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.buponline.com
email: info@buponline.com

ISSN: 2283-8910
ISBN: 979-12-5477-177-8
ISBN online: 979-12-5477-178-5
DOI 10.30682/sitlec45

Grafica: Alessio Bonizzato
Impaginazione: Sara Celia

Prima edizione: aprile 2023

Table of contents

PREFACE

Serena Zuccheri

The book *Studies on Chinese Language and Linguistics in Italy* gathers fourteen papers written by eighteen members of the AILC - Associazione Italiana di Linguistica Cinese (Italian Association of Chinese Linguistics). Founded in Rome on 19 June 2017, the Association was created by a group of young scholars who, in 2014, started a series of conferences known as the "Study Days on Chinese Linguistics"[1]. This book, in fact, was conceived in 2021, after the Sixth Edition of the Study Days at the University of Bologna (Forlì campus) ended. Since one of the main AILC objectives is to develop exchange and collaboration among Italian and foreign scholars in Chinese linguistics, this volume aims to make the topics and the research fields covered by the members of the Association also known abroad. Thus, the contributions, which underwent double peer review, range from theoretical

1 The last meeting (the Seventh Edition of the Study Days) was held in 2022 at the University of Bergamo.

to applied Chinese linguistics, presenting original research, methodological aspects, and results obtained in applied research, as well as the presentation of the state of the art related to specific topics and fields.

In my capacity, as the editor of this book, I would like to thank the Executive Committee of the Association for its support and the authors who enthusiastically participated in this publication project with their interesting and remarkable research. Finally, my gratitude goes to the Department of Interpreting and Translation at the University of Bologna, which encouraged the publication of this volume.

I sincerely hope this book can be a source of inspiration for other young scholars, just as the Study Days promoted by the AILC are for its members.

ON PREFIXATION IN MODERN CHINESE*

Giorgio Francesco Arcodia

Ca' Foscari University of Venice

1. Introduction

In the literature on Chinese word formation, the (possible) distinction between the processes of 'derivation' and 'compounding' is still an unresolved issue. Word-formation elements which display high productivity and always appear in a fixed position with respect to the base word (in a particular usage), such as *xué* 学 'branch of knowledge' (as in *xīnlǐxué* 心理学 'psychology') have been analysed as affixes, as 'affixoids', or just as compound constituents. Also, it is often claimed that many proposed affixes of Modern Chinese, as e.g. *-huà* 化 '-ise, -ify', developed following a foreign model: specifically, it has

* The *Pinyin* romanisation system and simplified Chinese characters have been used as a default throughout the article. However, traditional characters are also used when needed for consistency with the source. The glosses follow the general guidelines of the Leipzig Glossing Rules: additional glosses include MOD 'marker of modification' and SFP 'sentence-final particle'.

been suggested that they entered the Chinese lexicon as constituents in complex words coined in Japan, which in turn followed European models (Masini 1993). In the case of possible prefixes, which received less attention than suffixes in the literature (see Arcodia 2012b), it has even been claimed that they all (or virtually all) derive from a foreign (usually, English or Japanese) model (Jia 2019). However, this largely depends on the definition of prefix(oid) which one chooses to adopt: for instance, Zhao (2018) proposes a number of prefixoids for which a foreign origin seems unlikely.

In this paper, we propose a reassessment of prefixation in Modern Chinese. Following Arcodia (2012b), we discard the 'prefix' *vs.* 'prefixoid' distinction, also because grammaticalized morphemes in Chinese (as well as in most languages of the Mainland East- and Southeast Asian area; Bisang 2004) very often do not show the formal correlates of grammaticalization. In the framework of Construction Morphology (Booij 2010), we treat potential prefixoids as fixed slots in a construction. In this analysis, the main differences between affixes/affixoids and regular compound constituents lie in their fixed position, their stable selectional properties and, above all, in the fixed, conventionalized meaning they contribute, as opposed to the more 'open' interpretation for compound constituents. The sample items we chose for our analysis are drawn from a selection of the literature on the topic; following Basciano and Bareato (2020), we also rely on web corpora and searches for the analysis of the use of complex words.

We will show that potential prefixes in Chinese have different properties: there are class-maintaining prefixes, class-changing prefixes, as well as prefixes with ambiguous properties with respect to word-class assignment. We will compare 'native' patterns and patterns which seem to follow a foreign model, showing that they do not constitute coherent subsets in terms of their behaviour. We will argue that the differences between prefixes and suffixes in Chinese may be partly explained by the different role of lefthand constituents and righthand constituents in compounding (unlike e.g. Romance languages). However, as conventionalised constructions used for word formation, prefixation patterns also have properties which do not fit in the general picture of headedness and word-class assignment in the morphology (and syntax) of Modern Chinese: above all, the fact that the word class of 'prefixed' words is often inconsistent with that of the corresponding base (non-prefixed) word, as e.g. *màoyì* 贸易 'commerce' > *fēi-màoyì* 非贸易 'non-commercial', but both endocentric nouns and adjectives are generally right-headed in Chinese (Ceccagno and Basciano 2007). We will argue that this is a major difference between prefixed and suffixed items in Chinese, since the latter always seem to define the word class of the complex word; also, it can be taken as an argument in favour of analysing prefixes as a separate morphological phenomenon, distinct from suffixation and from compounding.

This paper is organised as follows. In Section 2, we shall provide a critical overview of key notions in word formation, focussing on the distinctions between derivation and compounding, and between prefixation and suffixation, and on their application to Chinese. In Section 3, we shall introduce our sample and research methodology, and we shall present the results of our survey, offering our analysis of the data collected. In Section 4, we provide a summary of our findings, as well as some hints for further research.

2. Word formation: derivation *vs.* compounding, prefixation *vs.* suffixation

In the general literature on word formation, the definition of derivation and compounding are notoriously thorny issues (for an overview, see Lieber 2017; ten Hacken 2017). Both derivation and compounding have the function of creating new words, and their difference is normally associated with the nature of their constituents: while compounding is traditionally defined as a combination of words (e.g. Fabb 1998), stems (Bauer 1998), or lexemes (Lieber and Štekauer 2009), derivation involves affixes, or even nonconcatenative exponents (e.g. tone change, ablaut). This, in turn, leads to the issue of the delimitation of 'lexemes' and 'affixes', which is also far from controversial. Due to space limitations, we cannot provide an in-depth critical discussion of these issues here: we shall limit ourselves to a concise overview of these *vexatae quaestiones*, focussing on the aspects which are most relevant for the research described in the present paper.

2.1 Derivation and compounding: an overview

The categories of compounding and derivation may have fuzzy borders. Compounds may be hard to distinguish from syntactic phrases, on the one hand, and from derivation, on the other hand; derivation, in turn, may overlap not only with compounding, but also with inflection (Lieber 2017). As pointed out by Dressler (2006) and ten Hacken (2017), the delimitation of these processes of word formation tends to be theory-dependent: for instance, not all theoretical frameworks assume a strict boundary between words and phrases (more generally, between morphology and syntax), or between compound constituents and affixes (see e.g. Booij 2010). Moreover, any criteria set up to distinguish between compounding and derivation (or between compounds and phrases) might not necessarily work for all (or even just most, or many) languages. For instance, in Dutch, words ending in a voiced obstruent (as e.g. *hoofd* 'head, main') undergo final devoicing: this happens also when a word is used as the first constituent in a compound (*hoofdingang* 'main entrance'), but not when a derivational or inflectional affix starting with a vowel is added (as in *tweehofdig* 'two-headed'; ten Hacken 2017). This criterion might be used to distinguish affixation from compounding, but its application is limited to Dutch or any other language that follows a similar pattern, if existent (and it works only for items with specific sounds, thus lacking general applicability).

In our view, what really matters in this case is the relevance of any notion to be defined. As ten Hacken (2017) puts it,

> The question is not whether a definition is accurate but whether it delimits a useful theoretical concept. In a pretheoretical sense, it is no problem to use a language-specific definition with criteria that are easy to apply.

Thus, while the aim of the present paper is understanding whether prefixation can be established as a distinct phenomenon in Chinese word formation, our main concern is not applying a specific cross-linguistic definition of prefixation based on formal criteria (i.e. a derivational affix located to the left of the word/root). We rather focus on the issue of whether prefixation in Chinese possesses features which set it apart from other forms of derivation (mostly, suffixation) and from compounding: in other words, we want to find out whether prefixation is a "useful theoretical concept" for our understanding of Chinese morphology. Following Haspelmath, we distinguish between 'comparative concepts', i.e. "concepts created by comparative linguists for the specific purpose of crosslinguistic comparison" (Haspelmath 2010: 665), and language-specific descriptive categories, which have reality and may be defined only in individual languages: here we deal with the latter, rather than with the former.

Going back to the issue of the borderline between different processes (and objects), the distinction between compounding and phrases, and that between derivation and inflection, are not particularly significant for the purposes of the present study: Chinese has (almost) no inflection, and the distinction between compounds and phrases, while being sometimes problematic[1], is not called into question when we try to distinguish compounds from derived words.

As for the borderline between compounding and derivation, while there may be different definitions of these two phenomena (again, often also theory- and language-dependent), the *communis opinio* is that the most fundamental distinction concerns form, rather than function: as already pointed out above, compounding and derivation share the same basic function, but involve items of a different nature. However, as pointed out by Lieber (2017),

> [w]here word formation involves the combining of robustly contentful free lexemes we can be confident that we have compounding. Where one or more of the

[1] Indeed, several diagnostics have been proposed in the literature to assess wordhood for Chinese, as e.g. lexical integrity, conjunction deletion between coordinate items, freedom of the constituents, semantic compositionality, the number of syllables, exocentricity, productivity, and more (see Chao 1968; Huang 1984; Duanmu 1998, *inter alios*). However, these sometimes give conflicting results. For instance, a verb-object construction as *dān-xīn* 担心 'worry', lit. 'carry-heart', may often be separated, thus lacking lexical integrity, but it has a opaque/lexicalised meaning, and may be followed by an object: it thus shows properties which are intermediate between compound and phrase.

> formatives involved in word formation is bound, semantically less robust, and fixed in position, we may be confident that we have affixal derivation. The fuzzy borderline between compounding and affixal derivation lies where we find the combination of bound forms that are not fixed in position, or semantically more robust than typical, or that derive historically from or are related to free forms.

Thus, despite the traditional emphasis on form in the delimitation of compounding and affixation, meaning actually does play a role in this connection: Lieber (2017) speaks of "robustly contentful" (for compound constituents) and of "semantically less robust" (for affixes), but in practice the type of meaning which may be expressed by affixes is not easy to delimit. For instance, Bauer (2002) analysed the semantics of derivation in a sample of 42 languages, showing that there is a wide range of meanings which are associated with derivation in grammatical descriptions (as e.g. 'payment for N' or 'have a pain in N'). According to Beard (1998: 57 ff.), there are as many as four types of derivation, based on their function/meaning:

> a. "Featural derivation", i.e. the processes which do not change the category of the base, but rather alter its "inherent features", as e.g. gender (Italian *studente* 'student' > *studentessa* 'female student').
> b. "Functional derivation", i.e. the processes that alter the lexical semantics of the base (Italian *pizza* > *pizzeria* "pizza parlour"; English *employ* > *employee*).
> c. "Transposition", i.e. the processes which change the word class of the base only (English *lonely* > *loneliness*).
> d. "Expressive derivation", also known as evaluative morphology, i.e. the processes which add meanings such as GOOD, SMALL or BAD to the base, without assigning a part of speech to the base and without shifting its reference (Italian *gatto* 'cat' > *gattino* 'kitty'; see Grandi and Körtvélyessy 2015).

Thus, while categories a., c., and d. do fit in the label "semantically less robust" and are generally easy to identify as derivational, the same does not necessarily hold for b., i.e. functional derivation, as the meanings which can be expressed may overlap with that of lexemes: that is, they may be "semantically more robust than typical", as Lieber puts it[2]. In other words, if inflection involves grammatical meaning, and compounding involves lexical meaning, derivation is somehow in between the two, from the semantic point of view.
There are thus several cases of items which have an ambiguous status between compound constituent (hence, lexeme) and affix. Take, for instance, the case of French prepositions as *avant* 'before' and *sur* 'on/over': these are used also as prefixes in words like *avant-guerre* 'pre-war (years)', *sur-exposition* 'overexposure'. Amiot (2005) holds that they should be analysed as prefixes, since they never change the gender of the base, they can combine with words belonging to different classes, they form endocentric nouns and they are used to convey at least one meaning which is different from that (or those) of the corresponding preposition: for instance, *sur* conveys the meaning 'excessively, in excess', rather than 'on/over' (e.g. *surcharge* 'overload'). Indeed, wordhood (or, better, 'lexemehood') and affixhood are not immutable statuses, and items which are in between the two stages may be hard to classify (Bauer 2005). Compare, for instance, German *-heit/-keit* '-hood, -ness', as in *Freundlichkeit* 'friendliness', and *-voll* '-ful', as in *ehrenvoll* 'honourable': they are both derivational suffixes which originate from a lexeme, a free word. However, in Modern German *-heit/-keit* lost the connection with the Old High German word *heit / heid* 'person, status, rank, nature, kind' from which it derives (Lightfoot 2005), and hence its affixal status is not generally called into question; *-voll*, on the other hand, may still be easily related to the adjective *voll* 'full'. Just as seen above for French *avant* and *sur*, the affixal status of *-voll* may be argued on the basis of semantic and distributional differences with the

[2] In fact, even apparent instances of transposition (c.) may be ambiguous in this respect. For instance, English *-ness* may be interpreted as a case of transposition, as its main function is that of turning an adjective into a noun, it can be analysed as carrying the meaning 'the state of being X' / 'the quality of being X' (e.g. *loneliness* as 'the state of being lonely'), arguably more lexeme-like.

corresponding adjective: *-voll* has a conventionalised meaning which only partly overlaps with that of *voll*, it is always bound, and is always found in the same position (attached to the right of the base).
In order to label the formatives whose status is ambiguous between lexeme and *bona fide* affix, mostly because they also occur as free lexemes, the terms 'affixoids', 'pseudo-affixes' or 'semi-affixes' have often been used (see Lieber 2017 and the references cited therein). However, as hinted at earlier, not everybody agrees on the necessity of having an additional category of affixoids. For instance, the difference between 'affix' and 'affixoid' may not be very relevant if we see word formation as construction-based, as in the framework of Construction Morphology (Booij 2010; henceforth, 'CxM'). In a nutshell, CM applies the basic principles of Construction Grammar (Goldberg 1995; Michaelis and Lambrecht 1996) to morphology: in CxM, all word formation patterns are constructions, i.e. form-meaning-function complexes. According to Booij (2007:34), "[w]ord formation patterns can be seen as abstract schemas that generalize over sets of existing complex words with a systematic correlation between form and meaning". Thus, for instance, the construction underlying all English and Dutch endocentric compounds may be represented as such (see Booij 2010):

(1) $[X_i\ Y_j]_{Yk} \leftrightarrow [\text{SEM}_j \text{ with relation R to } \text{SEM}_i]_k$

A schema as (1) represents a very high level of generalisation, with many variables which may be filled by a broad range of items; the main restriction included in the schema is that the construction is right-headed, i.e. the right-hand constituent (Y_j) defines the whole compound (as e.g. in *swordfish*). In CxM, "affixoids" are defined as "morphemes which look like parts of compounds, and do occur as lexemes, but have a specific and more restricted meaning when used as part of a compound" (Booij 2005: 114). However, the 'affixoidal' meaning is part of the construction, rather than of the item itself, and as such it is available only in its use within a specific word formation schema. The difference between affixoids and affixes proper, in this framework, is that the former bear a word class, as they still have an obvious connection with a lexeme of the language, whereas affixes do not have a word class of their own, as "they only exist as parts of complex words, and as parts of abstract schemas for these complex words" (Booij 2007: 34).

Thus, in a CxM analysis, the distinction between affixoids and affixes is relatively unimportant: what matters most, in our opinion, is that both affixoids and affixes with a lexemic origin undergo similar processes of semantic evolution and, above all, they do so within a construction, based on generalisations over a set of paradigmatically related words; while affixoids have a word-level equivalent, they do not exist *as such* outside a word formation schema, just as ordinary affixes. Moreover, as pointed out earlier, the affix *vs.* affixoid distinction is arguably even less relevant for Chinese. We'll get back to this in the next section.

2.2 Derivation and compounding: the case of Modern Chinese

The delimitation of derivation and compounding is even more controversial for Chinese. Indeed, the issue of what is an 'affix' and what is a 'root', or possibly a 'word' in the Chinese lexicon is about as old as Modern Chinese linguistics (see Pan *et al.* 2004 for a brief history of the debate). To sum up very briefly the work that has been done on the topic of affixes and derivation in Chinese word formation, we may just say that there is a very small number of morphemes almost devoid of meaning which are regarded by most authors as affixes, and a large number of very productive bound constituents which, as said above, have a 'floating' status, as they may be analysed as derivational affixes as well, as affixoids, or just as compound constituents. Thus, there appears to be no general consensus either on what constitutes a 'genuine' derivational affix in Modern Chinese, or on whether productive derivation exists at all in Chinese, and on whether it is a significant word formation process (see Pan *et al.* 2004 and Arcodia 2012b for an overview).

The lack of general agreement on the interpretation of the fundamental constituents and processes of word formation in Modern Chinese is related to the nature of the lexicon of this language. In Chinese, both grammatical and lexical morphemes can be either free or bound; the great majority of Chinese morphemes are lexical and correspond to roots. About 70% of Mandarin roots are bound, just as affixes (Packard 2000): however, the distinction between bound and free roots is not always clear-cut, since some bound roots can be used as free roots in specific contexts, and there are no formal or semantic differences between the two. Compare:

(2) 我不喜欢那个人。
wǒ *bù* *xǐhuan* *nà-ge* *rén*
1SG NEG like that-CLF person
'I don't like that person'

(3) a. 你是哪国人？
nǐ *shì* *nǎ-guó-rén*
2SG COP which-country-person
'where are you from?'
b. 我是法国人。
wǒ *shì* *Fǎ-guó-rén*
1SG COP France-country-person
'I am French'

In (2), the lexical morpheme *rén* 人 'person' is used as a free root, i.e. as a syntactic word, while in (3a-b) it is used as a bound constituent in complex words: note that there are no formal or semantic differences between those uses. In (3a-b), the lexical morpheme *guó* 国 is used as a bound lexical root: this is the only option available for this item, i.e. it can (almost) never be a syntactically free form. However, the type of meanings conveyed by free and bound lexical roots are the same, and both classes of roots are equally active in word formation (Sproat and Shih 1996; Basciano and Ceccagno 2009); thus, there appears to be no real reason to treat complex words formed by bound roots as different from those formed by free roots only (see Arcodia and Basciano 2017).

Just as seen in the preceding section for French and German (and other languages of Europe), in Chinese too we may find roots which appear in a specific position with a specific meaning, somehow different from that of the other uses of that root. As mentioned above, these have been treated either as affixes (e.g. Yip 2000), as 'affixoids' (*lèicízhuì* 类词缀 or *zhǔncízhuì* 准词缀; Ma 1995) or just as compound constituents (Dong 2004). As for items in the prefix position, we may take as an example *duō* 多 'many, much'. *Duō* 多 has different meaning and distribution when used as a free word (4a-b, 5) and as a prefix (6; exx. from Arcodia 2012b: 192-193):

(4) a. 她認識很多外國人

tā	*rènshi*	*hěn*	*duō*	*wàiguórén*
3SG.F	knows	very	many	foreigner

b. *她認識多外國人

tā	*rènshi*	*duō*	*wàiguórén*
3SG.F	knows	many	foreigner

'She knows many foreigners'

(5) 王朔的著作很多

Wáng	*Shuò*	*de*	*zhùzuò*	*hěn*	*duō*
Wang	Shuo	MOD	work	very	many

'Wang Shuo's works are numerous'

(6) 中國是多民族國家

Zhōngguó	*shì*	*duō-mínzú*	*guójiā*
China	COP	many-nationality	country

'China is a multiethnic country'

With the exception of a few set phrases, the adjective *duō* 多 'much, many' can modify a noun only if preceded by another modifier (Lü 1980), sometimes semantically redundant, as *hěn* 很 'very' in (4a); (4b) is therefore ungrammatical. Also, *duō* 多 can have a predicative function and appear after the topic, as in (5). In (6), however, *duō* 多 is conjoined to the noun *mínzú* 民族 'nationality' without being itself modified. Also, it carries a different semantic value from the adjective *duō* 多: whereas adding *hěn duō* 很多 to *wàiguórén* 外国人 in (4a) 'adds' meaning to the noun without altering its word class and distributional properties, in (6) the morpheme apparently turns the base noun into a non-predicative adjective, an attributive form. Here *duō* 多 does not only mean 'many' as in (4a), but rather 'having many X'[3]; in the formalism of CxM, it may be represented as such:

[3] Note that we do find *duō* 多 complex words belonging to the nominal class, as e.g. *duōbiānxíng* 多边形 'polygon' (lit. 'many-side-shape'). However, we analyse

(7) $[\textit{duō}\ [X]_{Ni}]_{ADJk} \leftrightarrow [\text{having many } \textsc{sem}_i]_k$

So, we have clear distributional and semantic differences between the adjectival use and the prefixal use of *duō* 多: therefore, we believe that *duō* 多 may be analysed as a prefix.

As for the prefix *vs.* prefixoid distinction, we already pointed out that its significance is limited in a constructional perspective: this is even truer for a language as Chinese, for at least two reasons. Firstly, as hinted at earlier, we often see items with lexical and (more) grammatical functions without (clear) formal differences, i.e. primary grammaticalization without secondary grammaticalization (in the sense of Traugott 2002)[4]; this is the case not only for Chinese, and it is also generally true for languages found in a region within East and Mainland Southeast Asia (Bisang 2004). While there are indeed cases of grammaticalised items which underwent significant formal changes (reduction) in Chinese too, this is the exception, rather than the norm: thus, having derivational (or, better, derivation-like) formatives which are identical to lexical (bound or free) roots is to be expected. Secondly, in the preceding section we pointed out that, in CxM, the fundamental difference between affixoids and affixes proper is that the former bear a word class: however, Chinese bound roots cannot be unambiguously assigned to a word class, unless we assume semantic criteria to distinguish word classes, since words and roots have no category-specific morphology or phonological features (Basciano 2017). Arcodia and Basciano go as far as to propose that, for Chinese, (CxM) word formation schemas are specified only for meaning and for the word class of the whole construction (word; Arcodia and Basciano 2018: 237-238):

> any element, free or bound can enter the schema as long as its semantics is compatible with it. In other

such cases as formed through a two-step process: *duōbiān* 多边 'multilateral, having many sides' + *xíng* 形 'shape' = *duōbiānxíng* 多边形 'polygon (in CxM terms, [[*duō* [X]Ni]ADJj [Y]Nk]Nl; see Arcodia 2012b: 194 for a similar proposal for *wú-* 无 '-free, -less'). On nominal uses for *duō* 多 complex words, see also Yang (2007).

[4] For instance, in Modern Chinese *zài* 在 is a verb, meaning 'to be at', but also a locative preposition ('at, in') and a marker of progressive aspect (Bisang 2004).

> words, the lexical category of the compound constituents is not specified […] The constituents […] can be in principle any root, free or bound, or any other word type chosen only by virtue of their semantics and inserted as such in the schemas. […] Once the individual item is matched to a constructional schema, features as word class assignment for constituents do not appear to be relevant, as what matters most to the language user are the properties of the construction as a whole.

While we do not necessarily agree *in toto* with this view, we do believe that one should not attach too much importance to the word class attributes of a formative to distinguish different degrees of affixhood, at least for Chinese[5].

Thus, to sum up, we treat as derivational affixes all instances of formatives which develop a different meaning from that of the corresponding lexeme, if this meaning is available only when the morpheme at issue is part of a complex word, appearing in a fixed position; also, we expect that, typically, the 'affixal' meaning be somehow more general or 'abstract' (for lack of a better word) than that of the corresponding lexeme (see Arcodia 2011: 127-128). Note that, in a CxM perspective, we treat those properties as being part of the construction, rather than belonging to the formative by itself.

2.3 Prefixation *vs.* suffixation

In the literature on Chinese word formation, prefixation has apparently received less attention than suffixation, and the criteria proposed to define prefixes seem to be even vaguer than those proposed for suffixes (Arcodia 2012b; see the overview in Xu and Cai 2007 and Yang 2007). Indeed, in the Western linguistic tradition too the recognition

[5] In the Chinese linguistic literature, opinions differ as to the necessity of having two distinct categories of 'affix' and 'affixoid': for instance, Sun (2000) argues against this distinction for Chinese, while Jia (2019), as well as many others (see the sources quoted in Arcodia 2012b) believe that these two categories actually have different properties, and should be kept apart. Due to space constraints, we shall not discuss the argumentations proposed for and against the differentiation of affixoids from 'true' affixes in Chinese.

of the existence of prefixes came much later than that of suffixes, and for a long time prefixing was regarded as a special kind of compounding, rather than derivation. This was because, among other reasons, many present-day prefixes in Standard Average European languages were actually prepositions or adverbs in Latin and Ancient Greek; also, just as seen above for French (Section 2.1), there still are prefixes whose phonological form is identical to that of prepositions (Montermini 2008). Also, the treatment of prefixes in the literature on word formation has sometimes suffered from a 'Eurocentric' bias: for instance, much emphasis has been put on the fact that prefixes do not normally bear a word class (i.e. the word class of prefixed words is the same as that of the base), but this does not necessarily hold for non-European languages (and, indeed, we have cases of class-changing prefixation also in the languages of Europe)[6]. To this we may add the fact that, generally speaking, there is a cross-linguistic robust trend to prefer prefixes to suffixes, both for derivation and for inflection (see e.g. Himmelmann 2014).
In a recent paper, Jia (2019) offers a discussion of the fundamental differences between prefixation and suffixation (she actually speaks of 'prefixoids' and 'suffixoids'; see fn. 5) in Chinese. The main differences she lists are the following:

a. While suffixes may be native or 'imported' (i.e. loans or calques of affixes in other languages, chiefly English and Japanese), virtually all prefixes have a foreign origin.
b. While suffixes add various types of meaning to the base they attach to, and bear a word class, the semantic contribution of prefixes tends to be the expression of logical notions as 'true', 'false', or 'similar'(/'divergent').
c. While suffixes may attach to words (or bound roots) or phrases, and always form words, prefixes generally attach to words (or bound roots) only.
d. While suffixes tend to offer background information, the focus being on the base, prefixes are focal, with the base offering background information.

[6] For instance, English *en-* in *ennoble*, or Italian *anti-*, forming adjectives from nouns (*antiforfora* 'anti-dandruff' Micheli 2020: 49).

e. Suffixes are generally more grammaticalised than prefixes, i.e. prefixes tend to retain a meaning closer to that of the corresponding lexeme, while suffixes tend to undergo more semantic bleaching.

Most of the above, however, require further discussion. Firstly, as hinted at in Section 1, the claim that virtually all prefixes derive from a foreign model largely depends on what items one chooses to treat as prefixes(/prefixoids). For instance, Zhao (2018) proposes a number of prefixes for which a foreign origin seems unlikely, as e.g. *bèi* 被 ‘PASS’ for ‘being X-ed falsely or forcedly’: see *bèi-jiù-yè* 被就业 ‘PASS-get-job’, ‘being counted as employed while unemployed’, or also ‘being forced to take up a job’ (see also Han 2012)[7]. Moreover, as we shall see in Section 3.2, even when a correspondence with a construction in a foreign language may be found, this does not entail that there is a (total or partial) overlap.

The same applies to b.: elements which can be claimed to be prefixes may express a fairly broad range of meanings, depending again on what one chooses to include in this category (see the case of *bèi-* 被). However, it is true that Chinese prefixes all modify the meaning of the base, and hence they are somehow functionally restricted: this is arguably related to a more general principle, namely that modifiers are invariably located to the left in Chinese, both in syntax and in word formation (we will get back to this later)[8]. On the other hand, class-changing prefixes seem to do more than just modifying the meaning of the base, and do bear a word class. Indeed, as mentioned in Section 1, prefixal formatives in Chinese can be either class-maintaining, class-changing, or both (Arcodia 2012b): for instance, *fù-* 副 ‘deputy, vice-’ is class-maintaining (*huìzhǎng* 会长 ‘president [of an associa-

---

7 Occasional examples of constructions as *he was suicided* (= he was actually murdered, but they made it look as if it were a suicide) may be found also in English (and elsewhere). However, this overlaps only partly with the Chinese pattern at issue here, and we believe the two are not directly related.

8 Interestingly, the range of meanings expressed by prefixation seems to be more limited than that of suffixation also in an Indo-European language of Europe as Italian. Prefixes tend to express meanings belonging to a rather restricted list of functions, as e.g. negation, quantification, repetition, while suffixes have a much broader range of functions (see Micheli 2020: 54-58, 62-64).

tion]' > *fùhuìzhǎng* 副会长 'vice-president'); *duō-* 多 'multi-' often forms (non-predicative) adjectives (6); *fēi-* 非 'non-, un, in-' may both be class-maintaining (*huìyuán* 会员 'member' > *fēihuìyuán* 非会员 'non-member') and class-changing (*màoyì* 贸易 'trade' > *fēimàoyì* 非贸易 'non-commercial'). Moreover, note that *duō-* 多 'multi-' forms items which are used not only as non-predicative adjectives, but also as adverbial modifiers (Yang 2007). This is the case e.g. for *duōfāngmiàn* 多方面 'many-faceted', 'in many ways' (example from the BCC corpus; note the use of the marker of adverbial modification *de* 地):

(8) 在阿富汗战争中，伊朗不仅与

zài	*Āfùhàn*	*zhànzhēng*	*zhōng*	*Yīlǎng*	*bù-jǐn*	*yǔ*
in	Afghan	war	in	Iran	not-only	with

沙特进行了合作，而且也

Shātè	*jìnxíng-le*	*hézuò*	*érqiě*	*yě*
Saudi	carried.out-PFV	cooperation	moreover	also

多方面地与美国进行了合作。

duōfāngmiàn-de	*yǔ*	*Měiguó*	*jìnxíng-le*	*hézuò*
in.many.ways-ADV	with	US	carried.out-PFV	cooperation

'During the Afghan War, Iran not only cooperated with the Saudis, but also carried out cooperation on many fronts with the US'

As for d., Jia's arguments are not really convincing: it remains unclear to us how prefixes, which, as said just above, mostly perform the function of modifiers, should be seen as more 'foregrounded' than suffixes. Point e., namely that suffixes are more 'bleached' than prefixes, may also be related to the type of meanings conveyed by prefixes. While the grammaticalization of derivational affixes, as said above, generally involves some degree of abstraction (see Arcodia 2011, 2012b), meanings conveyed by prefixes are often more generic, as e.g. negation or quantification, and hence more apt to be used as prefixes with a meaning very close to that of the corresponding lexeme (6); suffixes, on the other hand, sometimes derive from lexical morphemes with a 'richer' (for lack of a better word) intentional meaning (see e.g. *xìng* 性 'inherent property', 'immutable nature' > *-xìng* 性 'the property of X / connected with X'; Arcodia 2012b). This does not entail, however, that the mechanisms by means of which prefixes and suffixes are created are necessarily different: for instance, metaphorical abstraction (in the sense of Heine, Claudi and Hünnemeyer 1991) seems to be involved in the evolution of *luǒ* 裸 'naked' into *luǒ-* 裸 'lacking, unprepared' (e.g. *luǒhūn* 裸婚 'naked marriage', i.e. getting married without a car, a house, a wedding ceremony).

Thus, to sum up, what we may learn from the literature on word formation in Chinese is that prefixes and suffixes appear to have very different features. However, not all of those seem to be able to stand the empirical test; also, some differences between prefixes and suffixes may be related to general principles of word formation (and constituent order) in Chinese, or may not necessarily apply to all the items which can be considered as valid 'candidates' for affixal status.

3. Our survey

As said earlier, the aim of this paper is to assess whether prefixes form a consistent class of morphological formatives in Chinese, separate from suffixation and compounding, and whether they are significant for our understanding of Chinese word formation. To this end, we shall analyse the behaviour of a sample of items which are often treated as prefixes(/prefixoids) in the literature.

3.1 Sample and methodology

The formatives considered for this study were chosen from a selection of works on the topic (Xu and Cai 2007; Yang 2007; Hu 2018; Jiang and Li 2018; Zhao 2018; Jia 2019). Note that we decided not to include some items which are nearly universally regarded as prefixes, namely *dì* 第, added to cardinal numerals to form ordinals, *ā* 阿, a prefix expressing endearment (usually with kinship terms and names), *lǎo* 老 'old' and *xiǎo* 小 'young', used before names for people, respectively, older (or of the same age) or younger. The first of those is extremely regular and predictable, and is used with any number: it is actually somehow close to inflection, due to the generality of its meaning and to its universal applicability. The latter three are used with very specific subsets of lexemes and belong to the domain of evaluative morphology (see Arcodia 2015), which usually deserves a separate treatment due to its peculiarities. In Table 1, we list the items we considered for our study.

Item	**Gloss**	**Example**
非 *fēi-*	'non-, un-, in-'	非官方 *fēiguānfāng* 'unofficial'
副 *fù-*	'deputy, vice-'	副校长 *fùxiàozhǎng* 'vice-principal'
多 *duō-*	'multi-'	多功能 *duōgōngnéng* 'multi-functional'
高 *gāo-*	'high-'	高蛋白 *gāodànbái* 'high protein'
低 *dī-*	'low'	低糖 *dītáng* 'low sugar'
零 *líng-*	'zero'	零风险 *língfēngxiǎn* 'zero risk'
无 *wú-*	'-free, -less'	无糖 *wútáng* 'sugar free'
半 *bàn-*	'half, semi-'	半导体 *bàndǎotǐ* 'semiconductor'
准 *zhǔn-*	'quasi-'	准词缀 *zhǔncízhuì* 'quasi-affix'
类 *lèi-*	'pseudo-'	类词缀 *lèicízhuì* 'pseudo-affix'
超 *chāo-*	'ultra-'	超声波 *chāoshēngbō* 'ultrasonic wave'
反 *fǎn-*	'anti-'	反间谍 *fǎnjiàndié* 'counterespionage'
软 *ruǎn-*	'soft'	软暴力 *ruǎnbàolì* 'non-physical violence'
轻 *qīng-*	'light'	轻移民 *qīngyímín* 'short-distance immigrants'
被 *bèi-*	'forcedly, falsely'	被自杀 *bèizìshā* 'death claimed to be suicide'
裸 *luǒ-*	'naked, unprepared'	裸婚 *luǒhūn* 'marriage without material basis'
前 *qián-*	'former'	前总统 *qiánzǒngtǒng* 'former president'
闪 *shǎn*	'fast, unexpectedly'	闪退 *shǎntuì* 'to crash'
可 *kě-*	'-able'	可吃 *kěchī* 'edible'
好 *hǎo-*	'easy. pleasant to'	好听 *hǎotīng* 'pleasant to hear'
难 *nán-*	'hard, unpleasant to'	难看 *nánkàn* 'ugly'

Table 1. List of the items included in the present study.

