|-----------|---------|------|-------|
| <i>Ránhòu</i> use | Topic     | Overt   | Null | Total |
| CJ                | Same      | 4       | 26   | 30    |
|                   | Different | 2       | 0    | 2     |
| DM                | Same      | 40      | 20   | 60    |
|                   | Different | 42      | 8    | 50    |
| Total             |           | 88      | 54   | 142   |

Table 6. Cross-tabulation: Subject overtiness, *ránhòu* use, and Topic type.

Overall, the model that includes all the predictors is significant ( $LL = 44.72, p < 0.0001$ ) and has acceptable classification properties ( $C$ -index = 0.78). Table 7 displays individual coefficients in the model. The only significant estimate is *ránhòu* use ( $p < .001$ ), the odds of overt Subjects occurring after DM *ránhòu* being 13 times the odds of it occurring after CJ *ránhòu*. As for interaction of the explanatory variable, the results are non-significant ( $p > .05$ ). This suggests that the use of *ránhòu* as a DM favours the occurrence of overt Subjects independently from the Topic being different from or equal to the Topic in the sentence preceding *ránhòu*.

|                            | Estimate (log odds) | Standard Error | Z value (Wald) | $p$ -value | Estimate (odds) | 95% CI lower | 95% CI upper |
|----------------------------|---------------------|----------------|----------------|------------|-----------------|--------------|--------------|
| (Intercept)                | -1.87               | 0.54           | -3.49          | 0.00       | 0.154           | 0.05         | 3.95000e-01  |
| TopicB_different           | 17.44               | 1029.12        | 0.02           | 0.99       | 37425783.674    | 0.00         | NA           |
| TypeB_DM                   | 2.57                | 0.60           | 4.26           | 0.00       | 13.00           | 4.36         | 4.87440e+01  |
| TopicB_different: TypeB_DM | -16.47              | 1029.12        | -0.02          | 0.99       | 0.00            | NA           | 3.97778e+62  |

Table 7. Overt or null Subject: Logistic regression estimates.



As a post-hoc confirmation, we conducted another chi-square test, this time only maintaining the cases of same Topics in the sentences preceding and following *ránhòu*. The results are highly significant ( $p < .001$ ) with large effect size:  $\chi^2(1) = 22.77$ ,  $p = .0001$ ,  $V = 0.50$ , 95% CI [0.296, 0.71]. The cross-tabulated data reported in Table 8 demonstrate that overt Subjects are 2.09 times more preferred after DM *ránhòu* even if the previous sentential Topic is the same.

|                   |             | Subject           |                    |       |
|-------------------|-------------|-------------------|--------------------|-------|
| <i>Ránhòu</i> use | Frequency   | Overt             | Null               | Total |
| CJ                | Observed    | 4                 | 26                 | 30    |
|                   | Expected    | 14.67             | 15.33              |       |
|                   | $\chi^2$    | 7.76              | 7.42               |       |
|                   | PR [95% CI] | 0.16 [0.06, 0.42] | 6.22 [2.36, 16.37] |       |
| DM                | Observed    | 40                | 20                 | 60    |
|                   | Expected    | 29.33             | 30.67              |       |
|                   | $\chi^2$    | 3.88              | 3.71               |       |
|                   | PR [95% CI] | 2.09 [1.49, 2.95] | 0.48 [0.34, 0.67]  |       |
| Total             |             | 44                | 46                 | 90    |

Table 8. Cross-tabulation: Overt and null Subjects with same Topic.

Lastly, the results of the  $G$ -test assessing the relationship between DM type and the outcome variable are significant:  $G(8) = 24.46$ ,  $p = .002$  ( $< .01$ ). The cross-tabulated data in Table 9 are a further confirmation of the analysis conducted so far, with overt Subjects surpassing null Subjects in all cases and observed frequencies for overt Subjects being smaller than expected frequencies only for TH and CS.

|         |           | Subject |      |       |
|---------|-----------|---------|------|-------|
| DM type | Frequency | Overt   | Null | Total |
| TS      | Observed  | 21      | 1    | 22    |
|         | Expected  | 16.40   | 5.60 |       |
| TR      | Observed  | 4       | 0    | 4     |
|         | Expected  | 2.98    | 1.02 |       |
| TD      | Observed  | 2       | 0    | 2     |
|         | Expected  | 1.49    | 0.51 |       |
| TC      | Observed  | 12      | 2    | 14    |
|         | Expected  | 10.43   | 3.56 |       |
| TH      | Observed  | 10      | 5    | 15    |

|       |          |       |       |     |
|-------|----------|-------|-------|-----|
|       | Expected | 11.18 | 3.81  |     |
| VF    | Observed | 3     | 0     | 3   |
|       | Expected | 2.23  | 0.76  |     |
| CS    | Observed | 22    | 19    | 41  |
|       | Expected | 30.56 | 10.43 |     |
| CU    | Observed | 4     | 0     | 4   |
|       | Expected | 2.98  | 1.02  |     |
| Other | Observed | 4     | 1     | 5   |
|       | Expected | 3.73  | 1.27  |     |
| Total |          | 82    | 28    | 110 |

Table 9. Cross-tabulation: Subject by DM type.

### 5. Discussion

The data this study is based on confirm the relevant literature on CJs and DMs. In recent works (e.g., Fernandez-Salgueiro 2008), evidence is provided for an analysis of sentential coordination in terms of adjunction to VP<sup>3</sup>. For instance, a case like example 11 from English can be assumed to have the structure in Figure 1, in which both the silent objects in the matrix and the coordinate sentences are interpreted as co-referents of their antecedent ‘which book’:

- 11) [Which book]<sub>j</sub> has John bought <which book> and Mary read <which book>?

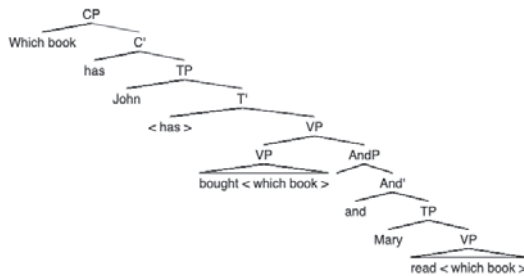


Figure 1. Structure of sentential coordination.

<sup>3</sup> In this work we will use the VP notation, without entering the details of a vP-shell analysis (Larson 1988) since it is immaterial for the purposes of our analysis.

Building on this analysis, it can be argued that cases like (3) above (proposed again in 12, in a simplified version, for convenience) have a similar structure. Specifically, the CJ *ránhòu* is an adjoined phrase with a TP in its Complement (Compl) position. Since only one C-domain is available in the structure, both the matrix and the adjoined NSs select what can be considered the relevant (and only feasible) A-Topic as antecedent. As a matter of fact, the DP *nǐ* is used to introduce a new Topic. Thus, both the NSs are interpreted as *nǐ* ‘you’:

12) 你先买一个，然后慢慢开始学。

*Nǐ*            *xiān*    *mǎi*    *yí-gè*,    *ránhòu* <*nǐ*>    *mànman*  
 2SG            first    buy    one-CL then    2SG    slowly  
*kāishǐ*        *xué*  
 start            study

‘First you buy one (computer), then (you) slowly start studying (it).’

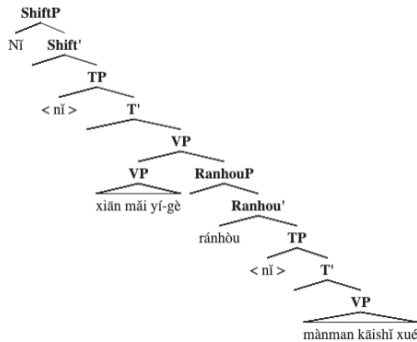


Figure 2. Structure of example 12.

These claims are in line with the literature on both temporal subordination and NS interpretation in Chinese. It is often claimed that the Subject of the *ránhòu*-clause can be left unexpressed when it corresponds to the matrix Subject (Madaro 2016). This can be explained in the light of the analyses conducted by Casentini (2022a) and Frascarelli and Casentini (2019): The NS is interpreted as co-referent of the A-Topic. In other words, both the matrix and the adjoined NS are silent low copies of the only possible A-Topic, since only one CP is available according to the adjunction analysis.