Since the formatives analysed have all been treated as prefixes(/prefixoids) in the literature, we rely on our sources for their basic characterisation. However, as stated in Section 1, we also rely on web corpora (chiefly, the BCC Corpus) and on raw web searches to validate our conclusions. Due to space limitations, we shall not discuss in detail the origins of each of these items and their pathway of diachronic evolution.

3.2 Data and analysis

As pointed out in Section 2.3, not all the formatives which may be regarded as prefixes in Chinese may be claimed to be foreign origin. Looking at the items in Table 1, we may remark that this applies also to those formatives which could be related to foreign word formation patterns. For instance, *wú* 无 'lacking X' is close to the English formatives'-free, -less', but with an obvious constructional difference. Moreover, even those who seem to correspond to prefixes in English (or other languages) do not necessarily overlap wholly with the foreign model, both in terms of meaning and in terms of distribution. For instance, *fēi-* 非 corresponds to more than one English formative, namely 'non-' and 'un-'; *duō-* 多 corresponds to *multi-* and *poly-*. Also, several words formed with *qīng-* 轻 to the left of the root do correspond to English complex words (or phrases) with *light*, as e.g. *qīngshí* 轻食 'light meal', and could well be calques (see Jiang and Li 2018); however, there are other *qīng-* 轻 words which do not seem to be directly related to English words (see *qīngyímín* 轻移民 'short-distance immigrants' in Table 1), and, anyway, the 'prefixal' meaning developed by *qīng-* 轻 does not completely overlap with the use of *light* in English word formation. For instance, *qīngshíshàng* 轻时尚 '(lit.) light fashion' is used to refer to a fashionable but simple, minimalistic lifestyle (green and low-impact): this shows that the meanings of *qīng-* 轻 is not only quite distinct from that of the corresponding adjective (*qīng* 轻 'light'), but also from the metaphorical senses of *light* in English.

Moreover, even if we limit our analysis to prefixes which have been claimed to be of foreign origin, we can see that their behaviour is not consistent in terms of word-class assignment. In Section 2.3, we provided an example for each type, namely a class-maintaining prefix (*fù-* 副 'deputy, vice-'), a class-changing prefix (*duō-* 多 'multi-', and a prefix which is sometimes class-maintaining, and sometimes class-changing (*fēi-* 非 'non-, un-, in-'): however, e.g. in Jia (2019) these three formatives are all seen as being coined due to the influence of English (they are basically described as calques). On the other hand, a clearly native prefix as the above-mentioned *bèi-* 被 'forcedly, falsely' appears to be class-changing: it can attach to words belonging to virtually any word class, and the construction is generally predicative, i.e. basically a verb (Han 2012). See the following example (from the BCC corpus):

(9) 大家都被小康了吧。

dàjiā	*dōu*	*bèi-xiǎokāng*	*le*	*ba*
everybody	all	PASS-moderately.affluent	PFV/PERF	SFP

'Let's say that everybody has been turned into 'moderately affluent''

In (9), the combination of *bèi-* 被 with the adjective *xiǎokang* 小康 'moderately affluent' is used as an inchoative verb[9]. Han (2012) points out that, in a minority of cases, *bèi-* 被 complex words may be used as topics or objects, which is the typical function of nouns: however, given the degree of word class flexibility of Modern Chinese, this is hardly surprising, as it is often the case that verbs can be 'nominalised' without any overt formal change (see Kwong and Tsou 2003).

Thus, in short, 'imported' and 'native' patterns of prefixation do not appear to form a consistent set, and even prefixes which likely follow a foreign model are not necessarily consistent with the 'source' construction. Incidentally, this is hardly unusual in Chinese word formation: there are plenty of examples of word formation patterns which are first coined after the model of a parallel construction in another language, but then develop independently in Chinese, as e.g. the suffix *-zú* 族 'clan, a category/group of people with common characteristics or behaviour' (*dītóuzú* 低头族 'smartphone zombies', lit. 'lower head clan'), originally from Japanese *-zoku* 族 (see Basciano and Bareato 2020).

The issue of word-class assignment is indeed relevant for our understanding of the place of the constructions considered here in Chinese word formation. In Chinese compounds, the head may be located either on the left or on the right, depending on the specific pattern: generally speaking, subordinate endocentric (in the sense of Bisetto and Scalise 2005) complex verbs are left-headed (as e.g. *kāibān* 开班 'open a class / course', *kāi* 开 'open, operate' being a verb); all other subordinate and attributive endocentric compounds are right-headed (Ceccagno and Basciano 2007). Thus, only verbs can be left-headed, while endocentric compounds belonging to all other word classes (and

9 This expression was created as a form of satire on statistics which are aimed at obtaining a certain result, rather than at providing a faithful picture of the situation, just as *bèi-jiù-yè* 被就业 'being counted as employed while unemployed' seen above (Section 3.2).

attributive endocentric compound verbs) are right-headed (with the exception of coordinate endocentric compounds, in which all constituents may be regarded as heads). Then, needless to say, we have subordinate, attributive and coordinate exocentric compounds, in which no constituent has the role of the head: in a CxM perspective, features as the word class of the complex word belong to the construction (and may be inherited from superordinate constructions; see Arcodia and Basciano 2018). Thus, Jia's (2019) proposal mentioned earlier that only suffixes, but not prefixes may bear a word class, and that the semantic contribution of prefixes is limited to a subset of meanings, would be in line with an analysis of Chinese word formation in which derivation is analogous to compounding: prefixes, as all modifiers, are not heads, do not determine the word class of the complex word, and provide a semantic contribution of a certain type; suffixes would be treated as heads (see e.g. Scalise 1990), determining the word class of the complex word.

However, as hinted at above, this is certainly true for suffixes, and might work for class-maintaining prefixes, but definitely not for class-changing prefixes. Indeed, the application of the notion of 'head' to derivation is problematic also for the languages of Europe (for an overview, see Arcodia 2012a): if we consider (only or mostly) headedness in the categorial sense (see Scalise and Fábregas 2010), this may apply to derivation too, but then it would be quite different from the notion of head used for compounding (which is much broader; Arcodia 2012a). Indeed, not everybody agrees on attributing head properties to either the base or the affix in derivation (see e.g. Zwicky 1985): in CxM, as mentioned earlier (Section 2.1), the semantic and categorial features of complex words are part of the construction itself in derivation; in endocentric compounding, the schema contains the specification of which constituent is co-indexed for word class (and other features) with the whole word (Booij 2010). Indeed, in a constructionist lexicon one may have schemas which generalise over what, at face value, look like 'left-headed' derived word, i.e. prefixed words whose word class (and selectional features) are not determined by the base, as e.g. English *be-* in *behead* (adapted from Booij 2010: 29):

(10) $[\textit{be}\ [\mathrm{X}]_{\mathrm{Nj}}]_{\mathrm{Vi}} \leftrightarrow [\textsc{remove}\ [\textsc{sem}_{j}]\ ]_{i}$

Thus, instead of attributing a word class (which would 'percolate' to the whole word) and other features to the prefix, in the framework of CxM we may say that (Arcodia 2012a: 382):

> […] the notion of 'head' in derivation is superseded by that of construction; the inconsistencies which result from the application of the syntactic notion of 'head' to derivation are not characteristic of an approach in which affixes are just exponents, the semantic contribution is a property of the construction and the identity or non-identity of the part of speech label of the base lexeme and of the derived word is also construction-specific.

This represents a major difference from what we said above about compounding: in the only productive construction for left-headed compounds in Chinese, categorial features are anyway co-indexed with that of the lefthand constituent.

In this approach, we may easily account for cases as *bèi-* 被, in which what was(/is) a passive marker, i.e. a function word, may turn a noun or an adjective into a verb. We may represent the construction underlying *bèi-* 被 complex words as such:

(11) $[\textit{bèi}\ [\mathrm{X}]_{\mathrm{X}j}]_{\mathrm{V}i} \leftrightarrow [\text{BE FALSELY/FORCEDLY}\ [\text{SEM}_j]\ ]_i$

The fact that items belonging to just about any word class may be used as bases in the schema in (11) is indicated by the 'X' in the word class slot for the variable, and the 'V' associated with the whole construction is not co-indexed with *bèi-* 被: it is a property of the construction, not of any of its constituents. Note that schemas as (11) are not created arbitrarily: as said above (Section 2.1), in CxM schemas are generalisations over existing words with a regular correspondence between form and meaning: in other words, the creation of *bèi-* 被 words provides language users with a stimulus to posit a generalisation, which guides them in the interpretation and creation of novel words formed according to the schema. This is different from simple analogy, since the connection with the model word(s) which is part of the process of analogy may be lost once a schema is established. Indeed, with the coining of new words according to a schema, the schema itself may develop further, and sub-schemas may be created.

Thus, to sum up, prefixal items of Modern Chinese do seem to possess features which set them apart both from suffixes and from regular compound constituents. While suffixes (or, better, constructions based on suffixes) always seem to determine the word class of the whole word, and they tend to adhere to the same general principles of constituent order and category/feature assignment as compounding, prefixes stand out as having distinct properties. Class-maintaining prefixes resemble compound constituents: their semantic contribution is similar to that of modifiers in attributive compounding, with which they share the lefthand position, and they never alter the word class (or other major features) of the complex word. Class-changing prefixes, on the other hand, are unlike any compound schema, not only since they do change the word class of the base lexeme, while in compounding the lefthand constituent determines the word class of the compound only in subordinate endocentric verbs but also because the word class they assign may not be associated with the prefixal item itself (as shown e.g. in 11).

Also, note that even though, as said above, class-maintaining prefixes are similar to modifiers in attributive compounding, there is indeed an important difference between the two (as hinted at in Section 1): as conventionalised constructions, prefixes tend to have a more 'regular' behaviour than compound constituents, as for semantic interpretation and selectional properties. Thus, for instance, in an attributive compound as *snail mail*, the modifier *snail* is used metaphorically to mean 'slow', and all other features of the noun *snail* are ignored; also, at least in some languages, as English (and other Germanic languages) and Chinese (Basciano, Kula and Melloni 2011), the interpretation e.g. of noun-noun compounds is less predictable, and may even be context-dependent, while derived words tend to have a regular interpretation (as specified in the word formation schema; on selection in derivation and compounding, see Scalise *et al.* 2005). Compare, for instance, the contribution of *lóng* 龙 'dragon' in the compounds *lóngchuán* 龙船 'dragon boat' (i.e. a dragon-shaped boat), *lóngtào* 龙套 'dragon costume' (costume with dragon designs), and *lóngwǔ* 龙舞 'dragon dance' (traditional dance in which a team of dancers manipulate a puppet dragon). A class-maintaining prefix as e.g. *qián-* 前 'former-', however, always contributes the same meaning to the complex word formed according to its schema: compare *qiánzǒngtǒng* 前总统 'former president', *qiánxiàozhǎng* 前校长 'former principal', and *Qiándōngdé* 前东德 'former East Germany'.

4. Summary and conclusions

In this paper, we tried to show that, in Modern Chinese, prefixation is a distinct word formation process with features which set it apart both from compounding and from suffixation. Firstly, based on Arcodia (2011, 2012b) we argued that derivation may be regarded as independent from (albeit related to) compounding, as derivational affixes (generally) emerge from compound constituents, but they develop different features from their source items/constructions: while a further distinction between affixes proper and affixoid might be viable for some languages and for some items, we maintain that it is not particularly significant for Chinese, especially in the perspective of CxM. Secondly, we proposed that prefixes may both be native to Chinese and (possibly) coined due to the influence of foreign languages, but these two sets are not consistent in terms of behaviour: more generally, while suffixes tend to share a number of fundamental features (chiefly, word-class assignment), prefixes are more varied, with different relationships with the whole word. Applying the formalism of CxM, we attributed those semantic and categorial features of prefixes to the word formation schema itself, rather than to individual constituents (i.e. either to the prefix or to the base). We then discussed how prefixes differ from modifiers in compounding, arguing that the former, but not necessarily the latter, provide a stable semantic contribution to the complex word, and may even be part of schemas which bear a word class independent from that of the base. Suffixes, on the other hand, are more consistent with compound constituents in this respect, since they conform to the most common model in Chinese word formation: namely, having the righthand constituent determining the word class of the complex word. However, again, we argued that it is the construction, rather than the suffix itself, which bears the word class.

Unfortunately, due to space limitations, we could not discuss the semantic evolution of the formatives considered here, nor could we provide quantitative measures of the 'ambiguity' of those prefixes which may be both class-maintaining and class-changing. Also, we could not discuss in detail the trends in part of speech assignment for class-changing prefixes: Yang (2007) suggests that prefixes can build both non-predicative adjectives and manner adverbs (see Exx. 6-8), but in our sample we have also constructions whose output is a verb (11). We leave this for further research.

References

Amiot, D. (2005) "Between Compounding and Derivation – Elements of Word-Formation Corresponding to Prepositions", in W. Dressler, D. Kastovsky, O.E. Pfeiffer and F. Rainer (Eds.) *Morphology and its Demarcations*. Amsterdam, John Benjamins: 183-195.

Arcodia, G.F. (2011) "A Construction Morphology Account of Derivation in Mandarin Chinese". *Morphology*, 21: 89-130.

--- (2012a) "Constructions and Headedness in Derivation and Compounding". *Morphology*, 22: 365-397.

--- (2012b) *Lexical Derivation in Mandarin Chinese*. Taipei, Crane.

--- (2015) "Chinese", in N. Grandi and L. Körtvélyessy (Eds.) *The Edinburgh Handbook of Evaluative Morphology*. Edinburgh, Edinburgh University Press: 352-360.

Arcodia, G.F. and B. Basciano (2017) "Morphology, Modern", in R. Sybesma, W. Behr, Y. Gu, Z. Handel, J.C.-T. Huang and J. Myers (Eds.) *Encyclopedia of Chinese language and Linguistics*, vol. 3. Leiden, Brill: 104-118.

--- (2018) "The Construction Morphology Analysis of Chinese Word Formation", in G. Booij (Ed.) *The Construction of Words. Advances in Construction Morphology*. Berlin, Springer: 219-253.

Basciano, B. and S. Bareato (2020) "Chinese Affixes in the Internet Era: A Corpus-Based Study of X-族 *zú*, X-党 *dǎng* and X-客 *kè* Neologisms", in B. Basciano, F. Gatti and A. Morbiato (Eds.) *Corpus-Based Research on Chinese Language and Linguistics*. Venezia, Edizioni Ca' Foscari: 237-279.

Basciano, B. and A. Ceccagno (2009) "The Chinese Language and Some Notions from Western Linguistics". *Lingue e Linguaggio*, 8(1): 105-135.

Basciano, B., N. Kula and C. Melloni (2011) "Modes of Compounding in Bantu, Romance and Chinese". *Italian Journal of Linguistics*, 23(2): 203-249.

Bauer, L. (1990) "Be-Heading the Word". *Journal of Linguistics*, 26: 1-31.

--- (1998) "Is There a Class of Neoclassical Compounds, and If So is It Productive?". *Linguistics*, 36: 403-422.

--- (2002) "What Can You Do with Derivational Morphology", in S. Bendjaballah, W.U. Dressler, O.E. Pfeiffer and M.D. Voeikova (Eds.) *Morphology 2000. Selected Papers from the 9th Morphology Meeting*. Amsterdam/Philadelphia, John Benjamins: 37-48.

--- (2005) "The Borderline between Derivation and Compounding", in W. Dressler, D. Kastovsky, O.E. Pfeiffer and F. Rainer (Eds.) *Morphology and its Demarcations*. Amsterdam, John Benjamins: 97-108.

Beard, R. (1998) "Derivation", in A. Spencer and A.M. Zwicky (Eds.) *Handbook of Morphology*. Oxford, Blackwell: 44-65.

Bisang, W. (2004) "Grammaticalization without Coevolution of Form and Meaning: The Case of Tense-Aspect-Modality in East and Mainland Southeast Asia", in W. Bisang, N.P. Himmelmann, and B. Wiemer (Eds.) *What Makes Grammaticalization? A Look from its Fringes and its Components*. Berlin/New York, De Gruyter: 109-138.

Bisetto, A. and S. Scalise (2005) "The Classification of Compounds". *Lingue e Linguaggio*, 4: 319-332.

Booij, G. (2005) "Compounding and Derivation – Evidence for Construction Morphology", in W. Dressler, D. Kastovsky, O.E. Pfeiffer and F. Rainer (Eds.) *Morphology and its Demarcations*. Amsterdam, John Benjamins: 109-132

--- (2007) "Construction Morphology and the Lexicon", in F. Montermini, G. Boyé and N. Hathout (Eds.) *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*. Somerville (MA), Cascadilla Proceedings Project: 34-44.

--- (2010) *Construction Morphology*. Oxford, Oxford University Press.

Ceccagno, A. and B. Basciano (2007) "Compound Headedness in Chinese: An Analysis of Neologisms". *Morphology*, 17: 207-231.

Chao, Y.-R. (1968) *A Grammar of Spoken Chinese*. Berkeley, University of California Press.

Dong, X. 董秀芳 (2004) 汉语的词库与词法. Beijing, Peking University Press.

Dressler, W.U. (2006) “Compound Types”, in G. Libben and G. Jarema (Eds.) *The Representation and Processing of Compound Words*. Oxford, Oxford University Press: 23-44.

Duanmu, S. (1998) “Wordhood in Chinese”, in J.L. Packard (Ed.) *New Approaches to Chinese Word Formation*. Berlin/New York, De Gruyter: 135-195.

Fabb, N. (1998) “Compounding”, in A. Spencer and A.M. Zwicky (Eds.) *Handbook of Morphology*. Oxford, Blackwell: 66-83.

Goldberg, A. (1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, University of Chicago Press.

Grandi, N. and L. Körtvélyessy (2015) “Introduction: Why Evaluative Morphology?”, in N. Grandi and L. Körtvélyessy (Eds.) *The Edinburgh Handbook of Evaluative Morphology*. Edinburgh, Edinburgh University Press: 3-20.

Han, Q. 韩清 (2012) “现代汉语新兴“被X”格式研究”. *Cangsang*, 3: 142-144.

Haspelmath, M. (2010) “Comparative Concepts and Descriptive Categories in Cross-Linguistic Studies”. *Language*, 86(3): 663-687.

Heine, B., U. Claudi and F. Hünnemeyer (1991) *Grammaticalization: A Conceptual Framework*. Chicago, University of Chicago Press.

Himmelmann, N.P. (2014) “Asymmetries in The Prosodic Phrasing of Function Words: Another Look at the Suffixing Preference”. *Language*, 90(4): 927-960.

Hu, Z. 胡孜娴 (2018) “语言接触中英汉语素的融合 ———以类词缀“软”的产生为例”. *Hunan Renwen Keji Xueyuan Xuebao*, 35(4): 99-102.

Huang, J.C.-T. (1984) “Phrase Structure, Lexical Integrity, and Chinese Compounds”. *Journal of the Chinese Language Teachers Association*, 19: 53–78.

Jia, Z. 贾泽林 (2019) “汉语类前缀和类后缀的不对称性分析”. *Hanzi Wenhua*, 241: 66-69.

Jiang, Y. 姜艳艳 and D. Li 李登桥 (2018) “当代新兴“轻X”词族探微”. *Wuyi Daxue Xuebao* (*shehui kexue ban*), 20(2): 75-79.

Kwong, O.Y. and B.K. Tsou (2003) “Categorial fluidity in Chinese and its implications for part-of-speech tagging”, in *Proceedings of* *the 10th conference of the European chapter of the Association for Computational Linguistics*. Budapest, Hungary, April: 115-118. https://aclanthology.org/E03-1081.pdf (visited 2023/02/20).

Lieber, R. (2017) "Derivational Morphology", in M. Aronoff (Ed.) *Oxford Research Encyclopedia of Linguistics*. Oxford, Oxford University Press.

Lieber, R. and P. Štekauer, (2009) "Introduction: Status and Definition of Compounding", in R. Lieber and P. Štekauer (Eds.) *The Oxford Handbook of Compounding*. Oxford, Oxford University Press: 3-18.

Lightfoot, D.J. (2005) "Can the Lexicalization/Grammaticalization Distinction Be Reconciled?". *Studies in Language*, 29(3): 583-615.

Lü, S. 吕叔湘 (1980) 现代汉语八百词. Beijing, Shangwu Yinshuguan.

Ma, Q. 马庆株 (1995) "现代汉语词缀的性质，范围和分类". *Zhongguo Yuyanxuebao*, 6: 101-137.

Masini, F. (1993) *The formation of Modern Chinese Lexicon and its Evolution towards a National Language: The Period from 1840 to 1898*. Berkeley, Project on Linguistic Analysis.

Michaelis, L.A. and K. Lambrecht (1996) "Toward a Construction-Based Theory of Language Function: The Case of Nominal Extraposition". *Language*, 72: 215-247.

Micheli, S. (2020) *La formazione delle parole. Italiano e altre lingue*. Carocci, Roma.

Montermini, F. (2008) *Il lato sinistro della morfologia. La prefissazione in italiano e nelle lingue del mondo*. Milano, Franco Angeli.

Packard, J.L. (2000) *The Morphology of Chinese. A Linguistic and Cognitive Approach*. Cambridge, Cambridge University Press.

Pan, W. 潘文国, B. Ye 叶步青 and Y. Han 韩洋 (2004) 汉语的构词法研究. Shanghai, Huadong Shifan Daxue Chubanshe.

Scalise, S. (1990) *Morfologia e Lessico*. Bologna, Il Mulino.

Scalise, S., A. Bisetto and E. Guevara (2005) "Selection in Compounding and Derivation", in W. Dressler, D. Kastovsky, O.E. Pfeiffer and F. Rainer (Eds.) *Morphology and its Demarcations*. Amsterdam, John Benjamins: 133-150.

Scalise, S. and A. Fábregas (2010) "The Head in Compounding", in S. Scalise and I. Vogel (Eds.) *Crossdisciplinary Issues in Compounding*. Amsterdam, John Benjamins: 109-126.

Sproat, R. and C. Shih (1996) “A Corpus-Based Analysis of Mandarin Nominal Root Compounds”. *Journal of East Asian Linguistics*, 5: 49-71.

Sun, Y. 孙艳 (2000) “现代汉语词缀问题探讨”. *Hebei Shifan Daxue Xuebao*, 23: 55-58.

Ten Hacken, P. (2017) “Compounding in Morphology”, in M. Aronoff (Ed.) *Oxford Research Encyclopedia of Linguistics*. Oxford, Oxford University Press.

Traugott, E.C. (2002) “From Etymology to Historical Pragmatics, in D. Minkova and R. Stockwell (Eds.) *Studies in the History of the English Language*. Berlin/New York, De Gruyter: 19-49.

Xu, Z. 徐志敏 and L. Cai 蔡亮 (2007) “汉语前缀与外来语素的前缀提取”. *Huaibei Meitan Shifan Xueyuan Xuebao*, 6: 133-136.

Yang, Y. 杨一飞 (2007) “新兴三音节准前缀词的性质、功能与发展”. *Yuyan Kexue*, 6(1): 48-57.

Yip, P. (2000) *The Chinese Lexicon. A Comprehensive Survey*. London, Routledge.

Zhao, C. 赵晨 (2018) “汉语类前缀类义指示作用与对外汉语类前缀词表设计——以新兴类前缀为例”. *Changjiang Congkan*, 4: 88-89.

Zwicky, A.M. (1985) “Heads”. *Journal of Linguistics*, 21: 1-29.

MIRATIVE *NǏ KÀN* 你看: AN ANALYSIS AT SYNTAX/ PRAGMATICS INTERFACE

Linda Badan
Ghent University

Yuan Huahung
National Taipei University of Business

1. Introduction

In this paper we offer an analysis at syntax/pragmatics interface of *nǐ kàn* 你看 'you look'. We provide syntactic evidence that *nǐ kàn* in certain contexts is a grammaticalised form that substantially differs from the verb *kàn* 看 'look'. In particular we focus on the use of *nǐ kàn* as intersubjective mirative marker (Tantucci 2021: 117-124; Van Olmen and Tantucci 2022: 167), which expresses the speaker's surprise (*mirativity*, De Lancey 1997) with respect to an event or a situation, as in the following example:

(1)	*Nǐ kàn*	*tā*	*jīntiān*	*xīnqíng*	*duō*	*hǎo!*
	你看	她	今天	心情	多	好!
	you-look	she	today	feel	very	well[1,2]

'**You look** [=surprise] how good she feels today!'

[1] The following abbreviations are used in glossing examples: ASP aspectual particle; CL classifier; DET determination particle 'de' 的; FP final particle; MOD adverbial/complement modifier 'de' 得; OM object marker 'bǎ' 把.

[2] Following the tradition of the generative grammar works, we based our

The novelty of our paper is twofold: methodological and theoretical. We first offer a number of systematic diagnostic tests to demonstrate that *nǐ kàn* as in (1) cannot be considered as a "genuine" verb. We argue that *nǐ kàn* substantially differs from its use as an imperative verb phrase and should be treated as a fixed form characterized by a bleached meaning and discourse related properties. The tests proposed in this paper are valuable syntactic tools to demonstrate that linguistic expressions are used as discourse markers not only with respect to the discourse context but also on the basis of certain syntactic restrictions. From the theoretical viewpoint, on the one hand, we propose that the mandatory presence of the second person singular pronoun is an overt realization of the *nǐ kàn*'s interactional property. We argue that this might be due to the high analyticity of the Chinese language: *nǐ* is "crystalised" together with the verb in a fixed form to underline the necessity of attracting the interlocutor's attention. Indeed, *nǐ kàn* can be also employed as discourse marker to take the turn in a dialogical exchange, just to call the attention of the hearer (Si 2013; Tantucci 2021; Van Olmen and Tantucci 2022):

(2)	*nǐ kàn*	*wǒ*	*bù*	*zhīdào*	*nǐ*	*zài*	*shuō*	*shénme*
	你看	我	不	知道	你	在	说	什么。
	you-look	I	not	know	you	ASP	say	what

'**Look**, I don't know what you are talking about'

On the other hand, we sketch a syntactic analysis within the generative grammar framework developed by Noam Chomsky in the 50's. In particular, we ground our proposal on the syntactic organization of the clausal spine postulated in the Cartographic Approach (Rizzi 1997;

research on retrospective data that we have tested each of them with 20 native speakers of Mandarin Chinese from the Beijing area. We composed a grammaticality judgments questionnaire. Every speaker was requested (i) to read the sentences (inserted within contexts), (ii) judge them as well-formed (grammatical), marginally well-formed (?), ill-formed (unacceptable or ungrammatical), (iii) describe their interpretation(s). For this article, we have then selected only the sentences that have been judged in the same way by all the informants.

Cinque 1999, and subsequent works). We defend the idea that a discourse marker is a linguistic item that undergoes syntactic restrictions and occupies a specific syntactic area in highest portion of an abstract structural hierarchy (see Section 5).

The article is organized as follows: In Section 2, we offer an overview of the main properties of discourse markers in general and we illustrate previous studies on *nǐ kàn*. In Section 3, we systematically illustrate a number of diagnostic tests to show that *nǐ kàn* differs from the verb *kàn* 'look'. In Section 4, we discuss the interpretation of *nǐ kàn* as a mirative marker, and, in Section 5, we offer a syntactic analysis of *nǐ kàn* as a discourse marker that occupies a high structural position in the clausal spine. In Section 6 we discuss some open issues and then, in Section 7, we conclude the paper.

2. Discourse markers: general properties

Discourse markers[3] are an intriguing cross-linguistic phenomenon that has been widely discussed in the linguistic literature from different perspectives (Schiffrin 1987; Fraser 1996, 1999, 2009; Biberauer *et al.* 2008; Biberauer and Sheenan 2011; Degand *et al.* 2013, among many others). Despite the varieties of approaches and terminologies, researchers agree in a number of properties that seem to be common among discourse markers cross-linguistically: Discourse markers (i) can belong to different grammatical categories, such as, for instance, verbs, adjectives, nouns, adverbs, conjunctions, which are grammaticalized in 'fixed' linguistic expressions; (ii) they are a phenomenon typical of spoken language (but not exclusively), which makes it difficult to identify their historical development of their grammaticali-

[3] In the literature, the expressions belonging to this class are defined in several different ways: discourse particles, discourse markers, modal particles, pragmatic markers. The multiple terminology reveals the difficulties in providing a unified classification and related linguistic analysis. For *nǐ kàn* analysed in this paper, we adopt the definition proposed by Bazzanella (2006: 456): "Discourse markers are external to propositional content; they are useful in locating the utterance in an interpersonal and interactive dimension, in connecting and structuring phrasal, inter- and extra-phrasal elements in discourse, and in marking some ongoing cognitive processes and attitudes".

zation path; (iii) they are multifunctional: their interpretation varies on the basis of their structural position and pragmatic context, thus (iv) their meaning is difficult to paraphrase or translate in another language. Moreover, some discourse markers are language specific and do not have a one-to-one correspondent in other languages; (v) they are uttered mainly to express the speaker's attitude and point of view over the proposition (very often in a dialogical exchange with reference to an interlocutor); (vi) they do not add anything to the truth's conditions of a proposition; (v) they are often characterized by a special prosody, i.e. they maybe be followed by a pause that 'separates' them from the rest of the utterance. As for this last property, in this paper we analyse only those cases where there is no audible pause between *nǐ kàn* and the rest of the enunciate. In fact, if a discourse marker is separated by the rest of the clause, it must be analysed as parenthetical, thus endowed with a completely different syntactic structure[4].

2.1 Discourse markers in Chinese

After the pioneering work by Chao (1968: 85), who defines expressions like *gānbēi* 干杯 as "minor sentences", discourse markers in Chinese have been widely studied. For a general overview, see for instance Feng (2008), who offers a typological classification of "conceptual" and "non-conceptual" pragmatic markers on the basis of "their inherent semantic import, rather than their contribution to the host clause" (Feng 2008: 1696). Conceptual markers encode certain conceptual information, expressing speaker's opinion with respect of a situation. These markers are for instance modal adverbial expression (*xìngyùnde shì* 幸运的是 'fortunately') which occupy the highest positions in the architectural spine of adverbs proposed by Cinque (1999). Non-conceptual markers do not express any conceptual messages (as *dànshì* 但是 'but'), but they give "just an indication that the speaker conceives a certain contrast holding between the adjoining propositions explicitly stated or contextually implied" (Feng 2008: 1697). However, in his analysis, Feng excludes *nǐ kàn* and similar forms from the domain of pragmatic markers, claiming that they do not provide any comment

[4] For a syntactic analysis of parenthetical versus non-parenthetical linguistic elements see for instance Giorgi (2011, 2016).

on the proposition, but they just have an interactive function. Differently from Feng, we will illustrate in detail below that *nǐ kàn* in certain contexts is a mirative marker that not only has an interactive function, but also does express the speaker's attitude (the sense of surprise) towards an event or a situation[5].

Other works are instead more focused on specific discourse markers, such as Miracle (1991) who proposes a discourse analysis of *hǎo* 好, *nà*(*me*) 那(么), *dànshì* 但是 and *kěshì* 可是, offering a detailed descriptions of their functions and comparing them to their English counterparts discussed by Schiffrin (1987). Lim (2011) focuses his attention exclusively on *wǒ juédé* 我觉得 'I think' used in an original corpus of natural conversations. Interestingly, Lim shows that *wǒ juédé* cannot be analysed simply as a discourse marker that achieved a level of grammaticalization (see also Huang 2003; Fang 2005). In particular, *wǒ juédé* should be analysed as a subjective speaker-oriented epistemic marker, which also expresses intersubjectivity, being often used "in consideration of addressee-oriented problems" (Lim 2011: 265) in conversational exchanges (see also Traugott 2012; Tantucci 2021; Van Olmen and Tantucci 2022)[6]. With this work, Lin also defends the view that routinized practices, such as stance-talking, are dialogical processes which include both speaker and hearer (Du Bois 2007).

2.1.1 Previous studies on nǐ kàn

Previous studies show that *nǐ kàn* 'you look' is grammaticalized as discourse marker (Zeng 2005; Wang 2006; Si 2013; Wang 2016 among

[5] Feng (2008), instead, analyses *wǒ kàn* 我看 'I see'/'I reckon' (on a par with *wǒ xiǎng* 我想 'I think', *wǒ cāi* (*xiǎng*) 我猜想 'I guess', *wǒ huáiyí* 我怀疑 'I suspect') as an example of epistemic marker, expressing the speaker's commitment to the truth of the proposition, as in (i):

(i) *Wǒ kàn,*	*Liú Xiáng*	*de*	*chénggōng*	*guīgōng*	*yú*	*tā-de*	*jiāoliàn*
我看,	刘翔	的	成功	归功	于	他的	教练
I think	Liu Xiang	DET	success	owe	to	his-DET	coach

'Liu Xiang's success is due to his coach, I think.'
(Feng 2008: 42)

[6] Inspired by Lim's (2011) work, we will get back to such co-occurrence of subjective and intersubjective functions for *nǐ kàn* in Section 4.

others): the semantic meaning of the verb *kàn* indicating the act of looking is bleached and the expression *nǐ kàn* assumes epistemic modality (Zeng 2005) (as in (3)), or introduces a new topic (as in (4)) (Miao and Han 2013; Yin and Li 2011; Si 2013).

(3) *Lǐ: Nǐ kàn zánmen zhè tímù qǐ shénme hǎo ?*
李: **你看** 咱们 这 题目 起 什么 好?
Li you-look we this title name what good
'Li: Look, what do we name this work?' (Zeng 2005: 18 ex 21-22)

(4) *nǐ kàn, yào bù women gǎitiān*
你看, 要 不 我们 改天
you-look want not we another day
yuē-shàng Lǎo Zhào yīkuài huí mǔxiào kàn-kàn
约上 老赵 一块 回 母校 看看。
appoint-up Lao Zhao together return mother-school look-look
'Look, do you think we can go back to the school that we graduated together with Lao Zhao?' Si (2013: 30)

Chen and Park (2006) distinguish two usages of *nǐ kàn*, one derived from *yī nǐ kàn* 依你看 'according to you look' (*nǐ kàn* 1 in (5a)) and one not derived from *yī nǐ kàn* (*nǐ kàn* 2 in (5b)).

(5) a. (*yī*) *nǐ kàn1, Xiǎo Wáng shì-bù-shì gè*
(依) **你看1**, 小王 是不是 个
according to you-look Xiao Wang be-not-be CL
hǎo rén
好 人。
nice person
'According to you, is Xiao Wang a nice person?'

b. (**yī*) *nǐ kàn2, Xiǎo Wáng zhēn shì gè hǎo*
(*依) **你看2** 小王 真 是 个 好
according to you-look Xiao Wang very be CL nice
rén
人!
person
'You look, Xiao Wang is really a nice person.' (Chen and Park 2006: 4)

For Chen and Park (2006), *nǐ kàn*1 takes an interrogative clause, and functions to inquiry the addressee (see also Yin and Li 2011). As a discourse marker, *nǐ kàn* 2 draws the attention of the addressee on the topic introduced by the clause which is a statement or a question. For Chen and Park (2006), *nǐ kàn*1 is an irrealis marker while *nǐ kàn*2 marks a realis situation (respectively in (6a) and (6b)):

(6) a.

nǐ kàn1	*wǒ*	*qù-bù-qù*?
你看1	我	去 不 去?
you-look	I	go-not-go

‘According to you, do I go?’

b.

nǐ kàn2	(*zhè*	*shì*	*hé*	*kǔ*	*ne?*)
你看2	(这	是	何	苦	呢?)
you-look	this	be	what	bitterness	PART

wǒ	*zěn*	*néng*	*bù*	*tóngyì*	*ne?*
我	怎	能	不	同意	呢?
I	how	can	not	agree	PART

‘You look, what did you make you complicated? How I cannot agree with you!’
(Chen and Park 2006: 10)

In a more recent work, Long *et al.* (2019) analyse *nǐ kàn* as a clause-taking imperative predicate pragmaticalized[7] in a discourse marker that attracts the addressee’s visual attention to the content of the following clause. The authors, on the basis of a *Conjoining pathway* (Brinton 2008), argue that the discourse marker *nǐ kàn* has a parenthetical structure, therefore it cannot be analysed as structurally integrated within the following sentence (the matrix clause), but instead it is combined with it (see also Long *et al.* 2019)[8]. However, to our knowledge, the distinction between *nǐ kàn* with and without comma intona-

[7] The discussion about the distinction between pragmaticalization and grammaticalization is vast, and it goes beyond the scope of this paper (for a resonated discussion see for instance Badan 2020).

[8] Specifically, Long *et al.* (2019) argue that *nǐ kàn* grammaticalized from a prosodically separated parenthetical to a prosodically unseparated parenthetical. The authors point out that this process (the *Conjoining pathway*) consists in “the loss of the phonetic gap and no syntactic operation is involved”. (Long *et al.* 2019: 24).

tion is not based on systematic prosodic studies which experimentally demonstrate the mandatory parenthetical status of *nǐ kàn*. Therefore, as mentioned above, in this paper we focus only on the use of *nǐ kàn* pronounced without any audible pause before the following sentence. All the scholars mentioned above also offer some scattered diagnostic tests to isolate *nǐ kàn* as a discourse marker. In the following section, we will list them and add further tests in a systematic way.

3. *Nǐ kàn* is a discourse marker: diagnostics

In this section, we briefly discuss the use of the verb-based discourse marker *nǐ kàn* 'you look' to express unexpectedness, i.e., a sense of surprise towards an event or a situation (Tantucci 2021; Tantucci and Van Olmen 2022)[9], as in (7). In (7) the original meaning of the verb *kàn* 'look' is bleached: *nǐ kàn* does not encode any act of looking, but its function is to express surprise with respect to the fact that Lisi today is feeling very good (in (7a)) or that Lisi's voice is wonderful (in (7b)), contrarily to the speaker's expectation. In both the examples, no visual perception is actually involved.

(7) a.	*nǐ kàn*	*Lǐsì*	*jīntiān*	*xīnqíng*	*duōme*	*hǎo*
	你看	李四	今天	心情	多么	好!
	you-look	Lisi	today	feel	so	good
	'*You look* how good Lisi feels today!'					
b.	*nǐkàn*	*Lǐsì*	*de*	*shēngyīn*	*duōme*	*měimiào*
	你看	李四	的	声音	多么	美妙!
	you-look	Lisi	DET	*voice*	*so*	*wonderful*
	'*You look* how wonderful Lisi's voice is!'					

However, before going into the details of the "surprise reading" (see

[9] More specifically, Van Olmen and Tantucci (2022) offer a study on a usage-based approach to parenthetical *nǐ kàn* (comparing also the parenthetical LOOK in Dutch, English and Italian), showing that *nǐ kàn* is used as attention-getting, mirative and intersubjective pragmatic marker. While Tantucci (2021: Chapter 4), focuses in children's usage of *nǐ kàn*, showing that the multi-functions of *nǐ kàn* as intersubjective, mirative and opinion-elicitation marker emerges at later stages of development than the directive usage of the same form.

Section 4), we first propose a number of diagnostic syntactic tests to demonstrate that *nǐ kàn*, in surprise contexts, must be treated differently from the "genuine" verb *kàn* 'look' under different respects.
(I) As previous studies on grammaticalization have shown (see for instance Hopper and Traugott 2003 and related works), an element that loses its lexical properties turning into a functional element tends to occupy a higher position within the clause. Indeed, *nǐ kàn* (not followed by a pause) does not have a free syntactic distribution, but undergoes a number of restrictions. First, it appears mainly in sentence initial position (cf. examples in (8)). The presence of *nǐ kàn* both in sentence internal and sentence final positions make the clauses ungrammatical or at least infelicitous.

(8) **sentence initial**

a.	*Nǐ kàn*	*tā*	*jīntiān*	*xīnqíng*	*duōme*	*hǎo!*
	你看	她	今天	心情	多么	好!
	you-look	she	today	feel	so	good
	'*Look*, how good she feels today!'					
	sentence internal					
b.	*tā	*nǐkàn*	*jīntiān*	*xīnqíng*	*duōme*	*hǎo!*
	*她	你看	今天	心情	多么	好!
	she	you-look	today	feel	so	good
c.	*tā	*jīntiān*	*nǐkàn*	*xīnqíng*	*duōme*	*hǎo!*
	*她	今天	你看	心情	多么	好!
	she	today	you-look	feel	so	good
d.	*tā	*jīntiān*	*xīnqíng*	*nǐ kàn*	*duōme*	*hǎo*
	*她	今天	心情	你看	多么	好!
	she	today	feel	you-look	so	good
	sentence final					
e.	*tā	*jīntiān*	*xīnqíng*	*duōme*	*hǎo*	*nǐ kàn!*
	*她	今天	心情	多么	好	你看!
	she	today	feel	so	good	you-look

(II) *Nǐ kàn* is a root phenomenon: so it cannot be embedded in subordinated clause (9). In (10), indeed, the presence of *nǐ kàn* is acceptable only if the clause is a quote of a direct discourse.

(9) *suīrán* (**nǐ kàn*) *tā* *kànqǐlái* *nàme* *cōngmíng*,
虽然 (***你看**) 他 看起来 那么 聪明,
although you-look he look-like so clever
dànshì tā *què* *bù* *huì* *shàng* *wǎng*
但是 他 却 不 会 上 网。
but he but not know go internet
'Although (*you look) he looks so clever, but he does not know how to use internet.'

(10) a. *tā* *shuō:* *"nǐ kàn* *nàme* *hǎo* *chī* *a!"*
她 说: 你看 那么 好 吃 啊!
she say you-look so good eat PART
'She said: "look, how delicious it is!"

b. **tā* *shuō* *nǐ kàn* *nàme* *hǎo* *chī* *a!*
*她 说 **你看** 那么 好 吃 啊!
she say you-look so good eat PART

(III) *Nǐ kàn*'s meaning is bleached to the point that it does not add anything to the propositional content of the sentence. Independently of the absence or presence of *nǐ kàn*, the truth values of the proposition remain unchanged (Zeng 2005):

(11) (*nǐ kàn*) *tā* *jīntiān* *xīnqíng* *duōme* *hǎo!*
(**你看**) 她 今天 心情 多么 好!
you-look she today feel so good
'(Look) how good she feels today!'

(IV) Evidences for the invariability of the form *nǐ kàn* as discourse marker is demonstrated by the fact that it cannot be modified. For instance, *nǐ kàn* cannot be negated (as in (12)) (Wang 2006) or be followed by an aspectual marker (as in (13) (Zeng 2005; Si 2013; Wang 2016; Wang 2006). *Nǐ kàn*, indeed, is insensitive to expression of aspect or tense (cf. (14a) with (14b)).

(12) **bié nǐ kàn/ nǐ bié kàn tā jīntiān xīnqíng duōme hǎo!*
***别** **你看** **你别看** 她 今天 心情 多么 好!
don't you-look you don't look she today feel so good

(13) **nǐ kàn le tā jīntiān xīnqíng duōme hǎo!*
***你看** **<u>了</u>** 她 今天 心情 多么 好!
you-look PERF she today feel so good

(14) a. *nǐ kàn zuótiān de huā duō*
你看 昨天 的 花 多
you-look yesterday DET flower so
měi ā!
美 啊！
beautiful PART
Lit. 'Look, how beautiful the yesterday flowers were!'

b. *nǐ kàn jīntiān de huā duō měi ā!*
你看 今天 的 花 多 美 啊！
you-look today DET flower so beautiful PART
Lit. 'Look, how beautiful the today flowers are beautiful!'

Importantly, notice that the second person singular pronoun *nǐ* 你 cannot be separated from *kàn* 看 by any other item. For example, the presence of an adverb is admitted only when *nǐ kàn* is interpreted as genuine verb indicating the act of looking (cf. (15a) with (15b)):

(15) a. *Nǐ xiànzài kàn tā xīnqíng duōme hǎo!*
你 **<u>现在</u>** **看** 她 心情 多么 好!
you now look she feel so good
'You are seeing now how good she feels.'

b. *Hǎo, nǐ jīntiān kàn diànyǐng,*
好, **你** 今天 **看** 电影,
OK you today look movie
dànshì míngtiān nǐ yào xuéxí
但是 明天 你 要 学习。
but tomorrow you must study
'Okay. You see movies today, but you have to study tomorrow.'

(V) *Nǐ kàn* cannot be reduplicated (cf. (16a) with (16b)), it cannot co-occur with auxiliary/modal verbs (17), and it cannot appear in serial verb constructions (18). The sentence in (18) is acceptable only if *nǐ kàn* is interpreted as a genuine verb and not as a discourse marker.

(16) a.

**nǐ*	*kàn-yī-kàn*		*/nǐ kànkàn*		
***你**	**看一看**		/你看看		
you	look-one-look		you look-look		
tā	*jīntiān*	*xīnqíng*	*duōme*		*hǎo!*
她	今天	心情	多么		好!
she	today	feel	so		good

b.

kàn-yī-kàn	*zhè*	*shù*	*huā*
看一看	这	束	花
look-one-look	this	CL	flower

'Look at this bunch of flowers.'

(17)

**nǐ huì*	*kàn*	*tā*	*biàndé*	*duōme*	*cōngmíng*	*ā!*
*你会	**看**	她	变得	多么	聪明	啊!
youwill	look	she	become	so	clever	PART

(18)

nǐ	*qù*	*kàn*	*tā*	*biàndé*	*duōme*	*cōngmíng*	*ā!*
你	去	**看**	她	变得	多么	聪明	啊!
you	go	look	she	become	so	clever	PART

'You go to see how clever she becomes.'

(VI) *Nǐ kàn* cannot be coordinated with any other predicate, while the genuine verb *kàn* can (cf. (19a) with (19b)):

(19) a.

nǐ kàn		*(yǔ/bìngqiě)*	*nǐ*	*tīng*
***你看**		(与/并且)	你	听
you-look		and	you	listen
tā	*de*	*shēngyīn*	*duōme*	*měimiào!*
她	的	声音	多么	美妙!
she	DET	*voice*	*so*	*wonderful*

cf. with:

	nǐ	*tīng*	*tā*	*de*	*shēngyīn*	*duōme*	*měimiào!*
	你	听	她	的	声音	多么	美妙!
	you	listen	she	DET	*voice*	*so*	*wonderful*
	‘Listen, how wonderful her voice!’						
b.	*tāmen*	*kàn*	*bìngqiě*	*sīkǎo*	*le*	*zhè-gè*	*tímù*
	他们	看	并且	思考	了	这个	题目。
	they	look	and	think	ASP	this-CL	subject
	‘They look at this subject and think about it.’						

(VII) *Nǐ kàn* cannot be in interrogative in the form A *bù* A:

(20)	* *nǐ kàn-bù-kàn*	*tā*	*name*	*cōngmíng*	*ā!*
	*你看不看	他	那么	聪明	啊！
	youlook-not-look	he	so	clever	PART

(VIII) *Nǐ kàn* as a discourse marker cannot take a complement (Si 2013):

(21)	**nǐ kàn*	*qīngchǔ*	(/*míngbái*)			*women*
	*你看	清楚	(/明白)			我们
	you-look	clearly				we
	shì fǒu	*xuǎnzé*	*yī-gè*	*xiàn*	*shìdiǎn*	
	是否	选择	一个	县	试点	
	if	choose	one-CL	county	try	
	(Si (2013: 11) slightly modified)					

With the diagnostics tests illustrated so far, we have demonstrated in a systematic way that *nǐ kàn* cannot be considered as a regular pronoun followed by a genuine verb, but it must be regarded as an invariable form whose meaning indicating the act of looking is bleached.

4. Mirative *nǐ kàn*

Zeng (2005) defines *nǐ kàn* as a subjectivity and epistemic marker (Si 2016: 387). On the one hand, *nǐ kàn* is a subjectivity marker in the sense that it reflects the speaker’s opinion: “The term subjectivity refers to the way in which natural languages, in their structure and their normal manner of operation, provide for the locutionary agent's expression of himself and his own attitudes and beliefs" (Lyons 1982: 102). On the other hand, the interpretation of *nǐ kàn* includes an epistemic component, in the sense that it expresses the speaker's degree of commitment to the truth-value of the proposition (Lyons 1977). However, as mentioned above, we focus on the usage of discourse marker *nǐ kàn* in contexts expressing surprise. In such contexts, the presence of *nǐ kàn* conveys not only subjectivity and epistemicity, but also evidentiality, which is a grammatical category used to indicate the nature of evidence for a given statement[10].The strict relation (and overlapping) between the concept of epistemicity and evidentiality is well-known and has been widely discussed in the literature (Palmer 1986; Kiefer 1994; de Haan 1999, 2001, 2005; Aikhenvald 2004)[11]. As Gonzáles *et al.* (2017: 68) notice, in fact, "speakers often make judgments on the basis of perceptual, reported, or inferred evidence, thus the relationship between epistemicity and evidentiality is often close and difficult to demarcate". However, *nǐ kàn* in the examples above conveys an additional aspect, which is unexpectedness. In contexts where the speaker utters *nǐ kàn* to express a sense of surprise, we follow Tantucci (2021) and Van Olmen and Tantucci (2022) considering the *nǐ kàn* is a mirative discourse marker. *Mirativity*, indeed, is a pragmatic phenomenon related to the expression of unexpected information (De Lancey 1997, 2001; Friedman 1980; Peterson 2010, among many others). The concept of mirativity combines the evaluation (thus epistemicity) by the speaker, that is the sense of surprise with respect to a certain fact or situation (thus the subjectivity/the speaker's point of view) with the evidentiality component, whose primary meaning is the source of information. Notice, indeed, that *nǐ kàn* is sensitive to exclamative

[10] It is possible to distinguish two types of evidentiality: direct and indirect evidentiality. With *direct evidentiality* we intend that the assertion is based on perceptual evidence; while with *indirect evidentiality*, we intend that the assertion is based on inference or hearsay. For a detailed overview of the concept of evidentiality and its realization in syntax cross-linguistically, see for instance Rooryck (2001a, 2001b), Aikhenvald and Dixon (2003) and the references cited there.

[11] A discussion about the link or overlapping of the two concepts, evidentiality and epistemicity, goes beyond the aims of this paper.

illocutionary force: it always requires an exclamation or a sentence that expresses surprise like a rhetorical question:

(22) a. exclamative: 你看 = mirative DM

nǐ kàn	*zhè tiānkōng*	*name*	*lán*	*ā!*
你看	这 天空	那么	蓝	啊!
you-look	this sky	so	blue	PART

'Look this sky is so blue!'

b. question: 你看 = NO mirative DM

nǐ kàn	*zhè*	*tiānkōng*	*lán-bù-lán?*
你看	这	天空	蓝不蓝 ?
you-look	this	sky	blue-not-blue

'Look if this sky is blue or not?

c. declarative: 你看= NO mirative DM

nǐ kàn	*tiānkōng*	*hěn*	*hēi*
你看	天空	很	黑。
you-look	sky	very	dark

'You see this sky is very dark'

As mentioned above, the mirative marker *nǐ kàn* conveys subjectivity, in the sense that it expresses the speaker's point of view. However, the function of *nǐ kàn* is not limited to this aspect. Uttering *nǐ kàn* in an exclamative clause, the speaker also aims at attracting the hearer's attention, a function that can be define as "intersubjectivity" in the sense of Traugott (2003) further developed by Fitzmaurice (2004). Fitzmaurice (2004) suggests that discourse markers like *you know*, *you see* etc. underwent a shift from intersubjective to "interactive": with *nǐ kàn*, indeed, the speaker grabs the interlocutor's attention and keeps the conversation going on (see Tantucci 2021; Van Olmen and Tantucci 2022)[12].

Interestingly, we observe that such interactive function of *nǐ kàn* seems to be morphologically explicit. Despite the fact that Chinese is a pro-drop language and *kàn* can also be used without pronoun and so be

[12] Our description of *nǐ kàn* as endowed with an intersubjective function is also in line with Liu (2006) who analyses *nǐ zhīdào* 你知道 used as discourse marker in dialogical exchanges.

interpreted as a simple imperative, with *nǐ kàn* in surprise contexts the presence of the pronoun *nǐ* 你 is mandatory:

(23)	*(*nǐ*)	*kàn*	*tā*	*jīntiān*	*xīnqíng*	*duōme*	*hǎo!*
	*(你)	看	她	今天	心情	多么	好!
	you	look	she	today	feel	so	good

As illustrated in detail above, the presence of the personal pronoun is fundamental to distinguish the grammatical status of *nǐ kàn* with respect to the simple verb *kàn* (see also Feng 2008). Moreover, in order to maintain the interpretation expressing the speaker's surprise, *nǐ* cannot be substituted by any other pronouns (cf. Zeng 2005):

(24)	**tā/tāmen/nǐmen/women*	*kàn*	*tā*	*jīntiān*	*xīnqíng*
	*他/他们/你们/我们	看	她	今天	心情
	he they you we	look	she	today	feel
	duōme	*hǎo!*			
	多么	好!			
	so	good			

In line with Feng (2008) (see also Lim 2011; Liu 2006) we argue that the mandatory presence of the second person singular pronoun in *nǐ kàn* is an overt morphological realization of the interactional aspect: *nǐ* is grammaticalized together with the verb to underline the requirement of attracting the interlocutor's attention. As (Feng 2008: 1696) notices "a second-person form serves to explicitly announce what sort of effect the speaker intends to have on the hearer" or "to oil the wheels of conversational exchange" (Fitzmaurice 2004: 428)[13].

We also propose that the overt realization of the second person pronoun is probably due to the high level of syntactic analyticity

[13] Other examples indicating the interactive discourse function indexed by the second-person pronoun are given by Feng 2008: 1696): "(*nǐ*) *xiǎng* (你) 想 'you think', (*nǐ*) *shuōshuō kàn* (你) 说说看 'tell me/us', (*nǐ*) *xiǎngxiǎng* (你) 想想 'you think', (*wǒ*) *jǐnggào nǐ* (我) 警告你 'I warn you', (*nǐ*) *tīng wǒ shuō* (你) 听我说 'listen to me', (*wǒ) fengquàn nǐ* (我) 奉劝你 'I advise you', (*wǒ*) *wèn nǐ* (我) 问你 'I ask you'".

of Chinese. For analytic language, we intend a language that expresses grammatical relationships mainly by functional elements (such as particles etc.) and word order (syntactic constructions) as opposed to the use of inflection (inflectional morphemes)[14]. If we compare *nǐ kàn* with the corresponding discourse markers based on the verb indicating the act of looking in other pro-drop languages like Spanish or Italian, in Spanish the second singular pronoun is never overtly realized (25a), while in Italian it is possible but not mandatory (25b):

(25) a. *¡(*Tu) mira que hablar así de nuestro hijo!*
you look that speak so of our son
"I am surprised that you, talking that way about our son!"

b. *Ma (tu) guarda che casino ha combinato!*
but you look that mess has done
"Look what a mess he has done!"

While the function of attracting the interlocutor's attention is optional in exclamations (a speaker can utter an exclamation in a self-talk, without the necessity to address it to an interlocutor (Holmberg 2010)), in other contexts, function of calling the hearer's attention is the main function of *nǐ kàn*. *Nǐ kàn*, indeed, can be also used as discourse marker to take the turn in a dialogical exchange (Si 2013), just to call the attention of the hearer (26), or to start a conversation (27):

(26) *nǐ kàn wǒ bù zhīdào nǐ zài shuō shénme*
你看 我 不 知道 你 在 说 什么
you-look I not know you ASP talk what
'**Look** I don't know what you are talking about'

[14] For a detailed discussion about the syntactic features to characterize Modern Chinese as an analytic language see Huang (2015).

(27) *nǐ kàn* *zhè* *shíjiān* *yī-tiāntiān* *guò* *dé*
你看, 这 时间 一天天 过 得
you-look this time one-dayday ASP MOD
fēi *kuài* *ā !*
飞 快 啊!
fly fast PART
'You Look (=surprisely), how the time is flying fast, day by day.'
(Si 2013:387)

For instance, the main employ of *nǐ kàn* in (28a), is attracting the hearer's attention (Cao 2010; Yin and Li 2011; Li and Yin 2011, among others) as demonstrated by the fact that the change of the second person pronoun to a first person pronoun makes the sentence unacceptable (28b):

(28) a. *nǐ kàn* *wǒ* *bù* *zhīdào* *nǐ* *zài*
你看 我 不 知道 你 在
you-look I not know you ASP
shuō *shénme*
说 什么
talk what
'**Look** I don't know what you are talking about'
b. **wǒ* *kàn* *wǒ* *bù* *zhīdào* *nǐ* *zài* *shuō*
*我 看 我 不 知道 你 在 说
I look I not know you ASP talk
shénme
什么
what

All in all, we then suggest that *nǐ kàn* in contexts expressing a sense of surprise, is a mirative discourse marker, a definition the includes its multi-faceted properties: the expression of subjectivity, the epistemic and evidential features, and (optionally) the attraction of the attention of an interlocutor.

5. The syntax of *nǐ kàn*

In this section we propose a syntactic analysis of *nǐ kàn* within the generative grammar framework (Chomsky 1957 and subsequent works), a theory of grammar that postulates underlying abstract forms that are the basis of all linguistic expressions. In particular, we adopt the Cartographic approach (Rizzi 1997; Cinque 1999),[15] which helps us to clearly identify the structural position of *nǐ kàn* used as discourse marker within the syntax of the clause. One of the main assumptions underlying the Cartographic framework is the correspondence between interpretation, informational value and syntactic position of a phrase. Among the objectives of Cartography, there is the identification of the various syntactic positions, with their interpretations (and, ideally, their prosodic outlines), as well as their mapping in a hierarchical order within the sentence. According to the Cartographic project, the sentences are structured in three main levels, which are organized in sub-units. The lowest level consists of the verb and its arguments and is referred to as the domain of verb phrase (VP). The VP is the syntactic area where the elements of the argument structure are encoded, i.e. the arguments consisting in the participants to the event or status expressed by the predicate. The arguments are typically realized through grammatical functions such as, for instance, subject and object. Immediately above the VP, there is the domain that represents the inflection of the verb (IP), where the temporal, number, person and structural case properties are realized. Then, above the IP, there is an area called Left Periphery (LP), which is a syntactic space that connects the sentence to the domain of the discourse:

(29) $[_{\text{Left Periphery}}$... Topic, Focus $[_{\text{Inflectional Phrase}}$...Tense, Number, Person, Structural Case $[_{\text{Verb Phrase}}$ ...Verb+its arguments]]]]

The LP encodes all the elements that appear on the left of the subject (hence its definition) and it is the syntactic domain where the information structure is realized. The information structure (term introduced by Halliday (1967) to denote the division of the discourse into information units) has been explored by a number of scholars within different linguistic fields (Chafe 1976; Prince 1981; Lambrecht

[15] For a detailed discussion and presentation of the Cartographic approach see for instance Cinque and Rizzi (2012), Shlonsky and Bocci (2019), Shlonsky (2010).

1994; Krifka 2008; Féry and Krifka 2008; Zimmermann and Féry 2010) and, among many definitions, can be conceived as "packaging of information that meets the immediate communicative needs of the interlocutors, i.e. the techniques that optimize the form of the message with the goal that it be well understood by the addressee in the current attentional state" (Féry and Krifka 2008: 123). In other words, with the information structure we intend the pragmatic criteria that regulate the information of the enunciation, which is commonly considered divided in old and new information. In syntax, an example of an item encoding old information is called topic (in (30)), while an element interpreted as new information is focus (as in (31)). Crucially the interpretation of the topic and focus structures is strictly depending on the context or the immediately previous discourse.

(30) A:	*nǐ*	*bǎ*	*wǒ*	*de*	*kèběn*		*fàng*
	你	把	我	的	课本		放
	you	OM	I	DET	textbook		put
	dào	*nǎ*	*le?*				
	到	哪	了?				
	arrive	where	PART				
	'Where did you put my textbook?'						
B:	[*nǐ*	*de*	*kèběn*		*a*$]_{Topic}$,		
	你	的	课本		啊,		
	you	DET	textbook		PART		
	wǒ	*fang*	*dào*	*nǐ*	*shūbāo*	*lǐ*	*le*
	我	放	到	你	书包	里	了
	I	put	arrive	you	bag	in	PART
	'Your textbook, I put it in the bag'						

(31)	*Lǐsì*	*nàme*	*hǎo*	*de*	*xuéshēng*	*ā!*
	李四	那么	好	的	学生	啊!
	Lisi	so	good	DET	student	PART
	$[_{IP}$*lián*	*nà-běn*	*shū*$]_{Focus}$	*tā*	*dōu*	*yǐjīng*
	$[_{IP}$连	那本	书]	他	都	已经
	even	that-CL	book	he	all	already
	kàn-wán	*le*				

看完 了
read-finish PART
'Lisi, such a good student! He read even that book.'

As we have shown above, contrarily to what is commonly claimed in the literature for discourse markers in general, the position of the mirative *nǐ kàn* is not free within the clause, but it undergoes several syntactic restrictions. *Nǐ kàn* interpreted as a mirative marker, must appear exclusively at the very beginning of the clause, as illustrated in the examples (8). The fact that *nǐ kàn* mandatorily appears on the left of the subject is a clear indication that it occupies a syntactic position in the Left Periphery (LP) of the clause. This is quite expected, since the mirative *nǐ kàn* is a linguistic item endowed with discourse related properties (as to topic and focus) which are generally encoded in the LP. However, in the formal linguistic literature, it has been widely demonstrated that elements such as discourse markers and vocatives, which at first sight seem to be syntactically free from the rest of the sentence, are encoded in a higher position with respect to the LP. In particular, we analyse then *nǐ kàn* as a discourse marker in terms of syntacticization (Haegeman and Hill 2013 among others), intended as a process of fixation both at the pragmatic and the syntactic level of grammar, "by which the interpretation and syntax of markers become increasingly fixated (Abraham 1991) to reflect typical interpretative and formal properties of speech acts" (Larrivée and Poletto 2018: 46). On the basis of comparable examples in a number of typologically different languages, scholars have postulated the existence of a syntactic area which is structurally higher than the LP: the Speech Act domain (Speas and Tenny 2003; Hill 2007; Giorgi 2010; Coniglio and Zegrean 2010, 2012; Haegeman and Hill 2013). The Speech Act domain is conceived as the highest portion of the syntactic clausal spine that encodes the performative aspect of illocutionary elements. Therefore, it is the domain that hosts the functional elements, such as discourse markers, that encode the codification of the speaker's attitude, his/her commitment towards the content of utterances and his/her relation towards the addressee.

(32) [**Speech Act**...[Left Periphery...[InflectionalPhrase...[Verb Phrase ...]]]]

We propose here the application of the same analysis for the mirative *nĭ kàn*. Remind that topic and focus are located in the LP. If *nĭ kàn* occupies a position in the Speech Act domain, it should then occur on the left of topic and focus (which are in the LP), that is in a structural position hierarchically higher than the LP, as it turns out in the examples (33)-(34): When *nĭ kàn* cooccur with a topic (cf (33a) with (33b)) or with a focus (cf. (34a) with (34b)) it must be always in a higher position, otherwise the sentence is (at least) infelicitous:

(33) a.

Nĭ kàn	*nà-gè*	*gūniáng,*	*tā*	*zěnme*	*name*	*cōngmíng!*
你看	那个	姑娘	她	怎么	那么	聪明!
you-look	that-CL	girl	she	how	so	clever

‘Look (=surprisely,) how that girl is so clever!’

b.

**Nà-gè*	*gūniáng*	*nĭ kàn,*	*tā*	*zěnme*	*name*	*cōngmíng!*
*那个	姑娘	你看	她	怎么	那么	聪明！
that-CL	girl	you-look	she	how	so	clever

(34) a.

nĭ kàn		*lián*	*zhè-běn*	*shū*
你看		连	这本	书
you-look		even	this-CL	book
tā	*dōu*	*yĭjīng*	*kàn-wán*	*le.*
他	都	已经	看完	了。
he	all	already	read-finish	PART

‘You look, even this book, he has finished reading it.’

b.

* *lián*	*zhè-běn*		*shū*	*nĭ kàn*
*连	这本		书	你看
even	this-CL		book	you-look
tā	*dōu*	*yĭjīng*	*kàn-wán*	*le.*
他	都	已经	看完	了
he	all	already	read-finish	PART

We would like to point out that differently from topic and focus, elements such as discourse markers are not part of the information structure, but are linguistic expressions that express the speaker’s point of view and are often characterized by performative properties linked to the illocutionary force of the clause (see Section 3). Additionally, notice that *nǐ kàn* shows a syntactic behaviour that is quite different with respect to adverbs although it can be interpreted as an adverbial "surprisingly". For instance, the position of the adverbial *chūhūyìliào-de*出乎意料地 'unexpectedly' is much freer. Differently from *nǐ kàn*, *chūhūyìliào-de* can appear lower than the subject (35a), or on the right of a topicalized element (35b) while, as we have seen above, *nǐ kàn* cannot.

(35) a. *nà-gè* *gūniáng* *chūhūyìliào-de* *zěnme* *name* *cōngmíng*
那个 姑娘 出乎意料地 怎么 那么 聪明
that-CL girl unexpectedly how so clever
ā!
啊!
PART
Unexpectedly, how clever that girl is!

b. ?*nà-gè* *gūniáng* *chūhūyìliào-de*
?那个 姑娘 出乎意料地
that-CL girl unexpectedly
tā *zěnme* *name* *cōngmíng* *ā!*
她 怎么 那么 聪明 啊!
she how so clever PART
'How that girl, she is so clever, unexpectedly!'

Moreover, even if they have a similar interpretation *nǐ kàn* and *chūhūyìliào-de* can cooccur. In that case, *nǐ kàn* must always be in a higher position:

(36) a. *nǐ kàn* *chūhūyìliào-de*
你看 出乎意料地
you-look unexpectedly
nà-gè *gūniáng* *zěnme* *name* *cōngmíng* *ā!*
那个 姑娘 怎么 那么 聪明 啊!
that-CL girl how so clever PART
'Look (=surprisingly), unexpectedly, how clever that girl is!'

b. **chūhūyìliào-de*			*nǐ kàn*		
*出乎意料地			你看		
unexpectedly			you-look		
nà-gè	*gūniáng*	*zěnme*	*name*	*cōngmíng*	*ā!*
那个	姑娘	怎么	那么	聪明	啊!
that-CL	girl	how	so	clever	PART

On the basis of the observations presented above, we argue that the final syntactic structures of the different elements in the distinct areas of the clausal spine is the following:

(37) $[_{\text{Speech Act}}$...Discourse markers $[_{\text{Left Periphery}}$... Topic, Focus $[_{\text{Inflectional Phrase}}$...Tense, Number, Person, Structural Case $[_{\text{Verb Phrase}}$ ...Verb+its arguments]]]]

6. Open issues

The analysis of *nǐ kàn* as discourse marker offered in this paper is far from being exhaustive. With this article, we focus on *nǐ kàn* as discourse marker conveying mirativity, however, there are several issues to take into account in order to determine the precise syntactic position in relation with other discourse-related elements. For instance, it would be interesting to explore the interaction of *nǐ kàn* with the final particles conveying the speaker's attitude. As demonstrated in the literature, final particles in Chinese can be categorized in two groups: particles that indicate the sentence type (like *ma* 吗 for interrogatives) and attitude particles (Paul 2014) that express the speaker's stance towards an event or a situation (as for instance *bei* 呗 that means 'no need to explain'). The co-occurrence of the discourse marker *nǐ kàn* and one of those attitude particles could yield to different pragmatic interpretations. Such a cooccurrence may be worth to be described in detail and it may open up issues regarding the syntactic positions of discourse markers and their interaction with attitude particles. In the formal linguistics literature, in fact, it has been shown that final particles and attitude particles merge in a hierarchical structure in the beginning of the clause, with the subsequent movement of the entire sentence on their left. Living aside the technicalities linked to the syntactic movement, if the attitude particles are in a domain which can be defined as Speech Act, why do discourse markers merge and stay in the same domain, without the need of any movement of the rest of the clause? Are we dealing with the same domain? Does the difference lie on the fact that particles are syntactic heads, while discourse markers are phrase (XP)? What is the syntactic derivation of the co-occurrence of discourse markers and final particles?

Another point of interest concerns the prosodic contour of *nǐ kàn*. As mentioned above, several studies claim that *nǐ kàn* is followed by a pause, grounding their analyses on the difference of *nǐ kàn* with or without the comma intonation (see for instance Long *et al*. 2019). However, the features of the comma intonation of *nǐ kàn* should be investigated with a detailed experimental analysis of its prosodic properties with respect to the following sentence.

All the issues raised above are not trivial and are worth to be investigated. We will leave them open for future research.

7. Conclusions

In this paper, we offered syntactic evidences to show that *nǐ kàn* can be used as a fixed linguistic expression defined as discourse marker. In particular, we proposed a set of tests to identify *nǐ kàn* as an intersubjective mirative marker on the basis of a number of syntactic criteria. These tests can be used as valuable diagnostics for future works on discourse markers in different languages. We have also discussed the obligatory presence of *nǐ* within the debate regarding the definition of Chinese as an analytic language. We have then offered a syntactic analysis within the Cartographic approach, showing that *nǐ kàn* occupies a syntactic position in the Speech Act domain above the Left Periphery. This proposal inserts Chinese discourse markers within a theoretical approach that identifies syntactic positions and their correspondent interpretations within a hierarchy of syntactic functional projections that encode performative and illocutionary features.