On the contrary, DMs present different structural properties, with respect to CJs. Different scholars (among others, Badan 2020; Casentini 2022b; Osa-Gómez 2012) claim that DMs are pragmatic constituents sitting in the Spec or in the head of specific functional phrases above CP, namely, SpeechActP. Furthermore, the latter can be further split into two layers, that is to say *grounding layer* (GroundP) and *responding layer* (RespP). The former is dedicated to the Speaker's attitude towards the proposition or the discourse content, whereas the latter (structurally higher than the former) is dedicated to what the Speaker wants the Addressee to do with the utterance. Nevertheless, in Lam (2014), Thoma (2016), and Wiltschko (2017), a further division of the GroundP in Ground<sub>Addressee</sub>P and Ground<sub>Speaker</sub>P is assumed. Specifically, as shown in 13, Ground<sub>Speaker</sub>P is dedicated to encoding the Speaker's attitude towards the utterance, while in Ground<sub>Addressee</sub>P it is encoded what the Speaker believes to be the Addressee's attitude toward the proposition:

13) [<sub>RespP</sub> [<sub>GroundAddresseeP</sub> [<sub>GroundSpeakerP</sub> [<sub>CP</sub>

Therefore, DMs introduce independent clauses, with their own C-domain. In the same way, the DM *ránhòu* does not introduce a subordinate clause, contrary to CJ *ránhòu*, thus explaining its functions linked to Topic 'management'. Specifically, it can be argued that when the DM *ránhòu* is used for TS, TR, TD, TC, TH, and TO, the Speaker gives the Addressee instructions to interpret the following utterance (for instance, 'pay attention to the Topic shift', 'let me finish my speech', or 'continue'). Therefore, it can be argued that when DM *ránhòu* signals TS, TR, TD, TC, TH, and TO, it sits in the RespP.

On the other hand, VF, CS, and CU play a role within the GroundP. More specifically, in these cases, the DM *ránhòu* can be argued to be located in Ground<sub>Speaker</sub>P, since it signals Speaker attitude toward the discourse content.

Hence, the second part of example (7) above (proposed again in 14 for convenience) can be analyzed as an independent clause with the DM *ránhòu* in RespP and the relevant Topic in Spec,ContrP, since it is a case of C-Topic<sup>4</sup>:

<sup>4</sup> As a matter of fact, if we consider example (7) above, the relevant Topic in (14)

- 14) 然后再一个星期不算  
*Ránhòu* [zài yí-gè xīngqī] bú suàn  
 then next one-CL week NEG count  
 ‘Then again another week doesn’t count.’



Figure 3. DM *ránhòu* in RespP.

Similarly, a sentence like that in 15 (extracted from the example for CU in Table 1 above) can be argued to have the structure in Figure 4, with the DP *tā* introducing a new Topic:

- 15) 然后它就要乘多少  
*Ránhòu tā jiù chéng duōshao*  
 then 3SG.N just multiply how much  
 ‘Then it will just multiply accordingly’

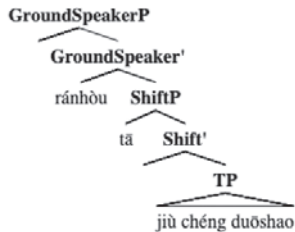


Figure 4. DM *ránhòu* in Ground<sub>Speaker</sub>P.

Finally, as is shown in Figures 3 and 4, the DM *ránhòu* is assumed to be in the head of either RespP or Ground<sub>Speaker</sub>P in order to account for

can be considered to be part of the ‘oppositional pair’ with respect to the (C-) Topic ‘this week’ in the first part of example (7).

the sentence final position of *ránhòu* as TO. Indeed, according to Osa-Gómez (2012), the Spanish DM *no* ‘no’ is derived by the movement of the whole CP to Spec,Ground<sub>Addressee</sub>P. Thus, it can be argued that a sentence like 16 below (extracted from the example for TO in Table 1) has the structure in Figure 5. Specifically, the first DM *ránhòu* is used to signal TC, thus it is located in Ground<sub>Speaker</sub>P, with the relevant silent Topic (*wǒ* ‘I’) in Spec,FamP (since it is a G-Topic). On the other hand, the second DM *ránhòu* with TO function is located in Ground<sub>Speaker</sub>P. Finally, the whole Ground<sub>Speaker</sub>P move to Spec,RespP accounting for its position above the TO *ránhòu*:

- 16) 然后明天想去那儿，去海滩，然后……  
*Ránhòu míngtiān xiǎng qù nàr, qù hǎitān, ránhòu...*  
 then tomorrow would like go there go sea then  
 ‘Then tomorrow I want to go there, to the sea, then...’