References

Aikhenvald, A.Y. (2004) *Evidentiality*. Oxford, Oxford University Press.

Aikhenvald, A.Y. and R.M.W. Dixon (Eds.) (2003) "Studies in Evidentiality". *Studies in Language*, 31(2): 474-478.

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) "Segnali discorsivi e sviluppi conversazionali", in F. Albano Leoni and R. Giordano (Eds.) *Italiano parlato. Analisi di un dialogo*. Napoli, Liguori: 137-157.

Biberauer, T., A. Holmberg, I. Roberts and M. Sheehan (Eds.) (2008) *The Final-Over-Final Constraint*. Cambridge (MA), MIT Press.

Biberauer, T. and M. Sheehan (2011) "Introduction: Particles through a Modern Syntactic Lens". *The Linguistic Review* (Special Edition on Particles in Syntax), 28: 387-410.

Cao, X. 曹秀玲 (2010) "从主谓结构到话语标记——"我/你V"的语法化及相关问题". *Hanyu Xuexi*, 5: 38-50.

Chafe, W.L. (1976) Givenness, Contrastiveness, Definiteness, Subjects and Topics, in C.N. Li (Ed.) *Subject and Topic*. New York, Academic Press: 27-55.

Chao, Y.-R. (1968) *A Grammar of Spoken Chinese*. Berkeley, University of California Press.

Chen, Z.陈振宇, and M.-X. Park 朴珉秀 (2006) "话语标记"你看""我看"与现时情态". *Yuyan Kexue*, 5(2): 3-13.

Chomsky, N. (1957), *Syntactic Structures*. The Hague/Paris, Mouton.

Cinque, G. (1999) *Adverbs and Functional Projections: A Cross-Linguistic Perspective*. Oxford Studies in Comparative Syntax. New York, Oxford University Press.

Cinque, G. and L. Rizzi (2012) "The Cartography of Syntactic Structures", in B. Heine and H. Narrog (Eds.) *The Oxford Handbook of Linguistic Analysis*. Oxford, Oxford University Press: 51-66.

Coniglio, M. and I. Zegrean (2010) "Splitting up Force, Evidence from Discourse Particles", in L. Brugè (Ed.) *University of Venice Working Papers in Linguistics*, 20: 7-34. Venezia, Dipartimento di Scienze del Linguaggio. Università Ca' Foscari.

--- (2012) "Splitting up force", in L. Haegeman, L. Aelbrecht and R.

Nye (Eds.) *Main Clause Phenomena: New Horizons*. Amsterdam, John Benjamins: 190-229.

Degand, L., B. Cornillie and P. Pietrandrea (Eds.) (2013) *Discourse Markers and Modal Particles: Categorization and description*. Amsterdam, John Benjamins.

De Haan, F. (1999) "Evidentiality and Epistemic Modality: Setting boundaries". *Southwest Journal of Linguistics*, 18: 83-101.

--- (2001) "The Relation between Modality and Evidentiality", in R. Müller and M. Reis (Eds.) *Modalität und Modalverben im Deutschen. Linguistische Berichte, Sonderheft* 9. Hamburg, H. Buske: 201-216.

--- (2005) "Encoding Speaker Perspective: Evidentials", in Z. Frajzyngier and D. Rood (Eds.) *Linguistic Diversity and Language Theories*. Amsterdam, Benjamins Publishing Company: 379-417.

De Lancey, S. (1997) "Mirativity: The Grammatical Marking of Unexpected Information". *Linguistic Typology*, 1(1): 33-52.

De Lancey, S. (2001) "The Mirative and Evidentiality". *Journal of Pragmatics*, 33(3), 369-382.

Du Bois, J.W. (2007) "The Stance Triangle", in R. Englebretson (Ed.) *Stancetaking in Discourse Subjectivity, Evaluation, Interaction*. Amsterdam, Benjamins Publishing Company: 139-182.

Fang, M. 方梅 (2005) "认证义谓动兵词的虚化". *Zhongguo Yuwen*, 6: 495-574.

Feng, G.W. (2008) "Pragmatic Markers in Chinese". *Journal of Pragmatics*, 40: 1687-1718.

Féry, C. and M. Krifka (2008) "Information Structure: Notional Distinctions, Ways of Expression", in P. van Sterkenburg (Ed.) *Unity and Diversity of Languages*. Amsterdam, John Benjamins: 123-136.

Fitzmaurice, S. (2004) "Subjectivity, Intersubjectivity and the Historical Construction of Interlocutor Stance: From stance Markers to Discourse Markers". *Discourse Studies*, 6(4): 427-448.

Fraser, B. (1996) "Pragmatic Markers". *Pragmatics*, 6: 167-190.

--- (1999) "What are Discourse Markers?". *Journal of Pragmatics*, 31: 931-952.

--- (2009) "An Account of Discourse Markers". *International Review of Pragmatics*, 1(2): 293-320.

Friedman, V. (1980) "The Study of Balkan Admirativity: Its History and Development". *Balkanistica*, 6: 7-30.

Giorgi, A. (2010) *About the Speaker: Towards a Syntax of Indexicality*. Oxford, Oxford University Press.

--- (2011) "The Syntax of Commas: An Analysis of Two Types of Parentheticals", in M.-L. Gao (Ed.) *Proceedings of GLOW in Asia VIII: Universals and Variation* 1. Beijing, Beijing University Press: 165-169.

--- (2016) "Integrated Parentheticals in Quotations and Free Indirect Discourse", in A. Capone, F. Kiefer and F. Lo Piparo (Eds.) *Indirect Reports and Pragmatics* 5. Switzerland, Springer International Publishing: 471-488.

González-Rodríguez, R., P. Roseano, J. Borràs-Comes and P. Prieto (2017) "Epistemic and Evidential Marking in Discourse: Effects of Register and Debatability". *Lingua*, 186-187: 68-87.

Haegeman, L. and V. Hill (2013) "The Syntacticization of Discourse", in R. Folli, C. Sevdali and R. Truswell (Eds.) *Syntax and its Limits*. Oxford, Oxford University Press: 370-390.

Halliday, M.A.K. (1967) "Notes on Transitivity and Theme in English: Part 2". *Journal of Linguistics*, 3(2): 199-244.

Hill, V. (2007) "Vocatives and the Pragmatics-Syntax Interface". *Lingua*, 117(12): 2077-2105.

Holmberg, A. (2010) "How to Refer to Yourself When Talking to Yourself". *New Castle Working Papers in Linguistics*, 16: 57-65.

Huang, C.-T.J. (2015) "On Syntactic Analyticity and Parametric Theory2, in A. Li, A. Simpson and W.-T.D. Tsai (Eds.) *Chinese Syntax in Cross-Linguistic Perspective*. Oxford, Oxford University Press, 1-48.

Huang, S.F. (2003) "Doubts about Complementation: A Functionalist Analysis". *Language and Linguistics*, 4(2): 429-455.

Kiefer, F. (1994) "Modality", in R.E. Asher (Ed.) *The Encyclopedia of Language and Linguistics*. Oxford, Pergamon Press: 2515-2520.

Krifka, M. (2008) "Basic Notions of Information Structure". *Acta Linguistica Hungarica*, 55(3): 243-276.

Lambrecht, K. (1994) *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge, Cambridge University Press.

Larrivée, P. and C. Poletto (2018) "Mapping the Syntacticisation of

Discourse: The Case of Sentential Particles", UCD *Working Papers in Italian Studies*, vol. 1. University College Dublin, Ireland.

Li, Q. 李君 and S.L. Yin 殷树林 (2011) "说提醒标记'你看'". *Qiushi Xuekan*, 38(3): 119-123.

Lim, N.E. (2011) "From Subjectivity to Intersubjectivity: Epistemic Marker *Wo Juede* in Chinese", in Y. Xiao, L. Tao and H.L. Soh (Eds.) *Current Issues in Chinese Linguistics*. Newcastle upon Tyne, Cambridge Scholars Publishing: 265-301.

Liu, L. 劉麗艷 (2006) "话语标记你知道". *Zhongguo Yuwen*, 5: 423-432.

Long, H.-P., B. Heine, G.-J. Ruan and M.-Y. Wu (2018) "The Grammaticalizational Relation between Two Modern Chinese wo xiang 'I think' Constructions". *Language Sciences*, 66: 212-225.

Long, H., X. Xu, M. Wu and F. Ursini (2019) "Formation of the Modern Chinese Clause-Taking Imperative ni kan 'you see': A Conjoining Pathway Account". *Lingua*, 232: 1-24.

Lyons, J. (1977) *Semantics*, volume 1&2. Cambridge, Cambridge University Press.

--- (1982) "Deixis and Subjectivity: Loquor, Ergo Sum?", in R.J. Jarvella, W. Klein (Eds.) *Speech, Place, and Action: Studies in Deixis and Related Topics*. London/New York, John Wiley: 101-124.

Miao, L.苗丽 and L. Han 韩蕾 (2013) "汉语口语话语标记"你看"的语用功能". *Yuwen Jianshe*, (3): 22-23.

Miracle, C.W. (1991) *Discourse Markers in Mandarin Chinese*. PhD Dissertation, The Ohio State University.

Palmer, F.R. (1986) *Mood and Modality*. Cambridge, Cambridge University Press.

Paul, W. (2014) "Why Particles are not Particular: Sentence-Final Particles in Chinese as Heads of Split CP". *Studia Linguistica*, 68(1): 77-115.

Peterson, T.R.G. (2010) *Epistemic Modality and Evidentiality in Gitksan at the Semantics-Pragmatics Interface*. PhD Dissertation. University of British Columbia, Vancouver.

Prince, E.F. (1981) "Towards a Taxonomy of Given-New Information", in P. Cole (Ed.) *Radical pragmatics*. New York, Academic Press: 223-256.

Rizzi, L. (1997) "The Fine Structure of the Left Periphery", in L. Haegeman (Ed.) *Elements of Grammar*. Netherlands, Springer: 281-337.

Rooryck, J. (2001a) "Evidentiality. Part 1". *Glot International*, 5(4): 125-133.

--- (2001b) "Evidentiality. Part 2". *Glot International*, 5(5): 161-168.

Schiffrin, D. (1987) *Discourse Markers*. Cambridge, Cambridge University Press.

Shlonsky, U. (2010) "The Cartographic Enterprise in Syntax". *Language and Linguistics Compass*, 4(6): 417-429.

Shlonsky, U. and G. Bocci (2019) "Syntactic Cartography", in *Oxford Research Encyclopedia of Linguistics*.

Si, L. 司罗红 (2013) "'你看'的语用功能及形成机制". *Huayin Shifan Xueyuan Xuebao* (*zhexue shehuikexue ban*), 3(38): 385-389.

Speas, P. and C. Tenny (2003) "Configurational Properties of Point of View Roles", in A.M. Di Sciullo (Ed.) *Asymmetry in Grammar. Volume 1: Syntax and Semantics*. Amsterdam, John Benjamins: 315-344.

Tantucci, V. (2021) *Language and Social Minds: The Semantics and Pragmatics of Intersubjectivity.* Cambridge, Cambridge University Press.

Traugott, E.C. (1995) *The Role of the Development of Discourse Markers in a Theory of Grammaticalization*. Paper presented at ICHL XII, Manchester.

--- (2003) "From Subjectification to Intersubjectification", in R. Hickey (Ed.) *Motives for Language Change*. Cambridge, Cambridge University Press: 124-140.

--- (2012) "Intersubjectification and Clause Periphery". *English Text Construction*, 5(1): 7-28.

Van Olmen, D. and V. Tantucci (2022) "Getting Attention in Different Languages: A Usage-Based Approach to Parenthetical Look in Chinese, Dutch, English, and Italian". *Intercultural Pragmatics*, 19(2): 141-181.

Wang, Y. 王媛媛 (2006) "普通話中'你看'變體的多角度考察". *Xuzhou Shifan Daxue Xuebao*, 32(1): 67-71.

Wang, T. 王婷婷 (2016) "'你看吧'与'你看'比较研究". *Xiandai Yuwen*, 7: 111-112.

Wei, X. 魏兴 and Q. Zheng 郑群 (2013) "西方语法化理论视角下对汉语话语标记'你看'的分析". *Waiguo Yuwen* (*shuangyuekan*), 5: 80-86.

Yin, S. 殷树林 and Q. Li, 李君 (2011) "说征询标记'你看'". *Langfang Shifan Xueyuan Xuebao* (*shehuikexue ban*), 27(3): 32-34.
Zeng, L. 曾立英 (2005) "'我看'與'你看'的主觀化". *Hanyu Xuexi*, 2: 15-22.
Zimmermann, M. and C. Féry (Eds.) (2010) *Information Structure: Theoretical, Typological, and Experimental Perspectives*. Oxford, Oxford University Press.

COMPLEX DEADJECTIVAL VERBS BASED ON OPEN SCALE ADJECTIVES IN MANDARIN CHINESE: A COMPARISON BETWEEN *JIĀ* 加+ADJ. AND *NÒNG* 弄+ADJ. VERBS

Bianca Basciano

Ca' Foscari University of Venice

1. Introduction

This paper focuses on Mandarin Chinese complex deadjectival verbs based on open scale adjectives, as those in (1a-b)[1]:

(1) a. […] 巴斯德只是把葡萄酒加热到摄氏50度，就消灭了不受欢迎的菌体。

Bāsīdé	*zhǐshì*	*bǎ*	*pútáojiǔ*	*jiā-rè-dào*	
Pasteur	only	OBJ	wine	increase-hot-up.to	
Shèshì		*wǔshí*	*dù*	*jiù*	*xiāomiè-le*
Celsius		50	degree	then	eliminate-PFV

[1] The glosses follow the general guidelines of the Leipzig Glossing Rules. Additional abbreviations include CMP = complement marker and MOD = marker of NP modification. Unless otherwise specified, examples come from the BCC corpus of Modern Chinese: http://bcc.blcu.edu.cn/ (visited 2023/02/20).

bù *shòuhuānyíng* *de* *jūn-tǐ*
not well-received MOD bacterium/germ-form
'Pasteur simply heated the wine to 50 degrees and eliminated the unwanted microorganisms.'

b. 我用了20分钟把烧卖弄热。
wǒ yòng-le èrshí fēnzhōng bǎ shāomài
I use-PFV twenty minute OBJ steamed.dumpling
nòng-rè
make-hot
'I warmed up the *shaomai* in 20 minutes.'

Both *nòng* 弄 'make' and *jiā* 加 'increase' have been analysed as causative light verbs (see Basciano 2013, 2019). In particular, *nòng* 'make' has been considered as a causative light verb combining with different change of state verbs, including adjectives (see Lin 2001; Feng 2003; Zhu 2005; Jie 2008; Basciano 2013), while *jiā* has been analysed as the spell-out of both the causative component and the increasing event in the logical representation of the complex (degree achievement) verb (Basciano 2019).
Basciano (2019) observes that, while *nòng* 'make' may seemingly combine quite freely with adjectives, *jiā* only combines with open scale adjectives. In cases like (1a-b), *nòng* and *jiā* combine with the same adjective, apparently with the same meaning. Following Rothstein (2008), Basciano (2019) speculates that *jiā* specifies the direction of change, meaning 'cause an increase in a certain property', without specifying a value, while verbs formed with *nòng* specify a value in the property range, without specifying the direction, meaning 'cause to have the value X in the property range'.
This paper aims at comparing these verbs, focusing on their aspectual behavior, in order to identify the differences between them. I will argue that *nòng* 'make' and *jiā* 'increase' mark different senses of degree achievements, and thus they are not freely interchangeable as V_1 in complex deadjectival verbs based on open scale adjectives. The data for this research are drawn from the BCC corpus of Modern Chinese (Beijing Language and Culture University, 15 billion characters).
The paper is organized as follows: Section 2 introduces the light verbs *nòng* and *jiā*, describing their characteristics. Section 3 presents my research questions and the data used for this research. Section 4 presents

the results, showing the different aspectual behaviour of the two kinds of verbs at issue. Section 5 discusses the results and puts forth hypotheses to account for the differences observed between the two kinds of verbs. Finally, in Section 6, I offer some concluding remarks.

2. *Nòng* and *jiā* as light verbs

Modern Chinese has only a few instances of lexical causatives, more precisely labile verbs (as e.g. Eng. *break*), such as *chén* 沉 'sink' and *kāi* 开 'open'[2]. Normally, change of state verbs cannot be used transitively (see Basciano 2017):

(2) a. 窗玻璃破了。

chuāngbōli	*pò*	*le*
windowpane	break	PFV

'The windowpane broke.'

b. *我破了窗玻璃。

wǒ	*pò-le*	*chuāngbōli*
I	break-PFV	windowpane

'I broke the windowpane.'

In Modern Chinese, causativity is mainly expressed either by periphrastic means or with resultative compounds, consisting of two elements, where the second constituent represents the resultant state brought about by the action expressed by the first constituent, as e.g. *tī-pò* 踢破 'kick-break', *cǎi-duàn* 踩断 'step-break'. In addition, Mandarin Chinese has a few phonetically realized causative light verbs, i.e. verbs that have a general and abstract semantic content (see e.g. Grimshaw and Mester 1988), as e.g. *dǎ* 打 'beat, strike, hit', *nòng* 'make, handle', *gǎo* 搞 'do' (Ōta 2003 [1958]; Lin 2001; Feng 2003; Zhu 2005; Jie 2008; Basciano 2013). These verbs often do not represent a particular action, origin, or manner, differently from V_1s in resultative compounds, but are rather bleached verbs whose only function is to form the transitive version

[2] There are relics of causatives formed by means of tonal contrast as well, as e.g. *liáng* 凉 'cool, cold' / *liàng* 凉 'make cool, let sth. cool'.

of change of state verbs (V_2), as in the following example from Zhu (2005: 227):

(3) a. 小蟲死了 。
Xiǎo *Chóng* *sǐ* *le*
Xiao Chong die PFV
'Xiao Chong died.'
b. 小強弄死 / 搞死了小蟲。
Xiǎo Qiáng *nòng-sǐ / gǎo-sǐ-le* *Xiǎo Chóng*
Xiao Qiang make-die/ do-die-PFV Xiao Chong
'Xiao Qiang killed Xiao Chong.'

The light verb *nòng* 'make' combines with different change of state verbs, including adjectives, as e.g. *nòng-duàn* 弄断 'make-break (intr.), break', *nòng-xǐng* 弄醒 'make-awake, wake up', *nòng-chén* 弄沉 'make-sink, sink (tr.)', *nòng-gān* 弄干 'make-dry, to dry', *nòng-àn* 弄暗 'make-dark, darken'. In these verbs, *nòng* does not express a particular action and is thus devoid of lexical content, having just a general causative meaning. It is an element involved in the causative/ inchoative alternation, having a causativizing function:

(4) a. 头发干了。
tóufa *gān* *le*
hair dry PFV
'The hair dried.'
b. 现在我得帮你把头发弄干。
xiànzài *wǒ* *děi* *bāng* *nǐ* *bǎ* *tóufa*
now I must help you OBJ hair
nòng-gān
make-dry
'Now I have to help you drying your hair.'

The root *jiā* too has been analysed as a special case of causative light verb, which forms the transitive variant of change-of-state verbs based on open scale adjectives involving an increase in the property denoted by the adjective[3] (Basciano 2010, 2019). Examples of this kind of verbs are: *jiā-kuān* 加宽 'increase-wide, widen', *jiā-cháng* 加长 'increase-long, lengthen', *jiā-shēn* 加深 'increase-deep, deepen', *jiā-qiáng* 加强 'increase-strong, strengthen' (see also Steffen Chung 2006: 196). *Jiā* 'increase' is generally added to open scale adjectives and only to those involving an increase in the degree to which an object possesses a gradable property. Those involving a decrease in some property require a V_1 that marks the negative direction of the change in degree, like *jiǎn* 减 'decrease, subtract' and *suō* 缩 'shrink' (Steffen Chung 2006: 197-198), as e.g. *jiǎn-duǎn* 减短 'decrease-short, shorten', *suō-xiǎo* 缩小 'shrink-small, reduce, narrow (make smaller)'[4].

Basciano (2010, 2019) points out that these verbs can be considered as degree achievements, as Eng. *widen*, *lengthen*, *dry*, which express events that describe the change undergone by an object with respect to the gradable property introduced by the base adjective (Hay *et al.* 1999: 132). Hay *et al.* (1999) introduce a function INCREASE in the semantic representation of these verbs, which they assume to be conveyed in English by the suffix *-en* (e.g. *widen*) or by a Ø morpheme (e.g. *dry*), which takes a gradable adjective and returns a description of an event involving some property undergoing a change in its degree. The logical representation assumed by Hay *et al.* (1999: 132) is as follows:

(5) $[[\text{INCREASE } (\phi)\ (x)\ (d)\ (e)]] = 1$ iff $\phi\ (x)\ (\text{SPO } (e)) + d = \phi\ (x)\ (\text{EPO } (e))$

[3] Note that the right-hand constituents of these complex verbs are considered as intransitive verbs and not as adjectives, since they are able to act as change-of-state verbs (see e.g Sybesma 1997; Zhang 2006; Liu 2010; Basciano 2019). The same goes for the right-hand adjectival constituent of complex verbs formed with *nòng* 弄 as a light verb seen above.

[4] Kennedy and Levin (2002) point out that verbs like *shorten* could be seen as involving a decrease in some property, i.e. a decreasing change that involves an increase in negative properties. Nevertheless, they assume that a change in the degree to which an object possesses some (gradable) property involves an increase, of a positive or negative degree. In Chinese, in contrast, different roots are used, marking the two directions.

INCREASE (ϕ) (x) (d) is true of an event e just in case the degree to which x is ϕ at the beginning of the event plus d equals the degree to which x is ϕ at the end of the event; i.e., just in case x increases in ϕ-ness by d. (Hay *et al.* 1999: 132)

Therefore, the logical representation of the sentence in (6a) is as in (6b):

(6) a. Kim lengthened the rope
b. $\exists e$, d [increase (*long* (*rope*)) (d) (e)]

Accordingly, *Kim lengthened the rope* is true if the length of the rope at the end of the increasing event equals its length at the beginning plus some unspecified degree of length[5].
Hay *et al.* (1999) assume that this logical representation is the one underlying both transitive and intransitive degree achievement verbs, which differ for the presence or absence of a causative component; they omit the external argument and the causative component from the logical representation[6].

[5] Note, however, that Kearns (2007: 43-44) points out that many transitive deadjectival verbs, in collocations as e.g. *lower the blind, lengthen my pants, widen the road*, and *lengthen the rope*, may have event structures based on action scripts, i.e. prototypical courses of actions performed by the agent, rather than on a gradual property change in the theme. For example, the script associated to *lengthen pants* "is very like that for shorten pants or alter pants, the main components being to unpick the old stitching, re-pin the garment to the new configuration, and re-stitch it" (Kearns 2007: 43). In this case, it is the sequence of actions that provides the event structure, and the predicate does not denote a caused gradual increase in the length of the pants; the event structure is not associated with a property scale, and thus the telicity of these examples does not provide evidence for contextually licensed closed scales for adjectives such as *low* and *long* (open scale adjectives). For example, Kearns (2007: 44) remarks that "the rope may be gradually lengthened by weaving or it may be lengthened all at once by tying a new piece to it". In other words, the property scale does not play a role in the aspectual interpretation of these examples. I will leave aside this issue here.

[6] They point out that the exact analysis of the causative is not central to what they intend to represent. In addition, they are not sure whether, in the analysis

Following Hay *et al.* (1999), Basciano (2010, 2019) proposes that *jiā* is the spell-out of one of the relevant parts of the logical representation, i.e. the increasing event (Basciano 2019: 231):

(7) a. 我们加宽了路面

wǒmen	*jiā-kuān-le*	*lù-miàn*
we	increase-wide-PFV	road-surface

‘We widened the road surface’

b. ∃e, d [*jiā* 加 ‘increase’ (*kuān* 宽 ‘wide’ (*lùmiàn* 路面 ‘road surface’)) (*d*) (*e*)]

In other words, *jiā* generically expresses the increasing event leading to the gradual change of state characteristic of degree achievements. However, Basciano (2019) points out that *jiā* also represents the causative component, thus acting as a causative light verb: in other words, it is the spell-out both of the increasing event in the logical representation and of the causative component.

3. Research questions and data

According to Basciano (2019), while *nòng* may seemingly combine quite freely with adjectives, *jiā* is subject to many restrictions: it is generally not found with closed scale adjectives, as e.g. **jiā-gān* 加干 ‘increase-dry’, **jiā-shī* 加湿 ‘increase-wet’, **jiā-píng* 加平 ‘increase-flat’ (cf. *nòng-gān* 弄干 ‘make-dry, to dry’, *nòng-shī* 弄湿 ‘make-wet, to wet’, *nòng-píng* 弄平 ‘make-flat, flatten’). In addition, as seen above, *jiā* is generally added to open scale adjectives involving an increase in the degree to which an object possesses a gradable property, while those involving a decrease in degree need a different V_1 marking the negative direction (see Sect. 2). *Nòng*, in contrast, is not subject to

of the causative alternation represented by intransitive/transitive pairs (*the soup cooled* vs. *I cooled the soup*), the causative component should be included in both the transitive and the intransitive forms (e.g. Levin and Rappaport Hovav 1995) or only in the transitive one (e.g. Hale and Keyser 1986; Hoekstra 1992, 2004; Ramchand 2008).

such kind of restriction: both positive and negative adjectives[7] are allowed (e.g. *nòng-cháng* 弄长 'make-long' / *nòng-duǎn* 弄短 'make-short' vs. *jiā-cháng* 加长 'increase-long' / **jiā-duǎn* 加短 'increase-short' / *jiǎn-duǎn* 减短 'decrease-short'). In my opinion, this is strictly connected to the semantics of these two (light) verbs: *nòng* is a 'make' verb, and its causative use derives from a process of abstractive generalization (Moreno 1993): creation of a physical object → creation of abstract entities → bringing about of an event (Basciano 2013); cross-linguistically, verbs expressing the meaning of 'make' tend to become markers of causativity (Moreno 1993). This verb combines with different kinds of change of state verbs, forming their causative variant (see Sect. 2); it does not provide any information about the direction of the change of state. In contrast, *jiā* means 'make a quantity or a degree higher': thus, it contains specific information about the direction of the change of state undergone by the object; it is precisely the 'increase' component which makes it compatible only with positive adjectives. We assume that, as in the case of *nòng*, this use of *jiā* can be seen as the result of a process of abstractive generalization: addition of a quantity (physical object; e.g. *jiā liáng shuǐ* 加凉水 'add cool water') → addition of a quantity (abstract object; e.g. *jiā yā* 加压 'raise the pressure') → increase in the degree of a gradable property. In the latter meaning, it expresses a change in a particular direction in the value of a scalar attribute; it is thus incompatible with negative adjectives. In order to express a decrease in the degree of a gradable property, a verb marking the 'negative' direction is required, as seen above (see Sect. 2)[8]. Therefore, in Chinese there is a clear distinction between the 'in-

[7] According to Hay *et al.* (1999), positive adjectives, like e.g. *long* and *wide*, are analyzed as functions from objects to positive degrees, while negative adjectives, as e.g. *short* and *narrow*, denote functions from objects to negative degrees. However, according to Caudal and Nicolas (2005), antonymy is not lexically but contextually determined, and, differently from Hay *et al.* (1999), they do not use negative degrees.

[8] An anonymous reviewer suggested that another possible reason for this difference may be that, while *nòng* +Adj. verbs are phrasal, *jiā* +Adj. verbs are lexical and, as such, are much more restricted. However, this view apparently is not supported by the syntactic behavior of *nòng* +Adj. verbs, which display non-separability of constituents, except for potential complement markers: indeed, aspect

crease' and 'decrease' functions, with different verbs marking the direction of change; a decrease in some property is not seen as involving an increase of a negative degree (cf. Kennedy and Levin 2002; see fn. 4). However, in some cases, *nòng* and *jiā* can combine with the same adjective, apparently with the same meaning, as e.g. *jiā-rè* 加热 'increase-hot' and *nòng-rè* 弄热 'make-hot' (see exx. 1a-b). Basciano (2019: 231, fn. 59), based on Rothstein (2008), hypothesizes that these two roots convey a difference in meaning. Rothstein (2008) points out that, while the English verb *cool* means 'undergo a decrease in temperature' (see also Hay *et al.* 1999), and not 'get a value in the cool range', *become cool* means 'get to have a temperature value in the (contextually determined) cool range', without specifying the direction of change:

(8) When I took the soup out of the fridge it was so cold that it burned my mouth, but after some time at room temperature, it had become pleasantly cool/ *it had cooled. (Rothstein 2008: 192)

Following Rothstein (2008), Basciano (2019) suggests that *jiā* specifies the direction of change, meaning 'cause an increase in a certain property', without specifying a value, while verbs formed with *nòng* specify a value in the property range, without specifying the direction, meaning 'cause to have the value X in the property range':

(9) a. *jiā-rè* 'increase-hot, heat, warm up' ('cause an increase in temperature');
b. *nòng-rè* 'make-hot' ('cause to have a temperature value in the (contextually determined) hot range').

In this paper we aim at answering the following research questions:

markers, if present, must follow the whole verb complex, and not the main verb. Another possible reason for this difference suggested by an anonymous reviewer is that the right-hand constituent in *jiā* +Adj. verbs is a noun rather than an adjective (*jiā* +N). I leave this issue for further research.

1. What is the aspectual behaviour of deadjectival verbs based on open scale adjectives with *nòng* and *jiā* as V_1?
2. Are *nòng* and *jiā* freely interchangeable in degree achievements?
3. May *nòng* be freely attached to open scale adjectives?
4. Are *nòng* and *jiā* used for different senses of degree achievements?

In order to answer these research questions, I examined the behavior of ten pairs of deadjectival verbs based on open scale adjectives, differing for the V_1 used (*nòng* and *jiā*) in the BCC corpus (Beijing Language and Culture University, 15 billion characters):

nòng-rè 弄热 'make-hot' – *jiā-rè* 加热 'increase-hot'
nòng-kuān 弄宽 'make-wide' – *jiā-kuān* 加宽 'increase-wide'
nòng-shēn 弄深 'make-deep' – *jiā-shēn* 加深 'increase-deep'
nòng-zhòng 弄重 'make-heavy' – *jiā-zhòng* 加重 'increase-heavy'
nòng-qiáng 弄强 'make-strong' – *jiā-qiáng* 加强 'increase-strong'
nòng-dà 弄大 'make-big' – *jiā-dà* 加大 'increase-big'
nòng-cháng 弄长 'make-long' – *jiā-cháng* 加长 'increase-long'
nòng-kuài 弄快 'make-fast' – *jiā-kuài* 加快 'increase-fast'
nòng-gāo 弄高 'make-high' – *jiā-gāo* 加高 'increase-high'
nòng-hòu 弄厚 'make-thick' – *jiā-hòu* 加厚 'increase-thick'

I considered the following factors:

- Ability to take imperfective markers (durative and progressive).
- Ability to be followed by the resultative *dào* 到 'up to', which sets a boundary to the event.
- Ability to take bounded measures of change.

4. Results

First of all, looking at the results of our search, we can observe that, generally speaking, with open scale adjectives *jiā* is much more common than *nòng* as V_1. Thus, there is a strong preference for using *jiā* to form deadjectival verbs from open scale adjectives; with some open scale adjectives only very few instances of *nòng* + Adj. can be found, as shown in Table 1.

Verb	Tokens
jiā-rè 加热 'increase-hot'	17920
nòng-rè 弄热 'make-hot'	79
jiā-kuān 加宽 'increase-wide'	1214
nòng-kuān 弄宽 'make-wide'	3
jiā-shēn 加深 'increase-deep'	17394
nòng-shēn 弄深 'make-deep'	1
jiā-zhòng 加重 'increase-heavy'	35319
nòng-zhòng 弄重 'make-heavy'	1
jiā-qiáng 加强 'increase-strong'	401530
nòng-qiáng 弄强 'make-strong'	4
jiā-dà 加大 'increase-big'	94490
nòng-dà 弄大 'make-big'	338
jiā-cháng 加长 'increase-long'	2724
nòng-cháng 弄长 'make-long'	33
jiā-kuài 加快 'increase-fast'	132680
nòng-kuài 弄快 'make-fast'	58
jiā-gāo 加高 'increase-high'	2041
nòng-gāo 弄高 'make-high'	24
jiā-hòu 加厚 'increase-thick'	8045
nòng-hòu 弄厚 'make-thick'	1

Table 1. Number of tokens in the BCC corpus.

From the aspectual point of view, the two types of verbs at issue display different behaviours. First of all, verbs with *jiā* as V_1 can generally be modified by imperfective markers (the progressive *zài* 在 / *zhèngzài* 正在 and the durative *zhe* 着), while verbs with *nòng* as V_1 cannot (no occurrences found in our sample):

(10) 中国正在加深改革、扩大开放 […]
Zhōngguó zhèngzài jiā-shēn gǎigé kuòdà
China PROG increase-deep reform enlarge
kāifàng
open.to.the.world
'China is deepening the reform and opening more to the world […]'

(11) 目前，上海正在加快"四个中心"的建设。
mùqián Shànghǎi zhèngzài jiā-kuài sì ge
at.present Shanghai PROG increase-fast four CLF
zhōngxīn de jiànshè
center MOD construction
'At present, Shanghai is accelerating the construction of the 'four centers'.'

(12) […] 经济全球化趋势正在加深 […]
jīngjì quánqiúhuà qūshì zhèngzài jiā-shēn
economy globalization trend PROG increase-deep
'[…] the trend of economy globalization is deepening […]'

(13) 这一切都加深着加重着他们相依为命的感觉，加深着加重着他们想要寻求一个属于他们两人小世界的渴望。
zhè yīqiè dōu jiā-shēn-zhe jiā-zhòng-zhe
this all all increase-deep-DUR increase-heavy-DUR
tāmen xiāngyīwéimìng de gǎnjué
they depend.on.each.other.for.survival MOD feeling
jiā-shēn-zhe jiā-zhòng-zhe tāmén
increase-deep-DUR increase-heavy-DUR they
xiǎngyào xúnqiú yī ge shǔyú tāmén
want seek one CLF belong.to they
liǎng rén xiǎo shìjiè de kěwàng
two person small world MOD thirst
'All of this is deepening and making heavier the feeling of depending on each other for life, it is deepening and making heavier their need for a small world for the two of them.'

This suggests that *jiā* +Adj. verbs are basically atelic, while *nòng* + Adj. verbs are basically telic. This is further proved by the ability of *jiā* +Adj. verbs to appear with 'for X time' expressions, which are compatible only with atelic verbs and set a temporal boundary to the event. For example, with the query *jiārè * fēn zhōng* 加热*分钟 'heat up for * minutes' we found 127 occurrences, while no occurrences were found of *nòngrè * fēn zhōng* 弄热*分钟 'heat up for * minutes':

(14) […] 放入微波炉加热4分钟。

fàng-rù	*wēibōlú*	*jiā-rè*	*sì*	*fēnzhōng*
put-enter	microwave	increase-hot	four	minute

'[…] put it in the microwave and heat it up for four minutes.'

As it is the case for English degree achievements based on open scale adjectives (see Hay *et al.* 1999), with these verbs telicity may emerge contextually[9] or by adding a bounded measure of change:

(15) 达罗沙把香烟放到唇边，吸进一大口，烟灰又加长了一毫米左右。

Dáluóshā	*bǎ*	*xiāngyān*	*fàng-dào*	*chúnbiān*	
Darousha	OBJ	cigarette	put-to	lip	
xījìn	*yī*	*dà-kǒu*	*yānhuī*	*yòu*	*jiā-cháng-le*
inhale	one	big-mouthful	ash	again	increase-long-PFV
yī	*háomǐ*	*zuǒyòu*			
one	millimeter	about			

'Darousha put the cigarette to his lips, inhaled a big mouthful, and the ash lengthened about one millimeter.'

9 In particular collocations and contexts, verbs derived from open scale adjectives, which are usually atelic, may be associated with closed scales and behave telically: for example, in *the tailor lengthened my pants*, real-world knowledge imposes a conventional maximal length for pants (see Hay *et al.* 1999; but see fn. 5). Hay *et al.* (1999) assume a contextual telos also for a sentence like *the soup cooled in ten minutes*. The insertion of an *in*-adverbial is possible because *the soup cooled* is more informative with a telic interpretation: the soup cooled to some bounded degree, i.e. room temperature.

(16) […] 把水渠加宽三尺、加深一尺 […]
bǎ *shuǐqú* *jiā-kuān* *sān* *chǐ*
OBJ ditch increase-wide three *chi* (1/3 meter)
jiā-shēn *yī* *chǐ*
increase-deep one *chi*
'[…] they widened the ditch three *chi* and deepened it one *chi*.'

In addition, *jiā* +Adj. verbs may be followed by the resultative *dào* 'up to', which sets a boundary to the event (see also ex. 1a):

(17) 牛奶加热到 37度左右，加入酵母搅拌 […]
niúnǎi *jiā-rè-dào* *sānshíqī* *dù* *zuǒyòu* *jiārù*
milk increase-hot-up.to 37 degree about add
jiàomǔ *jiǎobàn*
yeast stir
'When the milk reaches about 37 degrees (has warmed up to 37 degrees), add the yeast and stir it.'

(18) 最后村里出资 2 0 万元，将桥面加宽到了 9 米。
zuìhòu *cūnli* *chūzī* *èrshí* *wàn* *yuán*
finally village invest 20 ten.thousand Yuan
jiāng *qiáo-miàn* *jiā-kuān-dào-le* *jiǔ* *mǐ*
OBJ bridge-floor increase-wide-up.to-PFV nine meter
'Finally, the village invested 200000 Yuan and widened the bridge floor nine meters.'

(19) 除了把姚明将要用的床加长到2.4米、淋浴喷头加高到顶着天花板之外[…]
chúle *bǎ* *Yáo Míng* *jiāngyào* *yòng* *de*
except OBJ Yao Ming will use MOD
chuáng *jiā-cháng-dào*
bed increase-long-up.to
èrdiǎnsì *mǐ* *línyù-pēntóu* *jiā-gāo-dào*
2.4 meter shower-nozzle increase-high-up.to
dǐngzhe *tiānhuābǎn* *zhīwài*
push.to.the.top ceiling excluding

> 'Except for lengthening the bed where Yao Ming will sleep up to 2.4 meters and heightening the shower nozzle up to the ceiling […]'

In contrast, *nòng* + Adj. verbs are never followed by bounded measures of change nor by the resultative *dào* 'up to'.
To sum up, despite the verbs we considered are all formed from open scale adjectives, they display distinct aspectual behaviour: while *jiā* + Adj. verbs are basically atelic and can be made telic contextually or by adding bounded measures of change, *nòng* + Adj. verbs are telic.

5. Discussion

Since the two kinds of verbs at issue display distinct aspectual behaviour, I argue that *jiā* and *nòng* mark different senses of degree achievements. In this section, after presenting an overview of the main characteristics of degree achievements and on the properties of their base adjectives, I will put forth my hypothesis of the differences between these verbs.
Degree achievements are peculiar from the aspectual point of view since they display both telic and atelic behaviour according to standard diagnostics; thus, they do not necessarily entail the achievement of an endstate (Levin and Rappaport Hovav 1995: 172). Whether this ambiguity is related to the nature of the property of the scale denoted by the adjective or not is a matter of debate. According to Abusch (1986), the atelic sense of a deadjectival verb is 'become A-er', while the telic sense is 'become A'. Levin and Rappaport Hovav (1995: 129-130) and Jackendoff (1996: 331) consider the change of state described by these verbs as a movement along a path constituted of degrees of a property indicated by the adjectival base; Jackendoff (1996) points out that, if the path has a boundary, reaching the property described by the adjective, the sentence is telic, while, if the path is unbounded, going on indefinitely in the direction described by the adjective, the sentence is atelic. Kennedy and Levin (2002) observe that verbs of gradual change have as part of their meaning gradable properties; telicity is

not determined by a lexical diacritic (e.g. [±bounded]) or by some morphosyntactic features, but rather by the semantic properties of the degree of change.

According to Hay *et al.* (1999), the properties of the base adjective are crucial in determining the telicity of degree achievements: degree achievement verbs derived from closed scale adjectives are usually telic (*the clothes are drying* does not entail *the clothes have dried*), unless an adverbial or an explicit denial is added to cancel the 'completely' implicature (e.g. *I strengthened the rope, but not completely*); in contrast, degree achievement verbs derived from open scale adjectives are usually atelic (*the snow is slowing* entails *the snow has slowed*). However, Hay *et al.* (1999) point out that, in particular collocations and contexts, verbs derived from open scale adjectives may be associated with closed scales, displaying telic behaviour: for example, in *the tailor lengthened my pants*, real-world knowledge imposes a conventional maximal length for pants (see fn. 9). In addition, if a measure phrase providing an explicit bounded value is added, as e.g. in *Kim lengthened the rope 5 inches*, the predicate is always telic, regardless of the nature of the base adjective (Hay *et al.* 1990: 130).

Kearns (2007) holds a different view: she argues that telicity in degree achievements does not depend on the property nature of the scale of the base adjectives. She defends the traditional distinction between the two main senses of deadjectival verbs ('become A-er' and 'become A'; see Abusch 1986) and argues that there are two kinds of telic senses for deadjectival verbs, namely an achievement sense ('become A-er') and an accomplishment sense ('become A'). She points out that all deadjectival verbs can express at least the change of state 'become A-er' (comparative endstate); thus, since predicates which lexically entail an endstate are usually telic, all deadjectival verbs should be regarded as telic (see also Bertinetto and Squartini 2006). However, Kearns (2007: 36) points out that "although the comparative endstate is sufficiently identified to count as a telos, it is not uniquely specified, and thus a whole series of transitions to comparative endstates may fall under the same predicate". The process interpretation of these verbs, thus, comes from the repeatability of the non-unique comparative endstate, so that in *the room quietened for a few minutes* the interpretation is that 'throughout a period of a few minutes the room progressively became quieter and quieter'.

The telic accomplishment interpretation with deadjectival verbs is assigned the content 'become A' (where A is the positive form of the corresponding adjective), rather than 'completely' (giving the interpretation 'X becomes maximally A'). Thus, the interpretation of the implicature is given by the standard value of the property, and it is not dependent on the property nature of the scale (open scale adjectives vs. closed scale adjectives) but rather on the standard value ('become A'), which is a degree of the relevant property that constitutes the lower bound of the property region. Kearns (2007) assumes that a change-of-state accomplishment contains a process of iterated 'become A-er' transitions but has a necessary endpoint, i.e. the unique and non-repeatable transition 'become A', which provides the telos: beyond this endpoint the event cannot continue. Therefore, the telic sense in the traditional telic/atelic contrast is the accomplishment sense.

In verbs like *darken* or *quieten*, based on closed scale adjectives, the positive adjective does not lexicalize the maximal property value, so that the standard endstate may not coincide with the maximal value for *dark* (i.e. the standard value and maximal value for *dark* may differ): these verbs take both process and accomplishment senses freely.

In verbs like *clear*, *dry*, *empty*, also derived from closed scale adjectives, the positive adjective lexicalizes the maximal property value, and thus the standard value 'X is A' entails 'X is maximally A'; these verbs are associated with a default accomplishment sense[10].

Verbs like *cool* are based on open scale adjectives but have accomplishment senses which entail the standard endstate ('X is A'). Kearns (2007: 51) points out that "[t]here is no need to propose contextually given natural endpoints which are implicitly identified as contextual

[10] Kearns (2007: 64) argues that even deadjectival verbs with strong default accomplishment senses can have a process sense: all deadjectival verbs have the 'become A-er' sense, and this sense, as seen above, is interpreted as a process of iterated achievement-like transitions, with a duration adverbial.

upper bounds to the property scale. That is, there is no need to propose that an otherwise open property scale is closed just in case the verb is telic". Kearns (2007: 50) argues that in a sentence like *the soup cooled*, the telos is provided by the standard endstate 'X is A', and that the standard value is a normative value given by context and convention[11].

There are also deadjectival verbs based on open scale adjectives, like *wide* and *deep*, which do not take an accomplishment sense. The base adjectives of these verbs denote standard values with a lower bound partly inaccessible to modification (see also Caudal and Nicolas 2005), which correlates with the lack of an accomplishment sense for these verbs. Apparently, context cannot supply an upper bound for verbs like *widen*: *the gap widened in ten minutes* has only the achievement sense (within ten minutes the gap became wider), but not the accomplishment sense (the gap was widening throughout those ten minutes): cf. *the soup cooled in a few minutes*, which can have both the accomplishment sense (event duration: the soup was becoming cooler throughout a period of a few minutes, and at the end of that period the soup was cool) and the achievement sense (event delay: at the end of a few minutes the soup became cooler) (Kearns 2007: 36)[12].

---

[11] Kearns (2007: 50) points out that in the case of an event such as *the soup cooled*, the contextually determined telos correlates with contextual criteria for applying the positive adjective *cool*. For example, in *the soup cooled in ten minutes so we had to reheat it*, the context provides the salient presupposition that the soup should be eaten hot, so that the contextual standard for *cool* is 'too cool to eat'. In *the soup cooled in ten minutes so we started eating it* (before it cooled too much), the salient presupposition is that the soup can be too hot to eat, so the contextual standard for *cool* means 'cool enough to eat'. In both cases, the telos of the event is the onset of the state denoted by *the soup is cool*.

[12] See also the contrast between the following sentences (Kearns 2007: 52, 54):

a. *#The gap was half-wide/completely wide.*
b. *The soup was half-cool/completely cool.*

c. *#The gap widened (completely) in 90 seconds.*
d. *The soup cooled (completely) in ten minutes.*

Thus, while deadjectival verbs like *widen* only have a comparative endstate, verbs like *quiet*, *cool*, and *clear* have both the comparative endstate and a standard endstate.
To sum up, in Kearn's analysis, the availability of a telic accomplishment sense ('become A') depends on the characteristics of the standard value, rather than on the kind of scale of the adjectives (cf. Hay *et al.* 1999). Accordingly, the telic accomplishment sense is the strong default reading where 'X is A' entails 'X is maximally A', i.e. X bears the maximal possible degree of the relevant property; thus, deadjectival verbs like *empty* or *dry* have a default accomplishment sense. In contrast, at the other extreme, deadjectival verbs based on open scale adjectives that denote a standard value with a lower bound which is partly inaccessible to modification, like *widen* or *deepen*, cannot have an accomplishment telic sense. Thus, process resistant verbs are those based on closed scale adjectives for which the standard value entails that 'X is maximally A', while accomplishment resistant verbs are those based on open scale adjectives with an indeterminate lower bound. This is along the same lines claimed by Hay *et al.* (1999); however, the distinction is not based on the open or closed nature of the property scale, but rather on the standard value of the property. Accordingly, besides default accomplishments and process-only verbs, there are also verbs which display free variation; these verbs can be based both on closed and on open scale adjectives.
Based on Kearns' (2007) account of degree achievements, I argue that *nòng* always marks the telic accomplishment sense, i.e. 'cause to become A', whereby the interpretation of the implicature is given by the standard value of the property, and thus *nòng* +Adj. verbs are always telic. In contrast, *jiā* marks the process sense.
The fact that some base adjectives take *nòng* more easily than others (see Table 1) arguably depends on the characteristics of the standard

e. *The gap widened in ten minutes but it was still narrow.*
f. *!The soup cooled in ten minutes but it was still hot.*
Note that Kearns argues that when *completely* modifies an open scale adjective such as *cool*, the modifier is interpreted as 'indubitably, definitely, unquestionably'. Correspondingly, *the soup cooled completely* is interpreted as 'the soup cooled all the way to being (definitely) cool'.

value of the adjective involved. For example, only 3 occurrences of *nòng-kuān* 弄宽 'make-wide, widen' (vs. 1214 of *jiā-kuān* 加宽 'increase-wide, widen'), and no occurrences of *nòng-shēn* 弄深 'make-deep, deepen' (vs. 17934 of *jiā-shēn* 加深 'increase-deep, deepen') are found in the BCC corpus. Following Kearns (2007), I argue that this is due to the fact that open scale adjectives like *kuān* 宽 'wide' and *shēn* 深 'deep' denote a standard value with a lower bound which is partially inaccessible to modification, and the satisfactory telos for an accomplishment reading ('become A') is not identifiable. They are thus rarely found with *nòng*, which marks the telic sense, precisely because they resist the accomplishment meaning. The occurrence of these adjectives with the light verb *nòng* is possible only in certain contexts, as e.g.:

(20) 可那缝子细得连尾巴都塞不进去，就是用尽猿猴的气力也无法将缝隙扩大弄宽。

kě	*nà*	*féngzi*	*xì*	*de*	*lián*	*wěiba*
but	that	crack	thin	CMP	even	tail
dōu	*sāi-bù-jìnqu*			*jiùshì*		*yòngjìn*
all	squeeze.in-not-enter			even.if		exhaust
yuán-hóu		*de*	*qìlì*	*yě*	*wúfǎ*	*jiāng*
ape-monkey		MOD	strength	still	cannot	OBJ
féngxì		*kuòdà*		*nòng-kuān*		
crack		enlarge		make-wide		

'But that opening is so thin that even a tail cannot squeeze in it, nor could all the strength of an ape enlarge and widen it.'

I argue that here *kuān* is best interpreted as 'wide enough' rather than simply 'wide', and, accordingly, *nòng-kuān* 'make-wide' means 'make wide enough'. 'Wide enough' is not interpreted according to the contextual standard value for wide but is rather an implicit comparative, i.e. 'it is not wide enough for the current purpose': the implicit comparison is between the actual width of the opening and a specific width which is determined by the context (see Kearns 2007: 56).

A verb as *nòng-rè* 弄热 'make-hot', instead, is more common (79 occurrences of *nòng-rè* 'make-hot' vs. 17920 occurrences of *jiā-rè* 加热

'increase-hot, heat up') because it is based on an open scale adjective with a determinate lower bound. Thus, *nòng-rè* 'make-hot' expresses an accomplishment sense, where the telos is provided by the standard endstate 'X is A', which is a normative value given by context and convention.

Therefore, the reason why *jiā* is more common with open scale adjectives than *nòng* is arguably that those deadjectival verbs based on open scale adjectives are basically used in their process sense, which is marked by *jiā*. However, for a subset of those verbs, i.e. those based on open scale adjectives with a determinate lower bound, the accomplishment interpretation is possible too, and thus *nòng*, which marks the telic accomplishment interpretation ('become A'), can be used. In contrast, verbs based on open scale adjectives with an indeterminate lower bound, which are accomplishment resistant, are rarely found with the light verb *nòng* as V_1.

6. Conclusions

In this paper, using data drawn from the BCC corpus, I examined the behaviour of two kinds of complex deadjectival verbs based on open scale adjectives formed by means of two different light verbs, namely *nòng* +Adj. and *jiā* +Adj. verbs. I showed that these verbs, despite their apparently similar meaning, display distinct aspectual behaviour. Verbs formed with the light verb *jiā* as V_1 can generally be modified by imperfective markers (the progressive *zài/zhèngzài* and the durative *zhe*) and can appear with 'for X time' expressions, which set a temporal boundary to the event. This suggests that they are basically atelic. In addition, they can occur with bounded measures of change and with the resultative *dào* 'up to', which sets a boundary to the event, making the event telic. Thus, with these verbs telicity may arise contextually or by adding a bounded measure of change. In contrast, deadjectival verbs based on open scale adjectives with the light verb *nòng* as V_1 are always telic: they cannot be modified by imperfective markers and can neither occur with bounded measures of change nor with the resultative *dào* 'up to'. Therefore, these verbs differ in telicity, and the two light verbs are not freely interchangeable when forming deadjectival verbs based on open scale adjectives.

The data drawn from the corpus also highlight that *nòng* cannot be freely attached to open scale adjectives; quite to the contrary, its use is quite limited with open scale adjectives (see Table 1).

Based on Kearn's (2007) account of degree achievements, I argued that the light verb *nòng* always marks the telic accomplishment sense, i.e. 'cause to become A', and the telos is provided by the standard endstate 'X is A'. In contrast, complex verbs formed with *jiā* mark the process sense of degree achievements. Since the process sense is always available for degree achievements based on open scale adjectives, the light verb *jiā* is much more commonly used. The light verb *nòng* is found in those verbs in which the base is an open scale adjective with a determinate lower bound: verbs based on this type of adjectives allow both the process and the accomplishment sense. In contrast, when the base is an open scale adjective with an indetermined lower bound, and thus the satisfactory telos for an accomplishment reading ('become A') is not identifiable, *nòng* is rarely found as V_1.

Many issues still deserve further research. First of all, a more fine-grained analysis of deadjectival verbs based on open scale adjectives with determinate and indeterminate lower bounds may provide further insights into the phenomenon. In addition, a detailed study of the syntactic behaviour of these verbs could highlight other differences between them. Finally, an investigation of degree achievements verbs based on closed scale adjectives would provide a clearer picture of deadjectival verbs and could validate the hypotheses put forth in this paper.

Acknowledgments

I would like to thank the two anonymous reviewers for their comments and suggestions.

References

Abusch, D. (1986) "Verbs of Change, Causation and Time". *Report CSLI-86-50*. Centre for the Study of Language and Information (CSLI), Stanford University.

Basciano, B. (2010) *Verbal Compounding and Causativity in Mandarin Chinese*. PhD Dissertation, University of Verona.

--- (2013) "Causative Light Verbs in Mandarin Chinese (and Beyond)", in F. Montermini, N. Hathout and J. Tseng (Eds.) *Morphology in Toulouse. Selected Proceedings of Décembrettes 7*. Munich, Lincom Europa: 57-89.

--- (2017) "Causative Constructions", in R. Sybesma, W. Behr, Y. Gu, Z. Handel, C.-T.J. Huang and J. Myers (Eds.) *Encyclopedia of Chinese Language and Linguistics*, vol. 1. Leiden, Brill: 345-352.

--- (2019) "On the Event Structure of Chinese Resultative Compounds". *Cahiers de linguistique Asie Orientale*, 48: 173-241.

Bertinetto, P.M. and M. Squartini (1995) "An Attempt at Defining the Class of 'Gradual Completion Verbs', in P.M. Bertinetto, V. Bianchi, J. Higginbotham and M. Squartini (Eds.) *Temporal Reference, Aspect and Actionality*, Vol. 1, *Semantic and syntactic perspectives*. Torino, Rosenberg and Sellier: 11-26.

Caudal, P. and D. Nicolas (2005) "Types of Degrees and Types of Event Structures", in C. Maienborn and A. Wöllstein (Eds.) *Event Arguments: Foundations and Applications*. Berlin/New York, De Gruyter: 277-299.

Feng, S. 冯胜利 (2005) "轻动词移位与古今汉语的动宾关系". *Yuyan Kexue*, 4(1): 3-16.

Grimshaw, J. and A. Mester (1988) "Light Verbs and Theta-Marking". *Linguistic Inquiry*, 19: 205-232.

Hale, K. and S.J. Keyser (1986) "Some Transitivity Alternations in English", *Lexicon Project Working Papers*, 7. Cambridge (MA), MIT Center for Cognitive Science: 605-638.

Hay, J., C. Kennedy and B. Levin (1999) "Scalar Structure Underlies Telicity in 'Degree Achievements'", in T. Mathews and D. Strolovitch (Eds.) *SALT IX*. Ithaca, CLC Publications: 127-144.

Hoekstra, T. (1992) "Aspect and Theta Theory", in I.M. Roca (Ed.) *Thematic Structures: Its Role in Grammar*. Berlin/New York, De Gruyter: 145-174.

--- (2004) "Small Clauses Everywhere", in R.P.E. Sybesma, S. Barbiers, M. den Dikken, J. Doetjes, G. Postma, and G. Vanden Wyngaerd (Eds.) *Arguments and Structure: Studies on the Architecture of the Sentence*. Berlin/New York, De Gruyter: 319-389.

Jackendoff, R. (1996) "The Proper Treatment of Measuring Out, Telicity, and Perhaps Even Quantification in English". *Natural Language* *and Linguistic Theory*, 14: 305-354.

Kearns, K. (2007) "Telic Senses of Deadjectival Verbs". *Lingua*, 117: 26-66.

Kennedy, C. and B. Levin (2002) "Telicity Corresponds to Degree of Change", handout, *Topics in the Grammar of Scalar Expressions*. UCLA. https://semantics.uchicago.edu/kennedy/docs/telicity.pdf (visited 2023/02/20).

Levin, B. and M. Rappaport Hovav (1995) *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge (MA), MIT Press.

Lin, T.-H. J. (2001) *Light Verb Syntax and the Theory of Phrase Structure*. PhD Dissertation, University of California.

Liu, C.-S. Luther (2010) "The Positive Morpheme in Chinese and the Adjectival Structure". *Lingua*, 120: 1010-1056.

Moreno, J.C. (1993) "'Make' and the Semantic Origins of Causativity: A Typological Study", in B. Comrie and M. Polinsky (Eds.) *Causatives and Transitivity.* Amsterdam/Philadelphia, John Benjamins: 155-164.

Ōta, T. 太田辰夫 (2003 [1958]) 中国语历史文法. Beijing, Beijing Daxue Chubanshe.

Ramchand, G.C. (2008) *Verb Meaning and the Lexicon: A First-Phase Syntax*. Cambridge, Cambridge University Press.

Rothstein, S. (2008) "Two Puzzles for a Theory of Lexical Aspect: Semelfactives and Degree Achievements", in J. Dölling, T. Heyde-Zybatow and M. Schäfer (Eds.) *Event Structures in Linguistic Form and Interpretation*. Berlin/Boston, De Gruyter: 175-198.

Steffen Chung, K. (2006) *Mandarin Compound Verbs*. Taipei: Crane Publishing.

Sybesma, R.P.E. (1997) "Why Chinese Verb-*le* is a Resultative Predicate". *Journal of East Asian Linguistics*, 6(3): 215-261.

Xie, Z. 解正明 (2008) "轻动词假说与词典释义". *Cishu Yanjiu*, 4: 30-44.

Zhang, G. 张国宪 (2006) 现代汉语形容词. Beijing, Shangwu Yinshuguan Chubanshe.

Zhu, X. 朱行帆 (2005) "轻动词和汉语不及物动词带宾语现象". *Xiandai Waiyu*, 28(3): 221-231.

CHINESE LEXICOGRAPHY AND THE CRITICAL ANALYSIS OF LEXICOGRAPHICAL DISCOURSE

Chiara Bertulessi

University of Milan

1. Introduction

The present paper is concerned with lexicography in the People's Republic of China (PRC) and, specifically, with how this field and its products can be studied from a scholarly perspective that can be referred to as Critical Analysis of Lexicographical Discourse[1]. With this term I shall refer to a rather young critical approach to the study of lexicographical works (and mainly of language dictionaries), which finds its principal theoretical assumptions in the fields of both metalexicography (and, specifically, of critical lexicography) and Critical Discourse Analysis (CDA). As I also discuss in Bertulessi (2022), the development of this approach is very much indebted to the research carried out in rather recent times

[1] Drawing from Rodríguez Barcia (2012), in Bertulessi (2022), I employ the Italian term 'Analisi critica del discorso lessicografico', i.e., 'Critical Analysis of Lexicographical Discourse', which I shall also employ in this paper as a synonym to 'Critical Lexicographical Discourse Studies', proposed by Chen (2019).

by scholars such as Hornscheidt (2008), Rodríguez Barcia (2012, 2018) and Chen (2019), whose contributions have aimed to provide other researchers with what Hornscheidt (2008: 107) presented as a "a concrete research agenda for critical lexicographic research within critical discourse studies" or "a toolbox for critical lexicographic research". A few years after Hornscheidt, Rodríguez Barcia (2012, 2018), as part of her discussion of ideological implications in Spanish dictionaries, proposed a similar analytical approach, employing the term 'Análisis Crítico del Discurso Lexicográfico' (i.e., Critical Analysis of Lexicographical Discourse). In 2019, Chen also discussed a "discourse approach to critical lexicography", proposing the use of the English term 'Critical Lexicographical Discourse Studies' to refer to this approach. Together with earlier research and scholarly literature that focus on the interrelation of language and lexicography with society, culture, ideology (and power), the research carried out by these scholars constituted the theoretical and methodological basis for the critical study of Chinese lexicography and, specifically, of the *Xiàndài Hànyǔ Cídiǎn* 现代汉语词典 (XHC), which is presented in Bertulessi (2022).
To discuss specific issues, and for the sake of clarity, in the present paper it is still necessary to provide some fundamental definitions and briefly outline elements that constitute an essential part of the theoretical and methodological framework of Critical Analysis of Lexicographical Discourse. However, the main purpose of this paper is not to propose an in-depth discussion of the basic theoretical assumptions and methodological features of this approach: these have, in fact, already been largely discussed in the works cited above, including in Bertulessi (2022). On the contrary, the objective of the paper is to first provide an overview of this (rather young) analytical approach to metalexicography and, secondly, to connect these considerations to issues that regard the study of lexicographical works produced in the context of the PRC. To do so, in the sections that follow, I shall rely both on previous research carried out by other scholars as well as on my previous research, particularly with reference to the study of the XHC, which can be regarded as one of the most representative works of contemporary monolingual lexicographical discourse.
Based on these premises, the paper is structured as follows: Sections 2 and 3 present the theoretical background of Critical Analysis of Lexi-

cographical Discourse and the research carried out from this perspective. Section 4 focuses on lexicography in the context of the PRC; specifically, it is concerned with its relevance and on elements that, I argue, contribute to framing this field and the dictionaries produced in this context as interesting subjects for investigation from a critical perspective concerned with the relationship between lexicography and ideology. Section 5 briefly touches upon the issue of adopting paradigms of Critical Discourse Analysis in the study of Chinese (lexicographical) discourse. Finally, in Section 6 some concluding remarks with regards to the critical approach to lexicography discussed in this paper are outlined.

2. A critical approach to conducting research on dictionaries

As part of her discussion on the lexicographical treatment of ideologically loaded items in monolingual learners' dictionaries of English, Moon (2014) observed that:

> In relation to lexicography, ideology is where dictionaries collide with the social world: it brings in impolite and polite aspects of language, taboo items, evaluative orientation, connotation, and cultural allusion; the sublexicons, of course, of semantic fields such as politics, religion, ethnicity, sexuality, and so on; and above all the role of lexis, an unstable and mutable role, in naming and othering. (Moon 2014: 85)

Moon's words reflect what nowadays constitutes a commonly shared belief in metalexicography[2]: dictionaries are never neutral and, as such,

[2] Metalexicography is defined by Hartmann and James (2001: 93) as a "complex of activities concerned with the status of lexicography". Moreover, as stressed by Gouws (2020: 3), Wiegand *et al.* (2020: 322) see metalexicography as "the overarching theoretical domain of lexicography", thus encompassing different areas of dictionary research, including critical dictionary research (Gouws 2020: 3). Hornscheidt (2008: 107), discussing critical lexicographical research, points out that "Metalexicography, as a sub-discipline of linguistics, is concerned with the social and/or stately interests reflected in dictionaries".

they should never be regarded as objective containers of the words and meaning of a language (Moon 1989; Fishman 1995; Benson 2001). As the keywords 'critical', 'lexicographical' and 'discourse' suggest, the theoretical assumptions of the approach of Critical Analysis of Lexicographical Discourse can be found in earlier research carried out by scholars concerned with the influence of culture and ideology on lexicographical activities and its products, as well as in Critical Discourse Analysis. Critical lexicography as a shared direction of research did not gain much scholarly attention until the 1990s (Chen 2019; Hornscheidt 2011). This does not mean, however, that the ideological nature of dictionaries and the ideological implications of lexicographical compilation had been ignored by scholars before that time (e.g., Beaujot 1989; Moon 1989). However, 1995 saw the publication of the volume *Cultures, Ideologies and the Dictionary*, edited by Kachru and Kahane. In his introduction to the volume (which constituted a collection of 37 papers "on selected Western and non-Western languages"), Kachru (1995), commenting on the ten different parts that constitute the volume, meaningfully stressed that:

> These are, then, the ten heads of Rāvaṇa, presenting aspects of culture, ideology and power: a lexicographer has to encounter one or more of these in any serious lexicographical enterprise. And a discerning user of the dictionary has the responsibility to identify what type of power – e.g., ideological and cultural – a dictionary fosters. (Kachru 1995: lxv)

One of the basic assumptions of the critical approach to lexicography that is discussed in this paper is that the dictionary constitutes a form of discourse, i.e., lexicographical discourse (Rodríguez Barcia 2012: 137; Benson 2001). As observed by Rodríguez Barcia, the conceptualisation of the dictionary as discourse is not new; this can be traced back in earlier research, as emerges, for instance, in the work by Dubois conducted in the 1970s, which established a relationship between dictionaries and discourse and, specifically, between dictionaries and pedagogical discourse (Rodríguez Barcia 2012: 137). This concept also acquires a particular significance in an influential work by Benson published in the early 2000s.

In presenting the theoretical assumptions of his research on ethnocentrism in the *Oxford English Dictionary* (OED)[3], Benson suggests that "dictionaries are best thought of as representations of language, which implies that the language as it is recorded in the dictionary is something other than a simple reflection of the language 'as it is'" (Benson 2001: 23). Starting from the notion of lexicographical representation – that is, from the notion of the representation of language and its meaning as intrinsic to the compilation of a dictionary – the author incisively affirms that dictionaries constitute "a historically situated form of discourse through which certain linguistic communities have come to represent their languages to themselves and others" (Benson 2001: 24).

Following Fairclough (1989: 17), discourse is here defined as "language as social practice determined by social structures". This conception of discourse, which is embedded in the tradition of Critical Linguistics and Critical Discourse Analysis, is also strictly connected to the idea that discourse "is not produced without context and cannot be understood without taking the context into consideration" (Chen 2019: 376). Conceiving dictionaries as forms of representation of languages and as historically situated forms of discourse implies that the context in which the process of lexicographical compilation occurs cannot be ignored but, on the contrary, it should be taken into great consideration when approaching a critical analysis of a dictionary.

Finally, the interdependency between lexicography and context should be understood as bidirectional, or, to employ the words of Fishman (1995):

> we must interpret dictionaries in context and see them as both resultant of and constructive of their contexts. Indeed, this is what we do with other cultural artifacts. We recognize them as reflections of their contexts but as more than reflections we recognize them as constituents of those contexts, contexts which we must try to

[3] It is interesting to mention that Benson's work is particularly concerned with the representation of China in the OED.

> know by means of as many other artifacts and cultural behaviors as possible. (Fishman 1995: 34)

3. Critical Analysis of Lexicographical Discourse: a brief overview

Those outlined in the previous section constitute some of the principal theoretical assumptions the approach of Critical Analysis of Lexicographical Discourse draws from. This section aims to briefly present selected aspects of the research carried out in the field, by focusing on the linguistic and social contexts that have been studied by the scholars cited and, especially, to highlight common features that characterise their proposed analytical approach.

Starting from the assumptions that modern monolingual dictionaries "are regarded as an important (re)source for the construction, manifestation, and naturalization of public attitudes within society" and that the "publicly assumed function of dictionaries as sources of correct language usage has to be critically analysed" (Hornscheidt 2008: 107, 108), Hornscheidt focused on the issues of racism and colonialism which emerged from a set of monolingual dictionaries of Danish, German, and Swedish. In the author's view, racism and colonialism constitute useful examples to show how specific meanings, that are constructed and presented as 'neutral' by the lexicographical definitions in the dictionaries are far from being neutral or even objective. They also contribute to the reproduction and naturalisation of "certain attitudes" (Hornscheidt 2008: 108) and, therefore, of certain meanings that characterise specific and often dominant worldviews. The toolbox for critical lexicographical research proposed by Hornscheidt involves different elements being scrutinised in the dictionary, proceeding from the macro to the micro-level, i.e., from dictionary prefaces, to meaning explanations and usage examples, to metalinguistic comments. For each of these elements, the author outlines a set of questions that aims to guide the researcher in the analysis of how certain meanings are constructed in the dictionary (Hornscheidt 2008).

Rodríguez Barcia's research focuses on Spanish lexicography, which also serves as a starting point for the methodological and analytical proposal outlined in her 2012 article. Although in this contribution she does not reference Hornscheidt's work[4], the author also proposes a set of steps for the analysis of different elements of the dictionary, that belong to the introductory elements and the microstructure of the dictionary, i.e., the structure, features, and contexts of the lexicographical definitions, including labels of usage and metalinguistic comments[5] (Rodríguez Barcia 2012: 147). Moreover, the author stresses the importance of studying the context of production of the dictionary, ranging from the institutional context to the more general historical context in which the compilation has taken place (Rodríguez Barcia 2012: 145)[6].