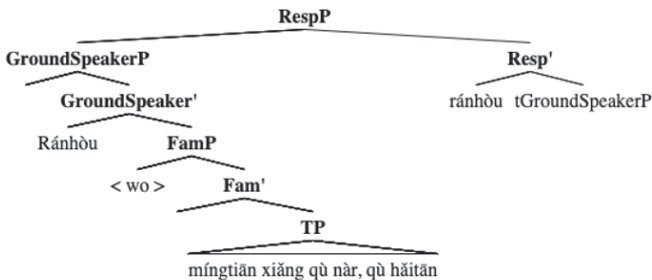


Figure 5. Movement of CP to Spec<sub>speaker</sub>P.

However, the few occurrences of *ránhòu* as TO (2 instances) cannot allow us to conduct an in-depth analysis. Thus, this question is left open for future research.

As a final point, the results of statistical analyses conducted show that the DM and the CJ *ránhòu* crucially differ in terms of overtness of the following Subject/Topic. Specifically, it has been shown that the relevant Subject/Topic is mostly overt when it follows DM *ránhòu*, whereas it is null when it follows CJ *ránhòu*.

In the light of the above discussion, the following distinction can be made. When *ránhòu* is used in its lexical function (i.e., CJ), the null

constituent is a proper grammatical Subject, which in turn can be either overt or null according to the IS. That is to say, when both the Subject in the matrix and in the adjoined clauses are co-referent of the same (and only one) A-Topic in the matrix C-domain, the adjoined Subject is null. On the contrary, *ránhòu* seems to trigger the [+ overt] feature for Topics when it is used as a DM. Thus, the relevant Topic following DM *ránhòu* should be overt, independently of its role in the discourse (i.e., A-, C-, or G-Topic). When realized as an A-Topic, it can serve as antecedent for the following NSs. Most remarkably, our data demonstrated a significant tendency of G-Topics to be realized as overt when following DM *ránhòu*. This in contrast with the usual behaviour of G-Topics, which in other contexts are generally realized as silent (see Section 2.2).

Before concluding, an exception concerning the use of *ránhòu* as a device for CS and TH needs to be accounted for. As seen in Table 9, the data for these two functions apparently depart from the patterns just described. While still presenting a bigger number of overt Subjects/Topics compared to null Subjects/Topics, in both cases the observed frequency for overt Subjects/Topics is smaller than the expected, while the observed frequency for null Subjects/Topics is bigger. Following, we will attempt an explanation for these differences.

In the case of TH, the relatively high occurrence of null Subjects/Topics might depend on discourse planning operations. In fact, as already mentioned, TH *ránhòu* signals the speakers' intention to maintain the turn while the interlocutor is obviously trying to cut in. Thus, it can be regarded as an 'emergency' linguistic device independent from Topic management. As a consequence, the choice (not) to realize an overt Topic might not depend on the TH function itself; rather, it might depend on other pragmatic, functional, or interactional factors that were not considered in the present work and should thus be addressed in future studies.

As for the case of CS *ránhòu*, two different hypotheses can be made. First, in our work we used CS as an umbrella term for a group of sub-functions that have not been considered separately. One of these, for instance, consists in introducing increments, i.e., grammatical extensions of the sentence (see Conti and Carella 2022; Wang 2018). In this case, *ránhòu* might be located in a different node of the syntactic structure, thus not triggering Topic overttness.

Alternatively, these differences can also be explained in the light of the grammaticalization process of *ránhòu* from CJ to DM. Specifically, as argued by Wang and Huang (2006: 995) “the basic temporality of *ránhòu* is exploited for discourse uses, as it grammaticalizes from temporal to non-temporal, from textual linking to conversational cohesion and discourse marker”. Thus, it can be hypothesized that CS *ránhòu* is still undergoing grammaticalization, and that this function still shares properties with the lexical use of *ránhòu* as a CJ, with possible consequences on the presence/absence of the [+overt] triggering trait.

## 6. Conclusions

The results of our analysis show that, in line with previous investigations, *ránhòu* is significantly more frequently used as DM than CJ in spoken Chinese. Like other DMs, one of the main properties of the DM *ránhòu* is multifunctionality. As a matter of fact, the present corpus-driven analysis demonstrates that it can perform at least nine functions, namely, TS, TR, TD, TC, TH, VF, TO, CS, and CU.

Furthermore, evidence is provided for a structural distinction between CJ and DM *ránhòu*. The former is argued to be adjoined to VP, whereas the latter is located in the functional domain above the CP. This structural ‘constraint’ can account for the interpretation of the null Subject following the CJ *ránhòu* as co-referent of the matrix Subject which, according to the present analysis, can be interpreted as the sentential Topic. Conversely, the statistical analysis shows that, when *ránhòu* is used as a DM, the following Topic is mostly overt, even in cases when the Topics in the two clauses connected by *ránhòu* are the same. The only exceptions are cases of TH and CS. This suggests that the grammaticalization process of *ránhòu* from CJ to DM might still be undergoing, at least for what concerns the CS functions, whereas in the case of TH the absence of an overt Topic might be related to other pragmatic or interactional factors.

Finally, it is argued that DM *ránhòu* is located in the head of either RespP or Ground<sub>Speaker</sub>P, accounting for the possibility to be in sentence-final position (i.e., TO function). However, only two instances of *ránhòu* as TO marker were found in our sample. Thus, further investigation should be conducted on this front.



## References

- Anthony, L. (2019) *AntConc (Version 3.5.8)*. Tokyo, Waseda University. <https://www.laurenceanthony.net/software> (visited 2022/04/09).
- Badan, L. (2020) “Italian Discourse Markers: The Case of *Guarda te*”. *Studia Linguistica. A Journal of General Linguistics*, 74(2): 303-336.
- Bazzanella, C. (2006) “Discourse Markers in Italian: Towards a Compositional Meaning”, in K. Fischer (Ed.) *Approaches to Discourse Particles*. Amsterdam, Elsevier: 449-464.
- Benincà, P. (2001) “The Position of Topic and Focus in the Left Periphery”, in G. Cinque and G. Salvi (Eds.) *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*. Amsterdam/London, Elsevier: 39-64.
- Benincà, P. and C. Poletto (2004) “Topic, Focus, and V2. Defining the CP Sublayers”, in L. Rizzi (Ed.) *The Cartography of Syntactic Structures. The Structure of CP and IP*, vol. 2. Oxford, Oxford University Press: 52-75.
- Benincà, P., G. Salvi and L. Frison (1988) “L’ordine degli elementi della frase e le costruzioni marcate”, in L. Renzi (Ed.) *Grande grammatica italiana di consultazione*, vol. I. Bologna, Il Mulino: 115-225.
- Biberauer, T., A. Holmberg, I. Roberts and M. Sheenan (2010) *Parametric Variation: Null Subject in Minimalist Theory*. Cambridge, Cambridge University Press.
- Brezina, V. (2018) *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge, Cambridge University Press.
- Büring, D. (1999) “Topic”, in P. Bosch and R. van der Sand (Eds.) *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge, Cambridge University Press: 142-165.
- Canavan, A., G. Zipperlen and J. Bartlett (2018) *CALLFRIEND Mandarin Chinese-Mainland Dialect Second Edition*. Philadelphia, Linguistic Data Consortium.
- Casentini, M. (2022a) *Topic-Comment Structure and the Interpretation of Null Subjects in Modern Chinese*. PhD Dissertation, Rome, Sapienza University of Rome.
- (2022b) “The Initial Sentence Discourse Marker 那 *na* and its Correlation with Sentence Final Particles”. *L’analisi linguistica e letteraria*, 32(3): 177-198.