Chen's works on the subject is particularly concerned with bilingual lexicography as a process of recontextualization, and the author specifically focuses on English lexicography and bilingual dictionaries in China (2015, 2017a, 2017b, 2018). In his 2019 paper titled "Towards a Discourse Approach to Critical Lexicography", the author provides a thorough discussion of the main features of the approach that he designates with the term 'Critical Lexicographical Discourse Studies' (CLDS), and provides an in-depth discussion of its theoretical assumptions and principles. Moreover, Chen devotes a consistent part of this contribution to methodology, thus systematising the elements that characterise the analytical framework and methods to carry out an analysis of dictionaries that relies on this approach. Specifically, after providing definitions of key concepts (i.e., 'critical', 'power', 'text', 'ideology', 'discourse'), the contribution focuses on defining what, in the dictionary, "becomes the target of critique" (Chen 2019: 372), thus outlining the fundamental principles of this approach. Moreover, with respect to the description of the theoretical framework, and similarly to Hornscheidt and Rodríguez Barcia, the author points to

4 The author does reference Hornscheidt (2008) with respect to critical lexicography in a later article (Rodríguez Barcia 2018: 198).

5 In the analysis of the definitions, the author suggests that a critical analysis concerned with issues of ideological nature also take into consideration the aspect of modality, i.e., the lexicographer's positioning with respect to the lexicographical text or definition (Rodríguez Barcia 2012: 149-150).

6 See also Rodríguez Barcia's contributions from 2016 and, in particular, 2018, in which the author describes critical lexicography as being related to the general framework of 'ethnolexicography'.

the importance of taking into consideration both the macro and the micro-levels in lexicographical research. In this respect, Chen (2019: 378-379) first outlines several "macro-level questions for critical lexicographical discourse studies", questions that are mostly concerned with the role of dictionaries and lexicographers in society. From this perspective, the macro-level represented by the historical and social context of production of a lexicographical work is also assigned an important role in Chen's research. Secondly, a set of micro-level questions are outlined, which are based on those put forward by Hornscheidt (2008) (Chen 2019: 379-380).

As acknowledged by Chen (2019: 367-368), the "systematic explanation of what CLDS is" found in his contribution from 2019 also constitutes the result of his previous research, within which feature case studies that concern English-Chinese bilingual lexicography and the often ideological implications of meaning recontextualization in these products. Specifically, Chen's critical lexicographical research has focused on two editions (1975 and 2000) of the *Xīn Yīnghàn Cídiǎn* 新英汉词典 (A New English-Chinese Dictionary) (Chen 2015; 2018), as well as on the bilingualised edition (English-Chinese) of the *Longman Dictionary of English Language and Culture* (Commercial Press 2004, 2011) (Chen 2017b). These case studies (which, due to space limits, cannot be further described here) provide interesting and valuable insights into the interplay between lexicography and ideology, insights that concern not only the relationship between English and Chinese (lexicography) in the PRC, but also the relevance that lexicography has acquired in specific historical moments and continues to acquire today in that specific national context[7].

Finally, in my research (esp. Bertulessi 2022)[8], I conducted a critical and diachronic study of the XHC and of selected entries from this authoritative monolingual dictionary of Modern Standard Chinese.

[7] Moreover, a contribution by Ding (2021) was published in 2021, in which the author relies on the CLDS framework outlined by Chen (2019) to investigate the 'Chineseness' as is displayed, on different levels, in the first volume (2015) of *The Chinese–English Dictionary* (unabridged, 1st volume, Fudan University Press, 2015).

[8] See also Bertulessi 2021a.

As I mentioned in the introduction to this paper, in my study on the XHC I relied on many of the elements that constitute an essential part of the theoretical and analytical framework of Critical Analysis of Lexicographical Discourse, as conceptualised and discussed by the scholars whose research I have (with no claim for exhaustiveness) presented in this section. In particular, the main purpose of the research was to investigate the ways in which the lexicographical discourse constructed by the XHC reflects, and at the same time, interacts with the political and social – and, therefore, ideological – context of compilation, also by adopting a diachronic perspective, which has involved the critical analysis of the prefaces and the selected entries in all the seven 'official' editions of this dictionary (1978, 1983, 1996, 2002, 2005, 2012, 2016) and in an earlier 'trial edition' distributed in 1973. In this respect, my research has also looked at the interplay between the official dominant political discourse and the lexicographical discourse constructed by the XHC.
Due to space constraints, it is not possible to further present and discuss the results and considerations that emerged in previous research carried out from this perspective on lexicography in China (Chen 2015, 2017b, 2018, 2019; Bertulessi 2020, 2021a, 2022; Ding 2021). However, what also emerges from these studies is that dictionaries produced in the context of the PRC in different historical moments represent a fertile area of investigation also from the perspective of critical lexicographical research. Yet, to date, although metalexicographical research devoted to Chinese lexicography appears to be very rich (especially in the PRC), studies of this kind are still very limited in number.

4. The relevance of lexicography in the PRC

Contemporary Chinese lexicography rests on a rich tradition with a very long history, which is commonly regarded as dating back at least to the early centuries of the Chinese Empire, or even earlier (Chen 1982; Yong and Peng 2008; Casacchia and Gianninoto 2012; Bottéro *et al.* 2015)[9]. Contemporary lexicography in the PRC is a very dynam-

[9] Some scholars mention the *Ěryǎ* 尔雅 (*Approaching Elegance*, III century BC ca.?) as the first lexicographical work in China (e.g., Casacchia and Gianninoto

ic field which, especially in the past few decades, has produced a great amount of products designed to fulfil different linguistic and cultural needs (from monolingual to bilingual dictionaries, from general purpose to specialised dictionaries and encyclopaedias, from dictionaries of Standard Chinese to dialect dictionaries), many of which are acknowledged for their high quality, scientific standards and practicality. The significant growth in the number of lexicographical products published the PRC (especially since the 1980s) has also been accompanied by the sustained development of the scholarly field of lexicography, as is demonstrated, among other things, by the creation of scholarly journals as the influential Lexicographical Studies (*Císhū yánjiū* 辞书研究, 1979), and the establishment of the Chinese Association of Lexicography (*Zhōngguó císhū xuéhuì* 中国辞书学会, 1992) (Klöter 2013: 884; Wei *et al.* 2014).

In introducing his overview of Chinese lexicography at the beginning of the 1990s, Creamer (1991) observed that:

> At times lexicography and lexicographers have been at the forefront of nation-wide language reform movements, and at other times both have been ignored or dismissed as irrelevant. Some dictionaries have been included among the great books of China and others have been suppressed or destroyed. While some lexicographers have enjoyed privilege and fame at court or in the scholarly community, others, in the not too distant past, have been ignominiously dispatched to the countryside to "learn from the peasants". (Creamer 1991: 2595)

Creamer's words reference the importance attained by lexicography in China in specific historical moments, in which, far from remain-

2015: 27-34; Xue 1982: 152), while others (e.g., Bottéro *et al.* 2015) affirm that the Chinese lexicographical tradition goes back to the first century and mention the *Shuōwén jiězì*说文解字 (*Explaining graphs and analyzing characters,* II century) as the "prototype of a Chinese dictionary" (Bottéro *et al.* 2015). However, in scholarly research on the subject, other works are sometimes mentioned as early instances of (proto-) lexicographical activities (see Creamer 1991: 2595). On the topic, see also Yong and Peng (2008), especially pp. 25-28 and 41-43.

ing confined within the borders of scholarly and linguistic discussions, issues involving dictionaries have intertwined with matters of social, political, and ideological nature. This is also particularly true with regards to the social and political context of the PRC, not only in the Maoist era, but also in more recent times (Bertulessi 2020, 2022; Lee 2014). Starting from this assumption and from examples that are related to the history of the XHC and of lexicography in the PRC, below I shall outline some selected elements that have characterised and, in some cases, continue to characterise this field and that, in my opinion, contribute to making dictionaries produced in this context potentially interesting subjects for investigation from the perspective of critical lexicography. In this respect, I am aware that the examples I draw from the history of the XHC do not of necessarily imply that all the dictionaries produced in the PRC were influenced by the same factors and discussions. However, given the central role that is commonly attributed to the XHC, both in the context of contemporary China and from a historical perspective, it seems reasonable to argue that this dictionary can still be seen as an authoritative 'representative' of lexicographical discourse in the PRC. Specifically, the sub-sections that follow focus on three elements: 1) lexicography and the standard language; 2) the scholarly debate on the 'ideological character' of dictionaries; 3) lexicographical planning.

4.1 Lexicography and the standard language

In his work on the XHC, Lee (2014) pointed out that dictionaries "can be highly political precisely because of two unique functions they perform in language standardization", that is: a) "dictionaries define words, the most fundamental component of language" and, b) "they reflect the boundaries of 'legitimate' words" (Lee 2014: 428). With regards to the history of lexicography in the PRC, dictionaries, and especially monolingual dictionaries, were very quickly assigned a key role in the national endeavour towards language standardisation, which was officially promoted as part of the Communist Party's agenda since the mid-1950s, and which constituted a key element of the wider project for the construction of the new nation (Luo and Lü 1956; Lee 2014; Bertulessi 2022: 43-54).

In this context, for example, the decision to start the compilation of a monolingual dictionary of the standard language, which would later

become the XHC, was formalised in 1956, when the State Council assigned to the Institute of Linguistics of the Chinese Academy of Sciences (created in 1949)[10] the task to compile a dictionary that would contribute to the promotion of the common language (*Pǔtōnghuà* 普通话) and to set and popularise its linguistic standards (Guowuyuan 1956; Lee 2014; Bertulessi 2022: 49-50). The political and ideological nature of certain projects for dictionary compilation became particularly striking in the era of the Cultural Revolution (1966-1976), when some lexicographers even became the target of denunciation campaigns for the allegedly 'reactionary' contents of their dictionaries, and certain editorial teams were also put under the control of the "workers-peasants-soldiers propaganda teams" (Chen 2018: 492-493; Lee 2014; Bertulessi 2020, 2022: 57-65).
Although not comparable to what occurred during the Cultural Revolution, in more recent times, specific issues that concern dictionaries have continued to attract attention and draw criticisms from the public and the scholarly community (Chen and Zhao 2014). An interesting example of this is represented by what was defined as a "linguistic lawsuit" directed at the lexicographers of the sixth edition of the XHC soon after its publication in 2012 (Pellin 2014). This lawsuit targeted what was regarded as an excessive amount of neologisms and, particularly, the list of words "starting with Western letters" (*xī wénzì kāitóu de cíyǔ* 西文字开头的词语), such as NBA or *pos jī* POS 机, through which, according to the over one hundred people that signed the document containing the charges, the XHC was harming the Chinese language (Pellin 2014: 220). In this respect, it is interesting to observe that Chen and Zhao (2014: 189, 192), in a contribution on the importance of critical lexicography, have chosen to mention this case among the events involving dictionaries that have stirred reactions in society in contemporary China.

4.2 The scholarly debate on the 'ideological character' of dictionaries

From a critical perspective, another element of interest is represented by scholarly discussions regarding the ideological nature of diction-

[10] In 1977, the Chinese Academy of Social Sciences was also established, and the Institute of Linguistics has, since then, been a part of this Academy.

aries and lexicographical compilation. For instance, as discussed in Bertulessi (2022), in the history of the XHC and, more in general, of lexicography in the PRC, compilers and scholars in the field have focused, in different moments, on the issues of the 'ideological character' (*sīxiǎngxìng* 思想性) or the 'class character' (*jiējíxìng* 阶级性) of the dictionary (e.g., He *et al.* 1960 [2004]; Min 1979)[11]. To my knowledge and at least with regards to the XHC, the relevance explicitly attributed to the 'ideological' and 'class character' of lexicography has diminished from the early 1980s, and the scholarly debates concerning the 'qualities', or 'character' (*xìng* 性) of the dictionary have tended to focus mostly on other lexicographical features (Bertulessi 2022: 85-88). Recently, however, these issues have once again been at the centre of some contributions from lexicographers involved in compilation of the XHC, who, commenting on the revision carried out for the seventh edition of the dictionary (published in 2016), have focused on the importance of expressing the dominant 'ideological positioning' of the historical context in which compilation occurs (Tan 2018; Pan 2018; Wang 2019; Bertulessi 2022: 88-91)[12].

These examples and those presented in the previous sub-section with regards to the standard language suggest that, also in the contemporary Chinese context, language dictionaries continue to (often explicitly) represent potential sites for the definition and re-definition of meanings (Fairclough 1989; Benson 2001) and that, as such, deserve to be further investigated from a critical and discursive perspective.

4.3 Lexicographical planning

Another important element that characterises lexicography in the PRC is lexicographical planning, which has been actively promoted by the

[11] The contribution was originally published in 1960 in the journal *Zhōngguó Yǔwén* 中国语文 (*Studies of the Chinese Language*)

[12] For instance, Tan (2018) refers to the inclusion, in the 7th edition of the XHC, of new words, set phrases and usages emerged in the 'Xi Jinping era', while Wang (2019) comments on the revision of lexicographical examples carried out for this edition with the objective of expressing a 'ideological character' the reflects elements of the social and political life of contemporary China. See, for example, the comment on the revision of the usage example to the entry *xīhuà* 西化 'westernisation' (Wang 2019: 635).

State since the 1970s, and contributes to showing how lexicographical activities are a field to which the country's political leadership has attached and continues to attach great importance.

Since 1975, relevant institutions of the PRC have issued three national plans for the compilation and publication of lexicographical works, adapting the number and the typology of the works included in the planning also to the linguistic and social needs of the different historical moments in which the plans were issued (Wei *et al.* 2014; Wei 2015; Bertulessi 2021b). Specifically, the first plan was issued in 1975, during a period generally considered as being characterised by strong political instability, and it outlined a list of 160 language dictionaries (monolingual and bilingual) to be published by 1985 (SPA 1975). The second plan lasted between 1988 and 2000, and provided for the compilation of 189 lexicographical works, among which featured not only language dictionaries, but also specialised dictionaries and encyclopaedias, which were deemed tools that could assist the wider project for economic reform and modernisation (SAPP 1989; Bertulessi 2021: 7-8). Finally, the latest plan was issued in 2013 and is expected to be implemented by 2025. It outlines the publication of 189 works, including both revisions and new projects (SAPPRFT 2013), which, following two revisions (2016 and 2017), reached 235 works. As shown by the research carried out by Wei *et al.* (2014) and Wei (2015), particularly with regards to the first two plans, the number of lexicographical works published in those phases was substantially higher than the number of those outlined in the documents, especially from the 1980s and even more the 1990s, when lexicography saw an unprecedented growth in the PRC (Wei *et al.* 2014; Wei 2015: 6).

Besides considerations on the vitality of this sector, lexicographical planning in China proves to be an interesting subject of research also from the perspective of critical lexicography, in that it highlights the relationship between this field and the interests that the State projects on it and its products. In this regard, as discussed in Bertulessi (2021b), the analysis of the documents introducing each national plan provides insights on "the ways in which the official political ideology participates in framing both the guiding principles and the objectives of lexicographical activities in the PRC" and on how reference works are presented as "tools in service of and influenced by not only cultural, linguistic, and social needs, but also ideological and political needs" (Bertulessi 2021b: 3).

5. Some remarks on Critical Discourse Analysis and Chinese (lexicographical) discourse

As already mentioned, the critical approach to the analysis of lexicographical discourse discussed in this paper has also been inscribed within the well-established area of research of Critical Discourse Analysis, or Critical Discourse Studies. In this respect, an issue that has not been dealt with in the paper, but which appears to need further investigation, is related to the implications of adopting paradigms from Critical Discourse Analysis[13] in the study of Chinese lexicographical discourse. Critical Discourse Analysis constitutes a rather wide, multifaceted area of research that was developed in the West – and, specifically, in the European academia. It emerged from studies in Critical Linguistics in the 1970s and 1980s and became a more established approach especially in the 1990s (Wodak 2001). Since then, it has developed into multiple strands and often multidisciplinary approaches. As incisively pointed out by Van Dijk, "CDA is a – critical – perspective on doing scholarship: it is, so to speak, discourse analysis 'with an attitude'". It focuses on social problems, and especially on the role of discourse in the production and reproduction of power abuse or domination" (Van Dijk 2001: 96). Several scholars have already discussed some of the implications of adopting Critical Discourse Analysis in the study of discourse(s) produced in the Chinese context (Cao 2014; Tian and Chilton 2014; Shi-xu 2014; Wei 2019; Chilton *et al.* 2012) and questions on whether it, as a perspective that has originated within the Western scholarly context and that places emphasis on aspects such as political commitment, can be applied to the Chinese context (Tian and Chilton 2014; Cao 2014). Among the different views on this topic, Tian and Chilton (2014) argue that Critical Discourse

[13] See also Wei (2019) for a discussion on Critical Discourse Analysis and Chinese discourse. Moreover, on this topic, Chilton *et al.* (2012) also provide insightful considerations concerning the conceptualisation of the concepts of 'critical' and 'critique' in the West and in the Chinese context.

Analysis "needs some tailoring and appropriating when being applied in China" (Tian and Chilton 2014: 197). Tian (2008) thus proposed a "wider angle critical perspective of CDA", which focuses more on socio-political transformation and, specifically, "aims to understand the workings and functions of discourse in the socio-political transformations, that is, to understand how discourse works to construct and represent social realities in the complex of social context" (Tian and Chilton 2014: 198-199). Starting from these considerations and from the assumption that dictionaries constitute forms of discourse (which are nonetheless historically situated, Benson 2001), this "wider angle" critical perspective appears therefore to be particularly relevant also within the study of Chinese lexicography, and, specifically, with regards to the possibility of critically investigating lexicographical discourse (also in diachronic terms) in relation to the transformations that have characterised and continue to characterise the social and political context of the PRC[14]. This issue will, hopefully, be the subject for further discussions.

6. Concluding remarks

The present paper has outlined some of the defining features of the critical approach to the study of lexicographical discourse. Moreover, the field of lexicography in China has been taken into consideration with regards to this approach and by highlighting some of the elements that may contribute to framing this field as a particularly interesting subject for investigation from this scholarly and analytical perspective within metalexicography.

What hopefully emerges from this contribution is that the so-called Critical Analysis of Lexicographical Discourse constitutes a (rather new) approach to the study of dictionaries and reference works and, from a more general perspective on the field of lexicography,

[14] In this respect, some of the results concerning the analysis of the lexicographical treatment of selected entries from the XHC presented in Bertulessi (2022), as well as from the case studies from Chen (e.g., 2015, 2018) can be interpreted from this "wider angle critical perspective" on discourse and socio-political transformations.

an approach that can be adopted by (and adapted to) the study of different lexicographical and linguistic traditions, as well as research questions and objectives. In this respect, this should not be conceived as a ‘model’ to be applied to one’s own object of study, but rather as a “conceptual and analytical framework” that, “with its limitations” (Chen 2019: 384) can contribute to the systematisation of methods and principles of the area of research concerned with lexicography and ideology which, as mentioned, is not new. Specifically, it provides a set of questions and guidelines concerning what, in lexicographical discourse, deserves to be placed under scrutiny within a research concerned with the interplay between lexicography and ideology in society. Besides the theoretical assumptions, elements common to the frameworks proposed so far include: the importance of looking at context, both as a preliminary step (e.g., the historical, but also the institutional or editorial context in which the compilation of a dictionary has taken place) and also in analysing meanings as they are presented from lexicographical discourse, starting from the premise that discourse is context-dependent; the need to focus on different elements of the lexicographical work(s) that are being analysed, going from the macro to the micro-level, i.e., from prefaces and other introductory materials, to the wordlist, meaning explanations, usage examples and metalinguistic comments in the definitions. Other methodological choices should, instead, be based on the peculiarities of the subject of study, which could, for instance, determine very often different methods for the selection of the corpus of entries to be analysed, as well as different approaches with regards to synchronic vs. diachronic analysis.

References

Beaujot, J.-P. (1989) “Dictionnaires et idéologie”, in F.J. Hausmann, O. Reichmann, H.E. Wiegand and L. Zgusta (Eds.) *Dictionaries: An International Encyclopedia of Lexicography*, First Volume. Berlin/New York, De Gruyter: 79-88.

Béjoint, H. (1994) *Tradition and Innovation in Modern English Dictionaries*. Oxford/New York, Clarendon Press and Oxford University Press.

Benson, P. (2001) *Ethnocentrism and the English Dictionary*. London/New York, Routledge.

Bertulessi, C. (2020) "'Un minestrone di feudalesimo, capitalismo e revisionismo': lo *Xiandai Hanyu Cidian* 现代汉语词典 tra rivoluzione, sovversione e reazione", in C. Chiara, F. Carbone, R. Coppola and B. Occhini (Eds.) *Sottosopra. Indagine su processi di sovversione*. Quaderni della Ricerca 6. Napoli, UniorPress: 109-120.

--- (2021a) "Critical Analysis of Chinese Lexicographical Discourse: A Case Study of *-zhuyi* 主义 Entries in the *Xiandai Hanyu Cidian* 现代汉语词典", in A. Alexiev and P. Zygadlo (Eds.) *China and the World: Language, Culture, Politics*, vol. 2. Sofia, St. Kliment Ohridski University Press: 22-29.

--- (2021b) "Lexicographical Planning and Official Ideology: The Case of China's National Plans for the Compilation and Publication of Lexicographical Works", in H. Van de Velde and F.T. Dolezal (Eds.) *Broadening Perspectives in the History of Dictionaries and Word Studies*. Newcastle upon Tyne, Cambridge Scholars Publishing: 1-29.

--- (2022) *L'ideologia nel discorso lessicografico cinese. Analisi critica dello Xiandai Hanyu Cidian* 现代汉语词典. Milano, LED Edizioni Universitarie.

Bottéro, F., W. Xu and W. Behr (2015) "Lexicographic Tradition", in R. Sybesma, W. Behr, Y. Gu, Z. Handel, C.-T. J. Huang and J. Myers (Eds.) *Encyclopedia of Chinese Language and Linguistics*. https://referenceworks.brillonline.com/entries/encyclopedia-of-chinese-language-and-linguistics/lexicographic-tradition-COM_00000226?s.num=0&s.f.s2_parent=s.f.book.encyclopedia-of-chinese-language-and-linguistics&s.q=lexicographic+tradition (visited 2023/02/20).

Cao, Q. (2014) "Introduction: Legitimisation, Resistance and Discursive Struggles in Contemporary China", in Q. Cao, H. Tian and P. Chilton (Eds.) *Discourse, Politics and Media in Contemporary China*. Amsterdam/Philadelphia, John Benjamins: 1-21.

Casacchia, G. and M. Gianninoto (2012) *Storia della linguistica cinese*. Venezia, Cafoscarina.

Chen, W. (2015) "Bilingual Lexicography as Recontextualization: A

Case Study of Illustrative Examples in a *New English-Chinese Dictionary*". *Australian Journal of Linguistics*, 35: 311-333.
--- (2017a) "The Discoursal Construction of the Lexicographer's Identity in a Learner's Dictionary: A Systemic Functional Perspective". *International Journal of Lexicography*, 30(3): 322-349.
--- (2017b) "Lexicography, Discourse and Power: Uncovering Ideology in the Bilingualization of Monolingual English Dictionaries in China". *Pragmatics and Society*, 8(4): 600-628.
--- (2018) "Identity and Affiliation: Exploring Chinese Lexicographers' Communal Identification in the 1970s". *Australian Journal of Linguistics*, 38(4): 484-518.
--- (2019) "Towards a Discourse Approach to Critical Lexicography". *International Journal of Lexicography*, 32(3): 362–388.
Chen, W. 陈伟 and Y. Zhao 赵彦春 (2014) "现代词典学：一个批评性视角". *Xiandai Waiyu*, 37(2): 189-198.
Chilton, P., H. Tian and R. Wodak (2012) "Reflections on Discourse and Critique in China and the West", in P. Chilton, H. Tian and R. Wodak (Eds.) *Discourse and Socio-Political Transformations in Contemporary China*. Amsterdam/Philadelphia, John Benjamins: 1-18.
Creamer, T. (1991) "Chinese Lexicography", in F. Josef Hausmann, O. Reichmann, H.E. Wiegand and L. Zgusta (Eds.) *Dictionaries: An International Encyclopedia of Lexicography*, Third Volume. Berlin/New York, De Gruyter: 2595-2612.
Ding, J. (2021) "Sinicization as Glocalization in the Chinese–English Dictionary". *Lexicography: Journal of ASIALEX*, 8(2): 166-187.
Fairclough, N. (1989) Language and Power. London, Longman.
Fishman, J.A. (1995) "Dictionaries as Culturally Constructed and Culture-Constructing Artifacts: The Reciprocity View as Seen from Yiddish Sources", in B.B. Kachru and H. Kahane (Eds.) *Cultures, Ideologies and the Dictionary: Studies in Honor of Ladislav Zgusta*. Lexicographica Series Maior 64. Tübingen, Max Niemeyer Verlag: 29-34.
Gouws, R.H. (2020) "Metalexicography, Dictionaries and Culture". *Lexicographica*, 36: 3-9.
Guowuyuan 国务院 (1956) "国务院关于推广普通话的指示". http://www.moe.gov.cn/jyb_xwfb/xw_fbh/moe_2128/moe_2326/moe_1144/tnull_14344.html (visited 2023/02/20).

Hartmann, R.R.K. and G. James (2001) *Dictionary of Lexicography*. London/New York, Routledge.

He, M. 何梅岑, H. Mo 莫衡 and C. Wu 吴崇康 (1960 [2004]) "词典里如何表现思想性", in J. Han 韩敬体 (Ed.) 《现代汉语词典》编纂学术论文集. Beijing, Shangwu Yinshuguan: 62-76.

Hornscheidt, A. (2008) "A Concrete Research Agenda for Critical Lexicographic Research within Critical Discourse Studies: An Investigation into Racism/Colonialism in Monolingual Danish, German, and Swedish Dictionaries". *Critical Discourse Studies*, 5(2): 107-132.

--- (2011) "Postcolonial Continuities in Danish Monolingual dictionaries: Towards a Critical Postcolonial Linguistics", in E.A. Anchimbe and S.A. Mforteh (Eds.) *Postcolonial Linguistic Voices: Identity Choices and Representations*. Berlin/Boston, De Gruyter Mouton: 265-298.

Kachru, B.B. (1995) "Introduction", in B.B. Kachru and H. Kahane (Eds.) *Cultures, Ideologies and the Dictionary: Studies in Honor of Ladislav Zgusta*. Lexicographica Series Maior 64. Tübingen, Max Niemeyer Verlag: lxiii-lxvi

Klöter, H. (2013) "Chinese Lexicography", in R.H. Gowus, U. Heid, W. Schweickard and H.E. Wiegand (Eds.) *Dictionaries. An International Encyclopedia of Lexicography, Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin/Boston, De Gruyter: 884-893.

Lee, S.Y. (2014) "Defining Correctness: The Tale of the 'Contemporary Chinese Dictionary'". *Modern China*, 40(4): 426-450.

Luo, C. 罗常培 and S. Lü 吕叔湘 (1956) "现代汉语规范问题", in 现代汉语规范问题学术会议秘书处 (Ed.) 现代汉语规范问题学术会议文件汇编. Beijing, Kexue Chubanshe: 4-5.

Min, J. 闵家骥 (1979) "谈谈语文词典的阶级性问题". *Cishu Yanjiu*, 1: 48-59.

Moon, R. (1989) "Objective or Objectionable? Ideological Aspects of Dictionaries". *English Language Research*, 3: 59-94.

--- (2014) "Meanings, Ideologies, and Learners' Dictionaries", in A. Abel, C. Vettori and N. Ralli (Eds.) *Proceedings of the XVI EURALEX International Congress. The User in Focus. Bolzano*, EURAC Research: 85-105.

Pan, X. 潘雪莲 (2018) “谈谈语文词典释义的思想性——以《现代汉语词典》第6版、第7版释义修订为例”. *Zhongguo Yuwen*, 6: 123-127.

Pellin, T. (2014) “The Sixth Edition of the Contemporary Chinese Dictionary and the ‘Linguistic Lawsuit’”, in B. Bock and M. Kozianka (Eds.) *Whilom Worlds of Words: Proceedings of the 6th International Conference on Historical Lexicography and Lexicology* (Jena, 25-27 July 2012). Hamburg, Kovac: 215-227.

Rodríguez Barcia, S. (2012) “El análisis ideológico del discurso lexicográfico: una propuesta metodológica aplicada a diccionarios monolingües del español”. *Verba*, 39: 135-159.

--- (2016) “*El Diccionario de la lengua española* (2014): análisis del nuevo discurso lexicográfico de la RAE”. *Lexis*, 40(2): 331-374.

--- (2018) “De la etnolexicografía a la lexicografía crítica”. *Revista de Investigación Lingüística*, 21: 186-206.

SAPP (State Administration of Press and Publication) (1989 [1991]) “关于全国辞书编写出版规划(1988–2000年)的报告”, in 中国出版年鉴 *1989*. Beijing, Guojia Chuban Nianjian: 473-478.

SAPPRFT (State Administration of Press, Publication, Radio, Film and Television) (2013) “关于印发《2013—2025年国家辞书编纂出版规划》的通知”. https://web.archive.org/web/20150320224120/http://www.gapp.gov.cn/news/1663/159617.shtml (visited 2022/09/07).

Shi-xu (2014) *Chinese Discourse Studies*. Basingstoke (UK), Palgrave Macmillan.

SPA (State Publishing Administration) (1975 [2013]) “国家出版事业管理局关于中外语文词典编写出版规划座谈会的报告”，in中国新闻出版研究院（Ed.）中华人民共和国出版史料*14 (1966-1976)*. Beijing, Zhongguo Shuji Chubanshe: 251-265.

Tan，J. 谭景春（2018）“谈谈语文词典收词的思想性———以《现代汉语词典》第7版增补条目为例”. *Zhongguo Yuwen*, 2: 224-225.

Tian, H. (2008) “Critical Perspectives on Discourse Studies”. *Foreign Language Teaching and Research*, 40(5): 339-344.

Tian, H. and P. Chilton (2014) “Issues in Discourse Approach to Social Transformations in China”, in Q. Cao, H. Tian and P. Chilton

(Eds.) *Discourse, Politics and Media in Contemporary China*. Amsterdam/Philadelphia, John Benjamins: 195-207.

Van Dijk, T. A. (2001) "Multidisciplinary CDA: A Plea for Diversity", in R. Wodak and M. Meyer (Eds.) *Methods of Critical Discourse Analysis*. London/Thousand Oaks/New Dehli, SAGE: 95-120.

Wang, W. 王伟 (2019) "谈谈语文词典举例的思想性———以《现代汉语词典》举例修订为例". *Zhongguo Yuwen*, 5: 632-636.

Wei, W. (2019) "Critical Analysis of Chinese Discourse: Adaptation and Transformation", in C. Shei (Ed.) *The Routledge Handbook of Chinese Discourse Analysis*. London/New York, Routledge: 95-120.

Wei, X. 魏向清 (2015) "国家辞书编纂出版规划的战略定位". *Cishu Yanjiu*, 1: 1-9.

Wei, X. 魏向清 et al. (2014) 中国辞书发展状况报告: 1978-2008. Beijing, Shangwu Yinshuguan.

Wiegand, H.E.M. Beißwenger, R.H. Gouws, M. Kammerer, A. Storrer and W. Wolski (Eds.) (2020) *Wörterbuch zur Lexikographie und Wörterbuchforschung*, vol. 3. Berlin/Boston, De Gruyter.

Wodak, R. (2001) "What CDA is about – a Summary of Its History, Important Concepts and Its Developments", in R. Wodak and M. Meyer (Eds.) *Methods of Critical Discourse Analysis*. London/Thousand Oaks/New Dehli, SAGE: 1-13.

Xue, S. (1982) "Chinese Lexicography Past and Present". *Dictionaries: Journal of the Dictionary Society of North America*, 4(1): 151-169.

Yong, H. and J. Peng (2008) *Chinese Lexicography. A History from 1046 BC to AD 1911*. Oxford, Oxford University Press.

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-constructional 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 Hinterölzl (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 Hinterölzl 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 Hinterölzl (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 languages. 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 without 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) $[_{\text{ShiftP}}$ A-Topic $[_{\text{ContrP}}$ C-Topic $[_{\text{FamP}}$ G-Topic $[_{\text{TP}}$ 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)[1]:

2) 张三很高兴.

$[_{\text{ShiftP}}$ *Zhāngsān*	$[_{\text{TP}}$	*<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.

[1] 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[2]:

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.'

[2] 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]$_i$*	*bú*	*suàn,*	***ránhòu***	*[zài*
So	this	one-CL	week	NEG	count	then	next
yí-gè		*xīngqī]$_j$*		*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]$_i$*	*huì*	*guòqu*
August	middle		1SG	will	over-go
shàng-bān,	***ránhòu***		*[wǒ]$_{j=i}$*		
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]$_i$*	*dōu*	*hěn*	*gāoxìng,*	***ránhòu***
So	very	many	people	all	very	happy	then
xīn	*lǎobǎn*	*lái*	*<tāmen>$_{j=i}$*		*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*>$_{i}$	*gàn*	*má*	*ne?*
Then	2SG	do	what	FP

‘Then what do you do?’

B: 然后准备上学。

Ránhòu	<*wǒ*>$_{j=i}$	*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 *Suǒyǐ hěn dà yí gè tīng,* ***ránhòu*** *chúfáng wòshì dōu a* ‘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 *Duì, wǒ zháojí zǒu, jīntiān, yīnwèi […]* ***ránhòu*** *wǒ zháojí zǒu* ‘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 *Shì Wèi de yí gè péngyou,* ***ránhòu*** *tā zhùzài Kensington* '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 […] *Jiù juézhe tè lèi ha,* ***ránhòu*** *wǒ, nà tiān jiù zài xiǎng* '[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 *Tā zhǐ yǒu yí suì duō yìxiē ba* 'He's just one year or so, doesn't he' B: 为啥呢？嗯 *Wèishá ne? En* 'Why? Uh' A: 然后，吃什么 […] ***Ránhòu****, chī shénme* 'Then, whatever he eats […]'
Verbal filler (VF)	Pause marker for lexical retrieval or local syntactic planning	然后呢，那个，那什么， 就那天小刚回来下班儿么 ***Ránhòu*** *ne, nà gè, nà shénme, jiù nàtiān Xiǎogāng huílái xiàbānr me* '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…… *Ránhòu míngtiān xiǎng qù nàr, qù hǎitān,* ***ránhòu****…* 'Then tomorrow I want to go there, to the sea, then…' B: 去海滩 *Qù hǎitān* '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 *Tā shuō lí xuéxiào yòu jìn,* ***ránhòu*** *pángbiān jiù shì dōu shì* shopping mall *ma* '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 *Dàjiā rúguǒ kǎo de bù hǎo,* ***ránhòu*** *tā jiù chéng duōshao* '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 *Zhōngguó, Yìndù,* ***ránhòu*** *Mòxīgē* '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 (κ), which is traditionally used for nominal variables (Brezina 2018: 90). Values of κ suggest a very good agreement for Subject overtness ($\kappa = 0.81$, $p = .00 < .01$, $z = 15.6$) and DM/

CJ use (κ = 0.83, p = .00 < .01, z = 11.6), and a substantial agreement for Topic type (κ = 0.77, p = .00 < .01, z = 14.9) and discourse functions (κ = 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 (χ^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): χ^2 (1) = 77.52, p = .00001.

	Frequency		
Use	**Observed**	**Expected**	χ^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**	**%**
Ránhòu 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: χ^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		
***Ránhòu* use**	**Frequency**	**Overt**	**Null**	**Total**
CJ	Observed	6	26	32
	Expected	19.83	12.17	
	χ^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	
	χ^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 overtness 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	
	χ^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	
	χ^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 overtness by Topic type.

The ability of *ránhòu* use (DM/CJ) and Topic (same/different) to predict Subject overtness 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		
***Ránhòu* 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 overtness, *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 form 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: χ^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		
***Ránhòu* use**	**Frequency**	**Overt**	**Null**	**Total**
CJ	Observed	4	26	30
	Expected	14.67	15.33	
	χ^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	
	χ^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[3]. 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]$_j$ has John bought <which book> and Mary read <which book>?

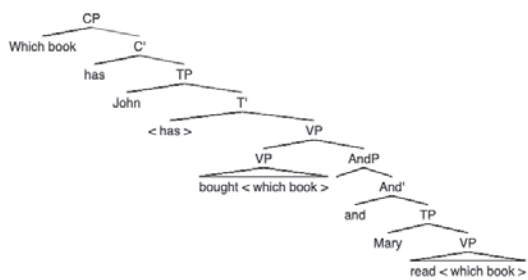


Figure 1. Structure of sentential coordination.

3 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).'

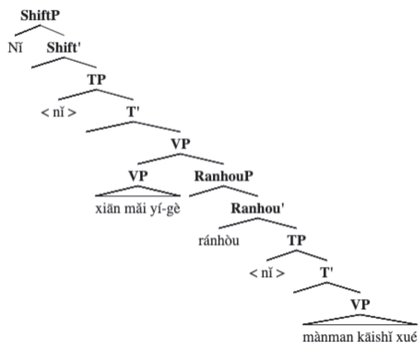


Figure 2. Structure of example 12.

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 $\text{Ground}_{\text{Addressee}}$P and $\text{Ground}_{\text{Speaker}}$P is assumed. Specifically, as shown in 13, $\text{Ground}_{\text{Speaker}}$P is dedicated to encoding the Speaker's attitude towards the utterance, while in $\text{Ground}_{\text{Addressee}}$P it is encoded what the Speaker believes to be the Addressee's attitude toward the proposition:

13) $[_{\text{RespP}}\ [_{\text{GroundAddresseeP}}\ [_{\text{GroundSpeakerP}}\ [_{\text{CP}}$

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 $\text{Ground}_{\text{Speaker}}$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[4]:

---

[4] 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.'

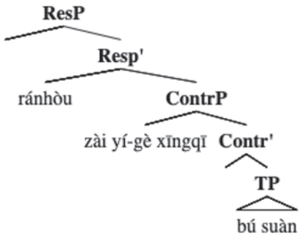


Figure 3. DM ránhòu in RespP.

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'

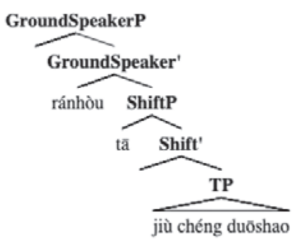


Figure 4. DM ránhòu in $\text{Ground}_{\text{Speaker}}$P.

Finally, as is shown in Figures 3 and 4, the DM *ránhòu* is assumed to be in the head of either RespP or $\text{Ground}_{\text{Speaker}}$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$_{Addressee}$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$_{Speaker}$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$_{Speaker}$P. Finally, the whole Ground$_{Speaker}$P move to Spec,RespP accountig 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…'

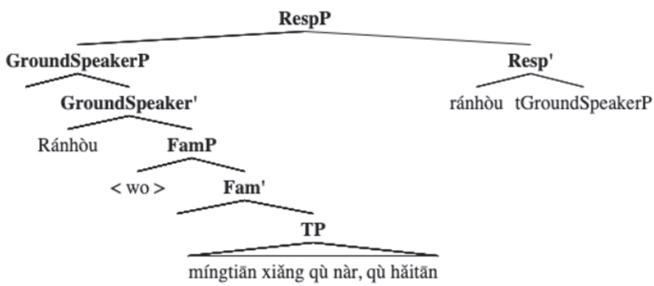


Figure 5. Movement of CP to Spec$_{Speaker}$P.

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 overtness.

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 $\text{Ground}_{\text{Speaker}}$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 Reexamination 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*. Berkley/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-Gòmez, 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 *ranhou* 然后 in a Corpus of Radio Conversations 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 *ranhou* 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 *ranhou* in Spontaneous Mandarin Conversation". *Chinese Language and Discourse*, 9(1): 1-25.

Wang, Y. (1997) "The Functions of *ranhou* 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.

INVESTIGATION ON SOME ITALIAN ARTISTIC TERMS ENTERED IN CHINESE: A DIACHRONIC AND SYNCHRONIC PERSPECTIVE

Feng Lisi

University of Milan

1. Introduction

This article is inspired by two lexical projects underway in Italy: LBC[1] and OIM[2]. LBC (Lessico dei beni culturali, 'Lexicon of cultural heritage') was founded in 2013 at the University of Florence. Its research demonstrates the global reach of Florence's cultural heritage, especially given the city's centrality in the history of art, as it tracks the reverberations of that heritage in the lexicons of nine foreign languages. The OIM (Osservatorio degli italianismi nel mondo, 'Watchdog for global Italianisms') is a database that collects words of Italian origin which appear in other languages, a considerable number of which belong to the field of art and architecture. Both projects incorporate words from

1 The website is open to the public: https://www.lessicobeniculturali.net/ (visited 2023/02/20).

2 Please visit the website: https://www.italianismi.org/ (visited 2023/02/20). The author of present paper is member of the research group of Chinese language.

the Chinese language. The extent of the dissemination of the language of Italian art is also reflected at the beginning of the eminent linguist Luca Serianni's *L'italiano nel mondo* (2011: 227) "l'espansione dell'italiano non è avvenuta attraverso le armi […] è stata garantita quasi esclusivamente dal suo spessore culturale" ('the expansion of Italian did not happen through force […] but was almost exclusively guaranteed by its cultural depth')[3]. In the OIM database, art vocabulary produces the second largest number of entries after music (Pizzoli 2019: 154).
The Italian Renaissance is a period of high interest to foreign scholars. The research question that this article will address is: how are Italian Renaissance artistic terms, which are considered as cultural symbols of the *bel paese*, understood and translated in Chinese? This article is divided into three parts: the first part will focus on the translatability and untranslatability of representative terms from the theory of visual art. The second part takes a diachronic perspective by considering the extent to which Renaissance art theories and techniques, such as *chiaroscuro* and *prospettiva*, have been introduced to China. Is there any trace of such an exchange in the history of language contact? The third part will investigate, from a synchronic perspective, how the most challenging abstract terms, such as Vasari's most representative term *disegno*, are translated and understood.

2. Translatability and untranslatability: Italian treasure art vocabulary

Translation is a productive process which connects two linguistic and cultural systems, emphasizing similarities and differences. Untranslatability is one of the central themes in the translation study. As Silvana Borutti observes, Heidegger's translators often leave concept-words, such as *Vorhandenheit* or *Gelassenheit*, untranslated (Borutti 2019). If we understand that the limits of translation are not an obstacle to understanding, it can allow us to productively engage with a culture that is very different from our own.
Because it has long provided a linguistic template in the discussion of visual art, Italian artistic vocabulary is now a cornerstone of artistic theory.

3 For example, the architectural term '*loggia*' mentioned by Serianni (2011) has entered into Estonian, Lithuanian, Latvian, New Greek, Georgian, Neo Hebrew.

This vocabulary is deeply influences that of other languages. The philologist and linguistic historian Matteo Motolese's 2012 monograph *Italiano lingua delle arti. Un'avventura europea (1250-1650)*, provides a detailed historical account of the spread of Italian art words across Europe[4].
Returning to translatability, the *Dictionary of Untranslatables* (2014) shows practically how to translate the "untranslatable" and demonstrates "philosophical thinking in translation". It is worth noting that as a guide on how to use the book, the first example given by the author is an Italian word *leggiadria*:

> The *Dictionary of Untranslatables* offers three types of entries. Among the "word-based" entries, some start from a single word in a single language, taken as "untranslatable", revealing a given constellation in time and/or space, such as *leggiadria*, which initially expresses the gracefulness of women in the Italian Renaissance and evokes for us the smile of the Mona Lisa. (Cassin *et al*. 2014: xxi)

This example clearly demonstrates the rich cultural connotations of these words and the challenges that come with them for translators. The object of the last section, *disegno*, is used in this dictionary 98 times in its original Italian form, as it is closely related to various other keywords, such as *concetto, leggiadria, genius, maniera, mimesi*s. The word is glossed as follows:

> *Dessein* is, along with dessin, one of the received translations of the Italian *disegno*. Eighteenth-century French broke with Italian tradition and, like German and English, separated the semantic fields of *dessein* and *dessin*. Nevertheless, *disegno* is to be thought of alongside "design", which not only means "drawing" (dessin) but also "the ability to grasp patterns".

4 I cite two examples analyzed in Motolese (2012): the form 'lambart' (<it. lombardo) was recorded in Catalan documents with the meaning 'master builder'; in Albrecht Dürer's writings was appeared the word 'quar' which corresponds to the Italian '*quadro*' with a simulation of the Venetian pronunciation.

The translatability and untranslatability of these words, which represent the treasures of Italian artistic and cultural heritage have attracted much attention from researchers from various fields (history of art, literature, linguistics) and from professionals (translators, tour operators, event organizers). The lexicon of critical terms, however, such as Leonardo's *sfumato* or Michelangelo's *non finito*, have not given rise to a semantic calque. Because of their difficulty to translate, even in modern European languages they have been kept in the original Italian. However, Chinese has some unique characteristics when it comes to translating foreign words, since it is "una delle poche lingue al mondo che meno frequentemente ricorre al semplice mantenimento di termini e sintagmi di origine straniera, […], i quali sono sempre resi nelle forme grafematiche della lingua di arrivo" (Bulfoni 2005: 107)[5]. From this point of view, the questions raised by Chinese translation strategies, such as whether to translate or not, how to translate, and the extent of the solution's vitality, are worth further reflection.

3. Historical translation solutions: *chiaroscuro*, *prospettiva*

This part intends to briefly review the artistic exchange between China and Italy in the sixteenth century, the critical period when European missionaries not only introduced Catholicism to China, but also brought about the first meaningful interactions between Chinese and Western languages, arts, culture and technologies. One of the major Western pioneers in the field of modern Chinese art, Michael Sullivan (1989: 1) argues that "the interaction between the cultures of Asia and the West is one of the most significant events in world history since the Renaissance". Beginning in the sixteenth century, Western thought, culture, and art had a huge impact on the East, but, conversely, Eastern culture did not have the same effect on the West.

At the end of the thirteenth century, Marco Polo, who served the Yuan Dynasty emperor Kublai Khan, had many opportunities to engage with Chinese art, but in his *Milione* he only mentions the luxurious decoration

[5] "One of the few language in the world which less frequently resorts to the simple maintenance of terms and phrases of foreign origin [...], which are always rendered in the graphemic forms of the target language".

of the palace buildings and some well-painted portraits of ancient Chinese rulers. For a long time thereafter, until the arrival of Jesuit missionaries, Western records of Chinese art were sparse (Sullivan 1989: 40-45). In general, exchange in the field of literature and art, compared with that in the field of technology (such as weapons, navigation) and commerce (silk, porcelain, plants, and spices), has left fewer written documents. This is perhaps because it requires a higher register of communication. In an essay discussing the language contact between China and Italy, Masini (2006: 17) examines some cultural keywords, such as *fēilùsuǒfèiyà* 斐录所费亚 filosofia 'philosophy', *fèixījiā* 费西加 fisica 'physics', *mòdáfèixījiā* 默达费西加 metafisica 'metaphysics', and those that have been adapted with a semantic calque and are still in use today, such as *dàxué* 大学 università 'university', *wénkē* 文科 materie umanistiche 'human Sciences' and *lǐkē* 理科 materie scientifiche 'scientific subjects'.

As far as the field of art is concerned, it must be mentioned that the Jesuit Library in Beijing contains a considerable number of books on art, especially architecture, from Italian authors, such as Andrea Palladio's *I quattro libri dell'architettura*, Marcus Virturius Pollio's *De architectura*, and Giovanni Rusconi's *Della Architettura*. The introduction of these Western books made the Beijing court increasingly interested in Western art (Sullivan 1989: 50-55). There are further noteworthy interactions in the field of art between China and Italy. Firstly, the Jesuit missionary Matteo Ricci (1552-1610) both brought paintings to China and transmitted artistic theories orally. Ricci arrived in Rome, a key cultural and artistic center at the time, when he was sixteen years old and studied in Rome until he was 25 years old before leaving for China. During these 9 years, there is no doubt that he encountered many of the artistic theories central to the Renaissance. Ricci's entry into Beijing to meet Emperor Wanli (1563-1620) took place on January 24, 1601, when the missionary was 48 years old and had already become an expert in Chinese language and culture. After being presented with numerous gifts from the Vatican, Wanli was drawn to the sacred paintings and claimed that the deities depicted there appeared to be alive. Wanli, however, soon began to feel uncomfortable observing such vivid and intense expressions. Compared to these paintings, the mechanical clock, another of the gifts he was honored with, gained better favour: this completely captured the emperor's attention and became the subject of discussion during numerous meetings between Ricci and Wanli. The emperor was curious about other things related to the West and began to ask Matteo Ricci how the buildings were built there. The Jesuit missionary was summoned many times to the imperial palace to paint the western buildings together with court painters (Fontana 2005: 202-209).

In these frequent exchanges and activities, we must ask whether Ricci contributed to the introduction of some Italian artistic theories and words? The case of the *Salus Populi Romani* (see Figure 1) is significant. Ricci carried a replica of this painting and hung it in a prominent place in his living room, so that many people could see it. For this reason, the *Salus Populi Romani* was well-known among Chinese artists, who were deeply influenced by the *chiaroscuro* technique. In this way, Ricci's painting gave rise to the style of *Zhōngguófēng shèng mǔzǐ tú* 中国风圣母子图 'Chinese style Virgin and Child'[6] (shown in Figure 2). He even genially used the *yin* and *yang* of traditional Chinese culture to explain the contrast between light and dark in chiaroscuro[7]. This creative cultural borrowing was particularly acute among his contemporary intellectuals[8], but does not appear to have obtained to the present day.

---

6 This work is archived in the Field Museum, labeled as '*116027 Chinese Madonna scroll*', photo by John Weinstein. Source at https://philippines.fieldmuseum.org/heritage/multimedia/79534, visited 2020/10/09. In 1909, Dr. Berthold Laufer brought this painting back to the Field Museum in Chicago. For more information about this scroll please see Martin P.S., 2003, and for further discussions about the Chinese Madonna's historical investigation please see in Dong L.H. 2013 and Li H., 2020.

7 In some works, this technique is also called the *āotūfǎ* 凹凸法 (concave-convex method).

8 For example, the landscape painter Wu Li 吴历 (1632-1718) was a famous figure who was influenced by the West art in the early Qing Dynasty. He had frequent contacts with Jesuits, and he expressed the following views on Chinese and European art: *Wǒ zhī huà bù qǔ xíngsì, bǐ quán yòng yīnyáng xiàngbèi xíngsì kējiù shàng yòng gōngfū* 我之画不取形似[…]彼全用阴阳向背形似窠臼上用功夫 (The paintings of our country do not pursue the similarity of external forms, and the paintings of their [European] countries are all focused on the rules and methods of light and dark contrast). For more information see Sullivan 1989: 55-65.

Figure 1. *Salus Populi Romani*, 9th-12th century (?), Basilica di Santa Maria Maggiore, Rome.

Figure 2. *Chinese Madonna scroll*, detail, Ming Dynasty (?), Field Museum of Natural History, Chicago.

A second source of interaction between China and Italy in the field of art is *Shìxué* 视学 *The science of vision* (1729), the first book in China to introduce Western theories of perspective[9]. In 1715, Giuseppe Castiglione (1688-1766), the Italian painter who probably had the greatest influence in China, started to work at the Beijing court. Castiglione became acquainted with another court painter, Nian Xiyao (1671-1738), the author of *The science of vision* (1729)[10]. Castiglione showed and explained to Nian the book of Andrea Pozzo, *Perspectiva pictorum et architectorum* (1693). From figure 3 and 4 we can directly see the effect of Andrea Pozzo on Nian; such correspondences appear in more than half of the book. Nian borrowed the theory of perspective and explained it via the calque *dìng-diǎn yǐnxiàn* 定点引线 (literally: 'fix a point and introduce lines').

Figure 3. Andrea Pozzo, *Perspectiva pictorum et architectorum* (1693).

[9] Before the introduction of a systematic explanation of perspective, Jesuit missionary Giulio Aleni is also a disseminator of Western vision theory. He wrote *Xìngxué cūshù* 性学牺述 (A general introduction to the study of the human nature), published in 1646, now in the Archives Romanum Societatis Iesu, Japonica Sinica II, 16, 21. For more information, please read Corsi (2013).

[10] See more in Corsi (2001).

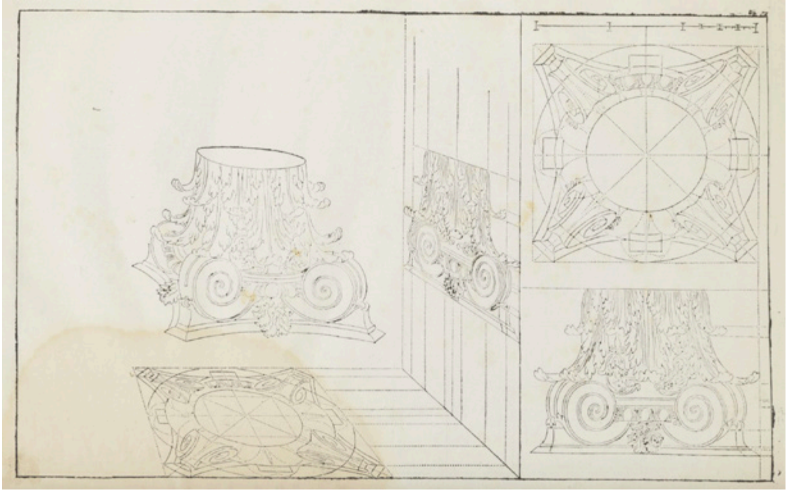


Figure 4. Nian Xiyao, *The science of vision* (1729).

4. Art terms in early bilingual dictionaries

The examples of *chiaroscuro* and *prospettiva* best demonstrate the dating and the manner in which Italian artistic terms entered into the Chinese language. Besides the sources detailed above, there is another type whose contribution cannot be ignored: the bilingual dictionary.

The manuscript *Dicionário Português-Chinês* 葡汉辞典 compiled between 1583 and 1588 is attributed to Michele Ruggieri and Matteo Ricci, and it may be the first foreign-Chinese bilingual dictionary in China. *Dicionário Português-Chinês* records Southern Mandarin, which is mixed with Fujian and Cantonese dialects. The basic vocabulary of Mandarin at the end of Ming Dynasty is almost the same as that of the modern Chinese lingua franca. There have been many linguistic analyses of this dictionary, such as from Yang Fumian 杨福绵 (1989), Levi Joseph (1998), Federico Masini (2003) and Wang Mingyu 王铭宇 (2014). The manuscript in which it is contained, *Japonica-Sinica I, 198* has 198 pages, from page 32 to page 156 it consists of Portuguese-Chinese glossaries, while the rest are linguistic, theological and scientific notes. A total of more than 6,000 Portuguese words have been collected. This part of article will focus on examining artistic vocabulary[11].

Words referring to occupation:

Portuguese	**Mandarin of South China**	**Reference**[12]
Artefice	absence of a translation	p. 46
Carpenteyro	*Mùjiàng, gōngshī* 木匠, 工师	p. 58
Capatero	*Píjiàng* 皮匠	p. 64
Imaginairo	*Huàjiàng, huàgōng* 画匠, 画工	p. 109
Lapidario	*Yùjiàng* 玉匠	p. 111

Words referring to colors:

Portuguese	**Mandarin of South China**	**Reference**
Cordeouro	*Jīnhuáng sè* 金黄色	p. 62

Words that describe qualities and represent philosophical abstractions:

Portuguese	**Mandarin of South China**	**Reference**
Fermoso	*Biāozhì, měi* 标志, 美	p. 105

Other words:

Portuguese	**Mandarin of South China**	**Reference**
Arte	*Shǒuyì, gōngfu* 手艺, 工夫	p. 46
artista	*Shǒuyì* 手艺	p. 46
Imagem	*Xiàngxíng, mú* 像形, 模	p. 109

11 For glossaries of Chinese and Italian terms of painting and sculpture please see two master theses: Aleci, M. (2017) *La scultura: storia, tecniche e materiali con repertorio terminografico italiano-cinese* and Grimaldi, N. (2018) *La pittura attraverso i secoli: aspetti storici e tecnici con repertorio terminografico italiano-cinese sulle arti pittoriche*.

12 The page numbers marked here correspond to the number in the lower right corner of original manuscript of the dictionary.

5. Chinese-Portuguese dictionary, Japonica-Sinica IV7[13]

This Chinese-Portuguese dictionary is presumed to have been compiled in the early seventeenth century and is found today in the Jesuit Archives in Rome, numbered Japonica-Sinica IV7. Most of the original cover has been lost, but there remains a fragment less than an inch wide on the inner cover, and only a few gilded letters remain. The Jesuit Archives have no record of when the original cover was replaced or what was on the original cover. According to a preliminary count, this manuscript contains approximately 7,228 entries. The dictionary is generally sorted according to the phonetic alphabet of Chinese characters. The entries include Chinese characters, words and phrases (Yang 2012). This part of the research aim to present some words in the field of art, as shown in the following tables.

Words referring to occupation:

Mandarin of South China	**Portuguese**	**Reference**[14]
Jiàng 匠	Official	p. 7
Rǎnjiàng 染匠	Tintureiro	p. 38

Words that describe qualities and represent philosophical abstractions:

Mandarin of South China	**Portuguese**	**Reference**
Jiàngxīn 匠心	Idea	p. 7
Jīngqiǎo 精巧	Oficial primo	p. 10
Róu 柔	Brand molle	p. 38
Xián 贤	Virtuoso, Sabio	p. 42
Yí 宜	Convem, Convenincia	p. 48
Yǎ 雅	Excellente	p. 49
Bù yǎ 不雅	Nao conven	p. 49

13 Full text is available online: https://ia803104.us.archive.org/9/items/Jap-Sin-4-7/JS-4-7.pdf (visited 2023/02/20).

14 The number of pages marked here corresponds to the number of pages of the free online dictionary manuscript document mentioned in the preceding footnote.

Yǎotiǎo 窈窕	Mulher fermosa	p. 51
Qiǎo 巧	Engenhoso	p. 64
Qiǎo shǒu 巧手	Buon mano, Buon official	p. 64

Other words:

Mandarin of South China	**Portuguese**	**Reference**
Cǎisè 彩色	Galanteo de cores	p. 4
Shuǐjīng 水精	Cristal	p. 10
Shuǐjīng 水晶	Cristal	p. 10
Shǒugōng 手工	O feitio da obra	p. 32
Fáng 坊	Partal botica di venda	p. 34
Fěnshì 粉饰	Ornar	p. 37
Rǎnfáng 染坊	Tinturaria	p. 38

6. Cultural-bound words: Vasari's *disegno*

From a synchronic perspective, Italian artistic terms entered Chinese mainly from the beginning of the 20th century. The following terms entered, respectively, via phonetic loan: *tǎnpéilā* 坦培拉 (tempera)[15], and semantic calque: *shībìhuà* 湿壁画 (fresco), *gānbìhuà* 干壁画 (secco), *míng'ànduìbǐ* 明暗对比 (chiaroscuro), *tòushì* 透视 (prospettiva)[16]. The interpretation of *chiaroscuro* as *míng'ànduìbǐ* 明暗对比 is somewhat similar to its interpretation in English as *light and shade*, even as mixed linguistic code partly in French and partly in English *claire and dark* in the seventeenth century (Motolese 2012: 139). *Prospettiva* derives from the late Latin *prospectivus* 'che assicura la vista' (which ensures vision), and thus its Chinese translation *tòushì* 透视 'see through' appears appropriate. The ways in which these words entered the Chinese language and the historical basis for their doing so are worth studying, but for now we are going to focus on the cat-

[15] Sometimes also called *jīdànhuà* 鸡蛋画 (literally: egg + painting), as this technique often uses eggs as a viscous agent.

[16] These words are collected in the aforementioned multilingual vocabulary project OIM (see Introduction).

egory of terms which are the most troublesome for translators, and that is abstract terms. They have rich ideological connotations and are inseparable from their historical and geographical environment. They are also considered as cultural-bound words, like *disegno* in the present case represents an extremely important concept in Vasari's theory of art.

I would first like to briefly explain the reasons why I chose the term *disegno* from Vasari's *Le Vite de' più eccellenti pittori, scultori e architettori* '*Lives of the Most Excellent Painters, Sculptors, and Architects*' (from now on *Le Vite*)[17], a foundational text in the artistic history of the European modern age and greatly contributed to the introduction of Italian artistic terms into other languages (Motolese 2012: 113). The problems of translating the term *disegno* have already been raised in translations of *Le Vite* into other languages: in the LBC database, this loanword is left in Italian in parallel texts of German, French and Spanish. With regard to translating the word into English, Biow (2018) writes that *disegno* is "a complex term", therefore he chose not to render it in English. Even in Li (2016) and Huang's (2019) work on Vasari, both written in Chinese, *disegno* is left in its original language.

Within the practice of translation, the languages and cultural systems involved partake in a process of "symbolic transformation", which brings together language, "consciousness" and "ontology" (Borutti 2020). Therefore, the term *disegno*[18] can be read as typical, in that it confirms the relativity of "translatability" and the absoluteness of "untranslatability" of a historical and cultural context.

The term *disegno* is in part an abstraction, linked to a conception of a mental state, which can variously be rendered in contemporary Italian as *pensiero*, *idea*, *piano*, *progetto*. The different contexts in which Varasi uses this term are listed below:

---

17 *Le Vite* is included in the aforementioned multilingual vocabulary project LBC (see Introduction) and the platform has also created a Vasari encyclopedia (with historical information on places, buildings, monuments, etc. described in the Vasari corpus).

18 For the background of the diffusion of the term design, which entered English as Italianism, and then in the form of *design* which returned to Italy as a return loan, see Cartago (2016).

> È necessario adunque che ella abbia corrispondenzia e che ugualmente ci sia per tutto attitudine, **disegno**, unione, grazia e diligenzia, le qual' cose tutte insieme dimostrino lo ingegno et il valore dello artefice. (Cap. III)[19]

In a more concrete sense, *disegno* is a hypernym, in its sense of a "work of art on paper". Its hyponym is *schizzo*, understood as the set of touches lightly hinted at with the pen:

> Gli schizzi chiamiamo noi una prima sorte di **disegni** che si fanno per trovare il modo delle attitudini et il primo componimento dell'opra, e sono fatti in forma di una ma[c]chia, accennati solamente da noi in una sola bozza del tutto. (Cap. XVI)

At the same time, Vasari understands the *disegno* as the set of outlines of the painted figures, called *profili*, *dintorni* or *lineamenti*:

> La pittura è un piano coperto di campi di colori, in superficie o di tavola o di muro o di tela, intorno a diversi lineamenti, i quali per virtù di un buon **disegno** di linee girate circondano la figura. (Cap. XV)

Disegno is a key concept in Vasari's art criteria, used frequently in his masterpiece as he unpacks its polysemy. Before investigating potential strategies for its translation, it is necessary to refer briefly to the two versions of *Le Vite* which have been translated into Chinese. The first one (Liu *et al.* 2003) consists of four volumes, on which several researchers have worked[20]. This translation was followed by that of

---

19 All examples cited are taken from *LV=Il Lemmario artistico nelle* Vite *di Vasari*, free online via: vasariscrittore.memofonte.it/lemmario.

20 In this contribution, account will be taken of the fact that the translations were not from Italian editions but are based on English translations. The version developed by the group led by Liu Yaochun is divided into three parts, according to Vasari's classification: each part corresponds to a phase of artistic development. Since the third part is very long, it has been divided into two volumes, for a total

Liu Mingyi (2004, 2005)[21] conceived as a compendium containing a selection of the most important lives.
Le Vite records the biography of more than two hundred artists, but to examine the validity of the Chinese rendering of the term *disegno*, this section focuses on the differences in the translation of this word in the part relating the life of Michelangelo. The table below lists four examples extracted from *La Vita di Michelangelo* (1550) to assess the validity of translations more accurately. V1 (Version 1) indicates the rendering of the term *disegno* in the translation of Liu Yaochun and colleagues (2003), while V2 (Version 2) refers to the rendering of the term in the translation of Liu Mingyi (2005). If this word was ignored in the translation, it will be marked as 'translation absent'.

i) Per il che, tornato Domenico, e visto il ***disegno*** di Michele Agnolo, disse: "Costui ne sa più di me".

V1. *Sùmiáo* 素描
V2. *Huà* 画

Huà is a hypernym of *sùmiáo*, as indicated in the definition of *Cíhăi* 辞海 (1999) "绘画的一种。主要以线条来描绘物体形象，不加彩色 ('A kind of painting. It mainly uses lines to describe the image of objects, without color')". According to the context, Michelangelo made a sketch without colouring, so the first *sùmiáo* 素描 translation is more precise.

ii) Alla quale opera non pensi mai scultore né artefice raro potere aggiungere di ***disegno***, né di grazia, né con fatica poter mai di finitezza, pulitezza e di straforare il marmo tanto con arte, quanto Michele Agnolo vi fece.

V1. *Gòutú* 构图
V2. translation absent

of four volumes and approximately 1890 pages.

21 Liu Mingyi's translation consists of about 450 pages. An updated edition was reprinted one year after publication.

In this context *disegno* is an entry that has an abstract meaning, parallel to *grazia*. *Gòutú* is a term used widely in plastic art, to indicate the organization of the position and relationship between objects, based on the aesthetic effect envisaged in the initial project (*Cíhǎi* 辞海, 1999). It is the first incarnation of the idea, that is, the first material elaboration of a mental concept. Therefore, here we could adopt another term as the alternative translation to *gòutú* 构图 that has an abstract meaning: *gòusī* 构思. According to the same authoritative dictionary *Cíhǎi* 辞海 (1999), *sī* 思 means: 'thought, idea, reflection'. This term refers only to mental activity, to thought and not to its material realization.

iii) Perché, volendo Michele Agnolo far porre in opera le statue, in questo tempo al Papa venne in animo di volerlo appresso di sé, avendo desiderio di fare la facciata della cappella di Sisto, dove egli aveva dipinto la volta a Giulio II. E già dato principio à ***disegni***, successe la morte di Clemente VII [...].

V1. *Cǎotú hé gòutú* 草图和构图
V2. translation absent

This is a somewhat peculiar example because Version 1 used two words in Chinese to render one Italian word. This also reflects the semantic richness of this term. *Cǎotú* 草图 refers to "初步画出的,不太精确的图 ('preliminary, imprecise drawings')". As mentioned above, *gòutú* 构图 refers to "the organization of the position and relationship between objects". *Disegno* in the original text contains both meanings, so it is appropriate to translate the term directly using both words *cǎotú hé gòutú* 草图和构图.

iv) E questo nell'arte nostra è quello esempio e quella gran pittura mandata da Dio a gli uomini in terra, acciò che veggano come il fato fa quando gli intelletti dal supremo grado in terra descendono, et hanno in essi infusa la grazia e la divinità del sapere. Questa opera mena prigioni legati quegli che di sapere l'arte si persuadono, e nel vedere i segni da lui tirati ne' contorni di che cosa ella si sia, trema e teme ogni terribile spirito sia quanto si voglia carico di ***disegno.***

V1. *Huìhuà* 绘画
V2. *Gòusī* 构思

In this context, *disegno* indicates the work completed after receiving instructions from God – we are talking about the *Last Judgment* painted by Michelangelo – and the meaning is not concrete, because it indicates the process of creation and invention. Thus, we pass from the divine idea to the realization of the work. Therefore Liu Yaochun's translation of *disegno* as *huìhuà* 绘画 'painting, drawing' is a solution characterized by a simplification of the concept. On the other hand, Liu Mingyi's translation, that is *gòusī* 构思, appears more adequate.

In addition to the two translations of the work, it is also necessary to consider that Yang Yongfang, one of the earliest scholars who studied Vasari in China, has rendered *disegno* as *jìqiǎo* 技巧, which means "*shúliàn de jìnéng* 熟练的技能 (qualified skills)" (*Cíhǎi* 辞海1999). Another Chinese scholar, Li Hong, in the work *Wǎsàlǐ hé tā de míngrén chuán* 瓦萨里和他的名人传 (Vasari and his Lives), decided instead to leave the word *disegno* in the original language, without any translation. According to the regulations in China on the publication of foreign works, the original language can be added in brackets in translations, but it is not recommended to leave only the foreign language without any translation. Following the present analysis of very different translation strategies, it can be deduced that current scholars are increasingly aware of the complexity of the semantic value of *disegno* and have not yet been able agree unanimously on a lexeme in the Chinese language.

However, the aim of this article is not limited to describing the diverse possible translations of this term, but instead to show various dynamic changes: from the obsolete solution *jìqiǎo* 技巧, to *gòutú* 构图, *gòusī* 构思, *sùmiáo* 素描, *cǎotú* 草图 and *huìhuà* 绘画 etc. More than six solutions have been used and discussed, with a trend towards directly adopting the original language. This because more and more Chinese scholars don't want to replicate the translating practices of their predecessors, because by reading a translation of a word like *disegno*, rather than the word itself, the understanding of the word precisely is deferred or occluded. Although with a long delay compared to other European languages, Italian artistic terminology has been received with growing acceptance and circulation in China, just as the translation and diffusion of the term *disegno* demonstrates in this contribution.