- Chafe, W. (1976) "Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View", in C.N. Li (Ed.) *Subject and Topic*. New York, Academic Press: 26-56.
- Conti, S. and G. Carella (2022) "Discourse Markers in Chinese and Italian: A Corpus-Driven Comparison of *ranhou* 然后 and *poi*". *L'analisi linguistica e letteraria*, 32(3): 151-176.
- Dong, X. 董秀芳 (2007) "词汇化与话语标记的形成". *Shijie Hanyu Jiaoxue*, 1: 50-61.
- Feng, G. (2019) "Pragmatic Markers in Chinese Discourse", in C. Shei (Ed.) *The Routledge Handbook of Chinese Discourse Analysis*. Abingdon/New York, Routledge: 216-229.
- Fernandez-Salgueiro, G. (2008) *Aspects of the Syntax of (TP-)Coordination, Across-the-Board Extraction and Parasitic Gaps*. PhD Dissertation. Michigan, University of Michigan.
- Féry, C. and S. Ishihara (2016) *The Oxford Handbook of Information Structure*. Oxford, Oxford University Press.
- Fischer, K. (2006) *Approaches to Discourse Particles*. Oxford, Elsevier.
- Frascarelli, M. (2007) "Subjects, Topics and the Interpretation of Referential Pro: An Interface Approach to the Linking of the (Null) Pronouns". *Natural Language and Linguistic Theory*, 25(4), 691-734.
- Frascarelli, M. and M. Casentini (2019) "The Interpretation of Null Subjects in a Radical Pro-Drop Language: Topic Chains and Discourse-Semantic Requirements in Chinese". *Studies in Chinese Linguistics*, 40(1), 1-45.
- Frascarelli, M. and R. Hinterhölzl (2007) "Types of Topics in German and Italian", in S. Winkler and K. Schwabe (Eds.) *On Information Structure, Meaning and Form*. Amsterdam/Philadelphia, John Benjamins: 87-116.
- Fraser, B. (1999) "What are Discourse Markers?". *Journal of Pragmatics*, 31: 931-952.
- Givón, T. (1983) "Topic Continuity in Discourse: An Introduction", in T. Givón (Ed.) *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Amsterdam/Philadelphia, John Benjamins: 5-41.
- Krifka, M. (2008) "Basic Notions of Information Structure". *Acta Linguistica Hungarica*, 55(3-4): 243-276.

- Kuno, S. (1976) "Subject, Theme, and the Speaker's Empathy: A Re-examination of Relativization Phenomena", in C.N. Li (Ed.) *Subject and Topic*. New York, Academic Press: 417-444.
- Lam, Z.W.-M. (2014) "A Complex ForceP for Speaker- and Addressee-Oriented Discourse Particles in Cantonese". *Studies in Chinese Linguistics*, 35: 61-80.
- Larson, R.K. (1988) "On the Double Object Construction". *Linguistic Inquiry*, 19(3): 335-391.
- Law, P. and J. Ndayiragike (2017) "Syntactic Tense from a Comparative Syntax Perspective". *Linguistic inquiry*, 48(4): 679-696.
- Lee, K.A. [1997] (2002) *Chinese Tone Sandhi and Prosody*. PhD Dissertation. Urbana/Champaign, University of Illinois.
- Li, C.N. and S.A. Thompson (1976) "Subject and Topic: A New Typology of Language", in C.N. Li (Ed.) *Subject and Topic*. New York, Academic Press: 457-489.
- (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley/Los Angeles/London, University of California Press.
- Liu, B. (2011) "Chinese Discourse Markers in Oral Speech of Mainland Mandarin Speakers", in Y. Xiao, L. Tao and H.L. Soh (Eds.) *Current Issues in Chinese Linguistics*. Newcastle, Cambridge Scholar: 364-405.
- Lopez, L. (2016) "Dislocations and Information Structures", in C. Féry and S. Ishihara (Eds.) *The Oxford Handbook of Information Structure*. Oxford, Oxford University Press: 402-421.
- Lü, S. 吕叔湘 (1999) 现代汉语八百词-增订本. Beijing, Shangwu Yinshuguan.
- Madaro, F. (2016) *La frase. Lingua cinese moderna standard*. Torino, Edizioni SEB27.
- Osa-Gómez, A. (2012) "Epistemic (Mis)Alignment in Discourse: What Spanish Discourse Markers Reveal". Unpublished manuscript.
- Perlmutter, D. (1971) *Deep and Surface Structure Constraints in Syntax*. New York, Holt, Rinehart and Winston.
- Pesetsky, D. (1987) "*Wh*-in-situ: Movement and Unselective Binding", in E. Reuland and A., ter Meulen (Eds.) *The Representation of (In)Definiteness*. Cambridge (MA), MIT Press: 98-129.
- Piccinini, C. (2020) "Analysis of the Pragmatic Uses of the Discourse Markers *na* 那 and *ránhòu* 然后 in a Corpus of Radio Conversa-

- tions in Chinese Language Recorded in Taiwan”, in *Italian Association for Chinese Studies. Selected Papers 3*. Venice, Libreria Editrice Cafoscarina: 107-120.
- Poletto, C. and G. Bocci (2016) “Syntactic and Prosodic Effects of Information Structure in Romance”, in C. Féry and S. Ishihara (Eds.) *The Oxford Handbook of Information Structure*. Oxford, Oxford University Press: 642-662.
- R Core Team (2020) *R: A Language and Environment for Statistical Computing*. Vienna, R Foundation for Statistical Computing. <https://www.R-project.org/> (visited 2023/02/20).
- Rizzi, L. (1986) “Null Objects in Italian and the Theory of *Pro*”. *Linguistic Inquiry*, 17: 501-558.
- (1997) “The Fine Structure of the Left Periphery”, in L. Haegeman (Ed.) *Elements of Grammar*. Dordrecht, Springer: 281-337.
- RStudio Team (2021) *RStudio: Integrated Development Environment for R*. Boston, RStudio, PBC. <http://www.rstudio.com/> (visited 2023/02/20).
- Saito, M. (2007) *Notes on East Asian Argument Ellipsis*. Unpublished manuscript. Nanzan University and University of Connecticut.
- Sankoff, G., P. Thibault, N. Nagy, H. Blondeau, M.-O. Fonollosa and L. Gagnon (1997) “Variation in the Use of Discourse Markers in a Language Contact Situation”. *Language Variation and Change*, 9(2): 191-217.
- Schiffrin, D. (1987) *Discourse Markers*. Cambridge, Cambridge University Press.
- Shi, D. (2000) “Topic and Topic-Comment Constructions in Mandarin Chinese”. *Language*, 76(2), 383-408.
- Stalnaker, R. (1974) “Pragmatic Presuppositions”, in M. Munitz and P. Unger (Eds.) *Semantics and Philosophy*. New York, New York University Press: 197-214.
- (2002) “Common Ground”. *Linguistics and Philosophy*, 25: 701-721.
- Stefanowitsch, A. (2020) *Corpus Linguistics: A Guide to the Methodology*. Berlin, Language Science Press.
- Su, L.I. (1998) “Conversational Coherence: The Use of *ranhou* in Chinese Spoken Discourse”, in S. Huang (Ed.) *Selected papers from the second international symposium on languages in Taiwan*. Taipei, Crane: 167-181.

- Thoma, S. (2016) *Discourse Particles in Bavarian German*. PhD Dissertation. Vancouver, University of British Columbia.
- Tsai, P. and W. Chu (2015) "The Use of Discourse Markers among Mandarin Chinese Teachers, and Chinese as a Second Language and Chinese as a Foreign Language Learners". *Applied Linguistics* 38(5): 638-665.
- Wang, C. and L.M. Huang (2006) "Grammaticalization of Connectives in Mandarin Chinese: A Corpus-Based Study". *Language and Linguistics*, 7(4): 991-1016.
- Wang, W. (2016) "Prosody and Discourse Functions of *ránhòu* 然后. With Implications for Teaching Mandarin Conjunctions at the Discourse Level", in H. Tao (Ed.) *Integrating Chinese Linguistic Research and Language Teaching and Learning*. Amsterdam, John Benjamins: 145-167.
- (2018) "Discourse Uses and Prosodic Properties of *ránhòu* in Spontaneous Mandarin Conversation". *Chinese Language and Discourse*, 9(1): 1-25.
- Wang, Y. (1997) "The Functions of *ránhòu* in Chinese Oral Discourse", in T.-F. Cheng, Y. Li, and H. Zhang (Eds.) *7th North American Conference on Chinese Linguistics (NACCL7) and 4th International Conference on Chinese Linguistics (ICCL4). Volume 2: Discourse, Historical Linguistics, Morphology, Phonology and Phonetics*. Los Angeles, GSIL (Graduate Students in Linguistics), University of Southern California: 380-397.
- Wiltschko, M. (2017) "Ergative Constellations in the Structure of Speech Acts", in J. Coon, D. Massam and L.D. Travis (Eds.) *The Oxford Book of Ergativity*. New York, Oxford University Press: 419-446.
- Xian, L. 鲜丽霞 and Y. Li 李月炯 (2015) "汉语话语标记研究综述". *Guangxi Shifan Xueyuan Xuebao*, 36(1): 122-127.