7. Conclusion

Art is an indispensable means for those outside of Italy to understand the country. How Italian art was introduced to China is not only the task of art history, but also the task of linguistics and translation studies. This article, as one of the first to address this topic from the perspective of the history of language contact, has discussed historically meaningful interpretations of two terms (*chiaroscuro* and *prospettiva*)[22] and has displayed more than 30 words extracted from two bilingual dictionaries, many of which are still in use today. For instance, with the continually improved status of artist and craftsman, today the word *jiàngxīn* 匠心 (literally: craftsmanship + heart) is used to express the excellence of the *maestro* in a philosophical way. This corresponds perfectly with *idea* as instantiated in the seventeenth century dictionary. Finally, abstract words have been discussed in terms of translatable relativity and untranslatable absoluteness, especially with the case of *disegno*, a typical and representative Italian Renaissance art theory word which poses great challenges for translators. However, there are some limitations on the scope of this paper and further research is needed. For example, the diachronic perspective is limited to the sixteenth century, while the synchronic view analyses only the word *disegno*. Two glossaries in Portuguese and Chinese are not well contextualized as there is yet to be a prolonged discussion of the legacy of the terms contained therein.

References

Biow, D. (2018) *Vasari's Words: The 'Lives of the Artists' as a History of Ideas in the Italian Renaissance*. Cambridge, Cambridge University Press.

Borutti, S. (2019) "Traduzione come conoscenza e rienunciazione". *Lingue Culture Mediazioni-Languages Cultures Mediation*, 6(2): 11-22.

[22] Although these two terms have now been replaced by new translation solutions.

Bulfoni, C. (2005) "La lingua cinese nell'era dell'informatica: alcune considerazioni preliminari". *Rivista degli Studi Orientali*, 78 (3/4): 107-115.

Cartago, G. (2016) "L'italiano e il design" in *Arte, design e moda: il mondo parla italiano*. Roma, Gruppo Editoriale L'Espresso: 47-72.

Cassin, B., E. Apter, J. Lezra and M. Wood (Eds.) (2014) *Dictionary of Untranslatables: A Philosophical Lexicon. Vol. 35*. Princeton, Princeton University Press.

Cihai 辞海 (2009[6]), Shanghai, Shanghai Cishu Chubanshe.

Corsi, E. (2001) "Perspectiva iluminadora e iluminación de la perspectiva. La versión del arte occidental de la perspectiva, de Nian Xiyao (1671-1738) en los prólogos a la 'Ciencia de la visión'". *Estudios de Asia y África*, 37(3): 375-418.

--- (2014) "La Diffusione delle conoscenze ottiche in Cina: il primato della visione nel Xingxue Cushu (1623) di Giulio Aleni SJ.", in M. Abbiati and F. Greselin (Eds.) *Il liuto e i libri. Studi in onore di Mario Sabattini*. Venezia, Edizioni Ca'Foscari: 231-240.

Dong, L. 董丽慧 (2013) "圣母形象在中国的形成, 图像转译及其影响—以'中国风圣母子'为例". *Wenyi Yanjiu*, 10: 132-142.

Fontana, M. (2005) *Matteo Ricci: un gesuita alla corte Ming*. Milano, Arnoldo Mondadori.

Huang, X. 黄小茵 (2019) "瓦萨里阿雷佐故居与《名人传》的诞生". *Meishu Guancha*, 2: 77-82.

Levi, J.A. (1998) *O diciónario Português-Chinês de Padre Matteo Ricci, S. J. (1552-1610): Uma abordagem histórico-linguistica*. New Orleans, University Press of the South.

Li, H. 李宏 (2016) 瓦萨里和他的名人传. Hangzhou, Zhongguo Meishu Xueyuan Chubanshe.

Li, H. (2020) "When Guanyin Encounters Madonna: Rethinking on Chinese Madonna from the Field Museum of Natural History, Chicago". *Buddhist-Christian Studies*, 40(1), 345-368.

Martin, P.S. (2003) "Stephen E. Nash", in S.E. Nash and G.M. Feinman (Eds.) *Curators, Collections, and Contexts: Anthropology at the Field Museum, 1893-2002*, *Fieldiana. Anthropology New Series*. vol. 36. USA, Field Museum of Natural History: 165-178.

Masini, F. (2003) "Notes on the First Chinese Dictionary Published in Europe (1670)". *Monumenta serica*, 51(1): 283-308.

Masini, F. (2006) "Rapporti fra spazio linguistico italiano e ambiente cinese", in E. Banfi and G. Iannaccaro (Eds.) *Società di Linguistica Italiana SLI 50. Lo spazio linguistico italiano e le "lingue esotiche". Rapporti e reciproci influssi.* Roma, Bulzoni: 7-25.

Motolese, M. (2012) *Italiano lingua delle arti: un'avventura europea (1250-1650)*. Bologna, Il Mulino.

Pizzoli, L. (2019) "Italiano e italianismi nel mondo: osservazioni sulla ricerca dei neologismi", in R. Bombi (Ed.) *Italiano nel mondo. Per una nuova visione*. Udine, Forum Editrice Universitaria Udinese: 151-158.

Serianni, L. (2011) "L'italiano nel mondo", in V. Coletti (Ed.) *L'italiano dalla nazione allo Stato.* Firenze, Le Lettere: 227-231.

Sullivan, M. (1989) *The Meeting of Eastern and Western Art*. California, University of California Press.

Wang, M. 王铭宇 (2014) "罗明坚、利玛窦《葡汉辞典》词汇问题举隅". *Liyun Yuyan Xuekan*, 1:138-150.

Yang, F-M. (1989) "The Portuguese-Chinese Dictionary by Matteo Ricci: A Historical and Linguistic Introduction", in *Proceedings of the Second International Conference on Sinology. Section on Linguistics and Paleography*. Taipei, Academia Sinica: 191-235.

Yang, H. 杨慧玲 (2012) "中国国家图书馆藏汉葡词典抄本及其史料价值". *Shixueshi Yanjiu*, 4: 118- 123.

Translations of Giorgio Vasari's *Le Vite*

Liu, Y. 刘耀春 (2003) 中世纪的反叛. Wuhan, Hubei Meishu Chubanshe.

Liu, Y. 刘耀春, Y. Bi, 毕玉 and L. Zhu 朱莉 (2003) 巨人的时代(一). Wuhan, Hubei Meishu Chubanshe.

Xu, B. 徐波, J. Liu 刘君 and Y. Bi 毕玉 (2003) 辉煌的复兴. Wuhan, Hubei Meishu Chubanshe.

Xu, B. 徐波, Y. Liu 刘耀春, X. Zhang 张旭鹏 and X. Xin 辛旭 (2003) 巨人的时代(下). Wuhan, Hubei Meishu Chubanshe.

Liu, M. 刘明毅 (2004) 艺苑名人传. Beijing, Zhongguo Renmin Daxue Chubanshe.

Liu, M. 刘明毅 (2005^2). 艺苑名人传. Beijing, Zhongguo Renmin Chubanshe.

INPUT-BASED AND OUTPUT-BASED INSTRUCTIONS: TEACHING ACTIVITIES FOR INTERROGATIVE CONSTRUCTIONS IN CHINESE AS A FOREIGN LANGUAGE TEXTBOOKS FOR ITALIAN LEARNERS

Gloria Gabbianelli

University of Urbino Carlo Bo

1. Introduction

In recent years in Italy, interest in the teaching of Chinese as a Foreign Language (CFL) has increased, including beyond higher education. This has created new needs to which both academic researchers and the Italian government have provided meaningful responses. Since 2010 Chinese language teaching was authorized to be included as a curricular subject in the curricular teaching of upper secondary schools (Langé 2018); in 2016, a group of experts including schoolteachers, academics, professors and representatives of the Embassy of the People's Republic of China and Confucius institutes in Italy, coordinated by the Ministry of Education, designed the *Sillabo della lingua cinese. Quadro di riferimento unitario per l'insegnamento della lingua cinese nella scuola secondaria di secondo grado* (Syllabus of Chinese language. A frame of reference for the teaching of Chinese in upper high schools) (Scibetta 2018). As a response to the increasing interest in studying Chinese language, over the last years several textbooks designed for the needs of new learners have appeared: new course books for high school, college and younger Italian-speaking learners have been published. Undoubtedly, this increase in CFL teaching materials has expanded the options of instructional materials for Chinese language teachers to choose from; however, it is auspicious that these new resources align with recent theories of Second Language Acquisition (SLA).

Although Foreign Language (FL) teaching has been increasingly integrated by digital tools and a number of online resources, several studies confirm that textbooks are still crucial in the language classroom. Studies have shown that both teachers and students still rely on textbooks (Garton and Graves 2014; Tomlinson 2012) and these resources turn out to be an effective reference for dealing with some difficulties caused by technological tools, progressively validated by the pandemic era (Trentin 2020). Furthermore, in some cases, textbooks are used as guidance in the curriculum and syllabus design, whereas they should be a tool that the teacher can adapt to the specific audience (Brown 2014). Given the role of textbooks in FL education, it is essential that they be in line with the newest SLA suggestions, to be a valuable support for teachers and learners.

This study aims to investigate the relationship between current SLA theories and the newest Chinese teaching materials for Italian-speaking learners. It examines five beginner-level CFL textbooks: three are currently used at high school level and two are destined for university learners. The study examines distribution and types of grammar instruction according to the classification of input-based and output-based.

2. Input-based and output-based instructions in second language acquisition research

In SLA research, both input and output are considered having a key role, and their application in the language teaching process has been investigated employing various pedagogical treatments (Benati 2017). Input refers to the language that learners are exposed to when reading or listening. SLA theories recognize the fundamental role of input in language acquisition: without being exposed to language input, acquisition does not take place (VanPatten 1996, 2003; Nassaji and Fotos 2004). Nevertheless, just being exposed to input is not enough. Indeed, since input is complex, the issue of the research is to investigate whether the exposition to the input is sufficient and how it should be presented to facilitate FL acquisition. In the Eighties, the *Natural Approach* (Krashen and Terrell 1983) recognized a fundamental role of input, claiming that input comprehension should always occur before L2 production. According to this approach, providing comprehensible input may support implicit learning (Krashen 1985), and only after having acquired L2 aspects through such comprehensible input are learners able to produce in the target language. In the Nineties some scholars supported the idea that the acquisition process does not occur in a totally implicit way. In order to produce acquisition, learners need to be guided to 'notice' the input by being exposed to language form (VanPatten 1990). VanPatten's method of *Processing Instruction* (PI) is a teaching model based on input-based instruction, which aims at guiding learners' attention to specific grammatical forms. VanPatten's model considers the psycholinguistic process which occurs during learners' exposure to the input (Figure 1).

Input → Intake → Developing system → Output
↑
Processing mechanisms
Focused practice

Figure 1. VanPatten's model of Processing instruction in foreign language teaching (VanPatten 2004: 6).

VanPatten reported the primacy of meaning on form (1990): when exposed to FL input the learner seems to process the input meaning before processing the form, which causes the need to direct learners' attention to the language form. Because acquisition does not occur automatically, or by just being exposed to comprehensible input, the PI model proposes some pedagogical steps to drive learners to connect a form-meaning relationship. Such pedagogical instruction aims to push learners to adopt strategies to correctly process the target form and meaning, including explicit grammar, information on learning strategies, and Structured Input (SI) activities (VanPatten 1993). Due to an initial learner's inability to convert the input into the intake, the learner should be guided to develop the grammar form through noticing techniques and SI, before any target language production; forcing the learner to produce the output may interfere with the ability of noticing the form. According to the PI's model, to guide learners to process the grammar form, instructions should focus on structured input activities based on a manipulated input, thus allowing the learner to absorb the input and convert it into intake. This helps learners develop the FL system and enables them to produce the output.

On the other hand, output is also recognized to have a crucial role in second language development (Swain 2005). Output is the language that the learner produces, either speaking or writing. SLA theories also support the role of output in the acquisition process. Swain's *Comprehensible output hypothesis* (1995) considers output effective for consolidating or modifying learners' existing knowledge and improving fluency. The recognized important role of output is to foster replication and automated employment of the target form without attention. In addition, by forcing learners to produce the target language and paying attention to the correct output they use to express an intended meaning, the output also contributes to creating new knowledge. VanPatten's (2003) *Processing theory* also considers output being important in promoting interaction, but not in contributing to the development of the linguistic system itself. Thus, output-based grammar instruction should provide opportunities for learners to use the target language, negotiating the meaning and using the target form for communication purposes in interactional situations. The research, overall, suggests the importance of enhancing the grammatical features in the input, as well as providing learners with opportunities to make form-meaning relationships.

The effectiveness of the two types of grammar instruction, based on input and output, has been investigated from different pedagogical models. A meta-analysis of comparative studies between comprehension-based and production-based instruction on L2 grammar learning (Shintani *et al.* 2013), observing 35 empirical studies, reveals that both types of instruction were effective for receptive and productive knowledge. The analysis reveals that instruction based on comprehension is more effective in immediate and delayed testing for receptive knowledge, while both types of instruction were equally effective in immediate testing for productive knowledge. Differently, instruction based on production activities became more effective in delayed testing. The authors also consider some variables, for example, their results show that comprehension-based instruction is more effective when learning is intentional rather than incidental, as well as when PI is adopted as the pedagogical process. Their findings conclude that combining both types of instruction may lead to more effective learning; however, while both are effective, they differ in the acquisition phase. During the initial phase of acquisition, the pedagogical procedures should push the learner to notice the input data and help them to convert input into intake; then, it should focus on output production for fixing the acquired knowledge (VanPatten 2003).

Although the teaching of grammar has been the subject of much research, production-based instruction and grammar teaching centred on accuracy of form and rule learning still seem to be dominant (Jean and Simard 2011; Shintani *et al.* 2013; Larsen-freeman 2015). Such a traditional approach to grammar teaching, based on explanations, examples, highly structured production practice and decontextualized activities, does not reflect effective practices for language acquisition according to the above SLA research, nor does it provide the input necessary for engaging learners' cognitive process for grammar acquisition. To favour the acquisition of the grammar system, this should occur in a meaning-focused way, and "grammar instruction should move from input only (manipulating input) to output practice. Grammar tasks should ensure that learners first process input language correctly and efficiently and then develop the competence to access the information about target features in their internal systems to create output" (Benati 2017: 384).

Considering that both input-based and output-based types of instruction are necessary for language achievement and are advantageous in favouring the acquisition process when applied in specific learning phases, we expect that pedagogical practice, as well as teaching materials, reflect these findings.

3. Studies on types and distribution of activities in FL textbooks

Studies have shown that many FL textbooks lack input-based activities that, according to SLA research, are necessary to foster FL acquisition (DeMil 2013; DeMil and Aubrey 2013). Despite some scholars claiming that the intent of the research itself should not be pedagogy-driven and that the roles of researchers and teachers are quite different (Larsen-Freeman 2009; 2015), the research in second and foreign language teaching also has the purpose to support the teaching practice. It is therefore useful that the findings of such research be integrated in pedagogical interventions, as well as in textbooks used in the classroom. However, some research projects, which have examined the type and distribution of activities targeting grammar instruction in FL textbooks, have shown that textbooks still tend to focus on production activities in grammar instructions.

Didi-Ogren and Goebel (2008) examined the introduction of verb vocabulary for discussing daily activities and verb conjugation paradigms in three commonly used university textbooks for studying Japanese as a foreign language in the United States. The research showed that numerous activities are designed to develop output but very few activities provide students with the opportunity to process the input and absorb it into the developing system. The authors present a set of suggested input-based activities to be integrated in the teaching practice. DeMil (2013) examined the grammar form of preterit tense in seven intermediate-level Spanish language textbooks. The classification adopted included the categories of input activity and output activities classified as mechanical drill, meaningful drill, or communicative drill. The research found that only one out of 48 was an input-based exercise. The author recommends that textbooks include more activities based on input, which is needed for acquisition. A study based on the same classification framework investigated the simple past tense in six beginner-level French language textbooks (DeMil and Aubrey 2013). The analysis counted a total of 65 activities; 12 were classified as input-based. The authors concluded that textbooks show some signs of considering new trends offered by SLA research. White and Farinelli (2014) conducted an analysis on six English FL books to determine the number of input-based activities. Their findings prove that the number of output-based activities was dominant, and the majority was included in the category of mechanical drills. They counted very few input-based exercises, concluding that the analysed materials lack input-based activities, and suggesting the integration of this type of instruction by creating new activities to be completed before the lesson to facilitate L2 acquisition. A previous study on the approach to grammar instruction in Italian beginner-level CFL textbooks also considered the categories of input and production-based activities (Gabbianelli 2020); findings of the study have shown that, similarly to investigations on different FL textbooks, production-based activities largely prevail; compared to input-based activities and among production-based activities, very few foster free communicative practice.

4. The current study

4.1 Purpose and RQs

The purpose of this study is to examine distribution and type of activities in five CFL textbooks recently published in Italy and to determine whether textbooks do integrate findings of language acquisition research. It aims to investigate how textbooks treat the concept of input, and how it is presented in teaching activities for grammar instruction. The research was guided by two questions:

(1) What is the distribution of activities, according to input-based and output-based classification in the analysed textbooks?
(2) How do these activities reflect the SLA research on grammar instruction?

5. Method

5.1 Textbooks

The present study examines five beginner-level textbooks for teaching Chinese as a foreign language (CFL). In order to observe the development trend of these types of textbooks in Italy, it was decided to examine the most recently published Italian textbooks, including textbooks destined to both university students and high school students. The analysed textbooks are the following:

Textbooks for university students:
1. *Comunicare in cinese*, 我会说汉语, vol. 1, Masini *et al.* 2021;
2. *Lingua cinese. Corso elementare*, 汉语初级教程, Bulfoni and Sun 2019.

Textbooks for high school students:
3. *Shangke le!* 上课了!, Varriano *et al.* 2022;
4. *Studiamo il cinese!* 咱们学中文吧!, Biondi *et al.* 2019;
5. *Parliamo Cinese*, 我们说汉语, vol. 1, Masini *et al*. 2018.

These textbooks were selected based on being the most recently published materials destined to Italian-speaking learners and are all in their 1st edition. Furthermore, in the introduction, four textbooks declare to adopt the communicative approach to Chinese language teaching, and to have designed the course book according to the communicative needs of learners, referring to the available frameworks for foreign language teaching. Specifically, the three textbooks destined for high school students refer to the *Syllabus of Chinese language* (2016), while the textbook, *Comunicare in cinese*, destined to university students, refers to the *Common European Framework of Reference for Languages* (CEFR 2001).
The three textbooks for high school teaching consist of a volume with two sections, one for classroom activities and one for exercises to do at home included in the final part of each book, whereas the two university textbooks present all the activities in the lesson section. All textbooks include CDs, while four out of five books, excluding *Lingua cinese. Corso elementare*, contain online supplementary resources consisting of an e-book, mp3 tracks, videos, supplementary fact sheets and character sheets. No further exercises are included in the online contents, thus, all the activities analysed in this study were presented in the paper course books.
Comunicare in cinese (2021) presents 18 lessons, each one divided into two parts (1 and 2) consisting of introduction activities, dialogue or text, vocabulary and grammar sections, followed by listening, speak-

ing, reading and writing exercises, and a final character writing section. A section for reviewing is presented every five lessons. In the final part, the course book includes glossaries, appendices and an index of vocabulary, grammar and communicative functions.

Lingua cinese. Corso elementare (2019) contains 18 lessons; each one is introduced by a vocabulary list, followed by the sections model sentences, grammar, conversation examples, reading and exercises. In the final part, an image of the characters presented in the lesson is included. Lessons 10 and 18 consist of reviewing exercises. The textbook also includes a glossary.

Among the textbooks for high school students the most recently published is *Shangke le!* (2022). This book presents a set of seven units of three lessons (1, 2, 3), each one with the subcategories warm up, dialogue, vocabulary, listening activities, a second part with the same subcategories followed by language and culture contents, grammar description, exercises, vocabulary practice and cultural curiosities. At the end of the book glossaries, maps and tables are included. This textbook also has a separate workbook for practicing character writing.

Studiamo il cinese! (2019) contains a set of nine modules of three units (1, 2, 3); each one is introduced by an illustrated comic strip presenting the language input, followed by the sections introduction activities, grammar, listening and pronunciation, writing and communication. At the end of each unit, the book presents a section of cultural contents. In the final part, it includes a language support section featuring grammar, language games, illustrated vocabulary, glossaries, and pinyin and radicals' tables. The textbook has a separate character workbook.

Finally, *Parliamo Cinese* (2018) consists of seven units, with each unit consisting of four lessons (a, b, c, d). Each lesson presents introduction activities, text or dialogue, a section on grammar, and exercises, which that are presented in three subcategories: listening and pronunciation, reading and writing, and communication. At the end of each unit a review and testing section, as well as a cultural part, are included. Glossaries at the end of the course book present a vocabulary index, communicative functions, grammar structures and a *pinyin* table.

5.2 Procedures

The analysis was developed in the following steps:

- selecting the grammar structure;
- reading all textbooks and selecting chapters where the grammar structure is presented;
- selecting the activities;
- numbering the activities according to their type.

To investigate textbooks' approach to grammar instruction, the author arbitrarily selected one grammar structure: interrogative construction. There are several ways in Chinese grammar to construct interrogative sentences: interrogative pronouns, alternative questions, affirmative-negative questions, or particles (Paternicò *et al.* 2021; Li and Thompson 1981). These different ways generally are not presented together, thus, this grammatical structure was chosen as it allows the analysis of different chapters of each textbook. Upon reading all textbooks, the author selected only those chapters that present the description of how to make interrogative sentences in the grammar section.

As the purpose of the investigation is to observe the approach to grammar instruction, only the exercises targeting the comprehension, or the production of the chosen grammar structure described in each chapter, were analysed. All the exercises for practicing and learning phonetic and character writing, such as pronunciation, matching character-pinyin, radical recognition, and listening recognition activities, were excluded from the analysis. General vocabulary activities such as matching images to one character and fill in the blank exercises focusing only on vocabulary knowledge were also not considered. As per method of classification, the study analyses the type and the number of activities referring to the grammar structure of Chinese interrogative sentences, according to the classification of input-based and output-based activities, based on Ellis (2002). *Input-based* activities are those that require students to do something to demonstrate they have comprehended or noticed the target form, without using the target language (e.g., selecting, matching, reordering, finding the intruder, etc.). *Output-based* activities are exercises in which students must produce the target form (e.g., fill in the blank or complete the sentence, reordering words to create a sentence, translation from Italian to Chinese, transformation, writing sentences or text, answering questions, creating a dialogue, etc.). *Input-based* activities were further subdivided according to their type in the categories of matching, selecting, and translation from Chinese to Italian. *Output-based* activities were classified according to Aski's (2003) typology, which includes four subcategories: *mechanical drills* in which students do not need to understand the meaning but only replicate a given pattern or form (e.g., transformation); *meaningful drills*, meaning-form oriented exercises with one possible solution (e.g., answer the question or produce a question for a given answer); or *communicative drills* that have more than one possible answer but still provide limited occasions for meaning negotiation and language construction (e.g., conducting an interview with given questions, asking one personal question and receiving the answer, role-play). In the present study these two types of drills are included in the same category named *meaningful drills*, as in Vold (2017). Finally, the category of *communication language practice* includes exercises in which learners produce the target language by simulating real communication to obtain previously unknown information (e.g., task-based activities).

6. Findings

Table 1 presents the results of the analysis of grammar instruction activities in the five analysed textbooks. Findings are shown according to the *input-based* and *output-based* classification of activities and to the subcategories *matching*, *selecting*, *translation* (from Chinese to Italian), and *mechanical drills*, *meaningful drills* and *communicative language practice*.

Findings show that the overall number of output-based activities is three times larger than the number of input-based activities. Each textbook presents a bigger quantity of output-based activities compared to input-based, and in all textbooks the highest number of exercises is included in the category *meaningful drills*.

Textbook	**Number of interrogative structure activities**	**Input-based activities**			**Output-based activities**		
		Matching	Selecting	Translation	Mechanical drills	Meaningful drills	Communicati-ve language practice
Comunicare in cinese 我会说中文 2021	35	4	2	0	3	25	1
Lingua cinese. Corso elementare 汉语初级教程 2019	13	1	1	0	1	10	0
Shangke le! 上课了！ 2022	33	3	11	0	1	18	0
Studiamo il cinese 咱们学中文吧！ 2019	54	4	1	2	12	35	0
Parliamo cinese 我们说汉语 2018	64	16	0	1	2	43	2
Total		28	15	3	19	131	3
		46			**153**		

Table 1. Distribution and types of interrogative structures activities in the analysed textbooks.

Comunicare in cinese contains four *matching* activities, such as matching questions to answers or two parts of a long sentence, and two *selecting* exercises; for example, activity N1 (see Appendix) requires learners to choose the correct position to collocate the interrogative pronoun, providing the opportunity to focus on the grammar form within a sentence. A different input activity (N2) asks students to complete the dialogue by placing the letters associated with the selected word. This book seems to prefer the use of exercises that employ short texts or dialogues over the use of discrete sentences, thus favouring the comprehension of the form in a contextualized situation. However, the textbook contains mostly production exercises, such as answering questions about a text, writing questions using given answers or images, or completing a dialogue with provided words. Some communicative drills require students to answer questions about personal information or refer to provided data (N3). This type of drill may be changed into real communication practice by including tasks, for example asking students to select one person among the images and describe them, then asking other students to identify the selected person. *Mechanical drills* included in this book are transformation activities, such as asking students to transform an affirmative or negative sentence into an interrogative. The book has one exercise that can be classified as *communicative language practice* (N4). It presents a movie poster and asks students to play a dialogue using a minimum of five sentences. The teacher may change this exercise to increase its communicative effectiveness and enable students to simulate real communication, for example by creating an info gap activity: one student sees the information on the movie poster, while the other student sees the information on a weekly agenda; they need to communicate to figure out whether they can go watch the film. In this way learners will employ the interrogative form and negotiate the meaning for real communicative purposes.

Lingua cinese. Corso elementare only contains two input-based exercises for practicing the analysed grammar structure. Activity N5 is a *selecting* activity that requires students to select the correct answers for a provided question. One *matching* activity (N6) asks students to match two sentences according to their meaning. Although these exercises allow students to focus on the grammar form of interrogative constructions, there are few compared to the production-based activities. The textbook presents 11 output-based activities: one is a *mechanical drill*, while the other activities are included in the *meaningful drills* category. These consist of four types of exercises: reordering provided words to create a sentence (N7), translating a sentence into Chinese (N8), answering questions about a text, and writing a sentence employing a provided word.

Shangke le! has a decent amount of input exercises compared to the total. In this textbook most of the input-based activities are in the homework section. According to the PI model, grammar instruction should move from constant exposition to input to focus on production, thus many input-based activities should be placed in the initial phase of exposition to grammar. This book includes 11 *selecting* exercises, many which ask the student to find the error in the sentence, or to select the word that is not needed in the sentence (N9). Like the above input-based selecting activities, in the homework section there are some odd-one-out exercises (N10). In contrast to the activities that employ sentences, these activities only focus on form, while according to the PI, to lead to language acquisition, each activity should keep the focus on meaning (Wong and VanPatten 2003). In the in-class exercises, each lesson includes a reordering dialogue task with provided sentences and, among *matching* activities, many include matching sentence halves (N11). There are no communication language practice activities, and the majority of output-based activities are *meaningful drills*. For example, filling in sentences with the correct pronoun or conjunction based on the meaning, answering questions (N12), writing questions using a given answer, and sentence translation.

Studiamo il cinese! presents one *selecting* and four *matching* activities, including matching text and image or a Chinese sentence with the Italian version (N13). Among the translation activities N14 requires learners to first observe and then translate some sentences into Italian. These based-on-input exercises support learners' language acquisition; however, the higher number of exercises examined in this book are still in the output-based category. For example, there are 11 *mechanical drills* which ask students to repeat the same question to all their classmates and collect their answers, to answer questions according to a given pattern and provided words, or transformation activities. One is a game (N15) that asks students to change their surnames and introduce themselves to the class; in turn, students ask each other questions about the classmates' new identity and when a mistake is made, the roles change. There are no communicative language practice activities and 35 *meaningful drills*. Some of them require students to write questions using given information, or answer questions about themselves, a text, or an image (N16). This book also has sentence translation exercises. There are different types of communicative drills, like creating a dialogue assuming the identity of a character using provided information, or role play activities. One requires learners to use a given set of questions to interview other students pretending to be a celebrity.

Parliamo cinese contains only one L2 to L1 *translation* activity and 16 *matching* exercises; these ask students to match a picture with a written sentence, or to match answers and questions (N17). Such input activities allow students to focus on the link between form-meaning; however, like in *Shangke le!*, in this textbook most of the input-based exercises are also situated in the homework section, thus intended for the practice of the acquired knowledge more so than for the introduction of the grammatical form. Few input-based activities are in the section that introduces grammar: for example, N18 requires learners to select the correct interrogative form for completing the dialogue by placing the associated letters. These activities only appear in the first units of the book, which present gradual training of the writing skill, as is typical of Chinese teaching. Such input activities should, however, also be offered later, allowing learners to focus on the use of grammatical form in exercises separated from those focusing on writing production. There are a total of 43 *meaningful drills*, many which require students to complete discrete sentences or dialogues using interrogative grammar forms. Other examples of output-based activities are those in which students are asked to write questions for the provided answers or images, or to translate sentences. The book includes communicative drills promoting the speaking skill that can be transformed into real communicative practice. For example, one exercise asks to describe the family in a provided image. The teacher can add one other image or more, asking student A to choose one, while student B must ask questions to guess which is the family that student A has selected. This textbook contains two *communicative language practice* activities for the analysed grammar form. The first is an information gap activity (N19) in which students work in pairs; they first fill in their daily school schedule, then each student must complete the other one's schedule by asking questions. In the second information gap activity (N20) students look at two different sets of sentences that are the reverse of one another; student A's text contains the information to complete student B's text and vice versa. In each sentence, some parts are missing, and learners must ask questions to complete their text.

7. Discussion

Our findings have shown that the five analysed textbooks emphasize output-based activities for grammar instruction. With reference to the RQ1, these results indicate that there is a large discrepancy between the number of input and output activities. In addition, the two textbooks that present a bigger quantity of input-based activities include most of these types of exercises in the homework section. But exercises based on input are intended to be in the initial phase of the acquisition process, in order to help learners process the input and convert it into intake, and then lead the development of the FL system, while in final sections, more language production activities should support the development of fluency and automaticity. These data seem to confirm that the analysed textbooks do not fully recognize the role of input in the grammar acquisition process and the grammar instruction adopted remains traditional (RQ2). Shintani *et al.* (2013) suggests that this emphasis on production should be motivated by the goal of foreign language teaching, that is to "enable students to speak and write in the L2, and it would seem self-evident that this can best be achieved by having students speak and write" (2013: 1). The focus on output-based exercises may also be due to the role that active production has for evaluating the success of a lesson; in particular when the communicative approach is adopted, students' participation is measured in terms of production (Shintani *et al.* 2013: 1). Another reason why grammar teaching is still so traditionally based on output activities may be attributed to the resistance of the publishing industry to take the risk of issuing innovative teaching materials that might not be well received by instructors, resulting in the preference of traditional instructional procedures (Fernández 2011). Given the crucial role of the input for language acquisition, input-based activities need to be integrated. As suggested by some scholars (White and Farinelli 2014; DeMil 2013), if there is a lack of input-based activities in teaching materials, instructors should integrate these types of exercises to support the output activities already present in textbooks. Providing supplementary activities to be completed at home prior to class and using class time to develop communicative activities may be an asset.

Our results also reveal that most output-based activities are meaningful drills. These types of drills, although focusing on the relationship between form and meaning and creating short and simple communicative occasions, do not require any meaning negotiation nor interpretation of other speakers' messages. We only counted a total of three communicative language practice activities in the analysed textbooks, which is too small a quantity. Interactional activities enable learners to create true communication to obtain previous unknown information. These exercises, which support learners in exploring form-meaning relationships and participating in language knowledge construction, need to be supplemented. According to these findings, we recommend that textbooks integrate more input-based and more communicative activities to support students' acquisition of CFL.

Finally, although the activities identified in the books are mostly output-based, mechanical drills were reduced in favour of production exercises that focus on meaning-form relationships. In addition, textbooks present explicit grammar description, which is common to both PI and Traditional Instruction models (TI), but do not follow the typical TI sequence, which moves from mechanical drills to meaningful drills, and then to communicative drills (White and Farinelli 2014). In fact, they seem to offer a mixed sequence of activities. This indicates that in Italian CFL teaching materials some attempts at change have already been applied. Hopefully, they will be guided by the SLA findings in how they follow a sequence of activities that reflects the acquisition process and open the path to the creation of more effective teaching materials.

8. Conclusion and limitations

Grammar instruction is a key issue in second language acquisition research, with instructors and researchers constantly focusing their attention to find the most effective way to teach and learn grammar. Because textbooks are also central for grammar instruction in the FL classroom, this study has examined to what extent the five most recently published CFL textbooks for Italian-speaking learners integrate SLA findings, supporting teachers' practice and learners' language acquisition. The newest issues of teaching materials for teaching CFL in Italy, although showing a minor number of input-based activities compared to output-based, reveal signs of change from the traditional way of teaching grammar. In order to facilitate students' language acquisition, textbooks need to be changed to include more input-based activities, which draw learners' attention to the FL grammar form for meaning.

One limitation of a study examining teaching materials is to ignore how teachers employ such materials in the FL classroom, including supplementary online resources. Some instructors might integrate textbook activities with input-based exercises or communicative tasks needed for language acquisition; certainly, if the textbooks themselves already contain such content, it would not be the teachers' job to create new activities, so they could devote their energy to students' other needs. A second limitation is that this analysis only considers chapters targeting the grammatical structure of the interrogative construction and considering more grammar items may achieve different results. It is hoped that this analysis, and the overview of research on the role of input in second language acquisition, will support teachers in creating supplementary input-based activities and textbook authors for forthcoming publications.

References

Aski, J. (2003) "Foreign Language Textbook Activities: Keeping Pace with Second Language Acquisition Research". *Foreign Language Annals*, 36(1): 57-65.

Benati, A. (2017) "The Role of Input and Output Tasks in Grammar Instruction: Theoretical, Empirical and Pedagogical Considerations". *Studies in Second Language Learning and Teaching*, 7(3): 377-396.

Biondi, M., H. Wang and Y. Zhao (2019) *Studiamo il cinese!* 咱们学中文吧!. Recanati, Eli edizioni.

Brown, D. (2014) "The Power and Authority of Materials in the Classroom Ecology". *The Modern Language Journal*, 98: 658-661.

Bulfoni, C. and X. Sun (2019) *Lingua cinese. Corso elementare* 汉语初级教程. Milano, Unicopli.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, U.K., Press Syndicate of the University of Cambridge.

DeMil, A. (2013) "Communicative or Communication: What your Textbook is Teaching". *International Journal of Instructional Technology and Distance Learning*, 10(6), 23-36.

DeMil, A. and J. Aubrey (2013) "Beginning French Textbooks: Are They Teaching Communication?". *Florida Foreign Language Journal*, 9(1): 19-29.

Didi-Ogren, H.H. and J.J. Goebel (2008) "The Role of Input in JFL University-Level Teaching Materials: An Examination of Three Widely Used Textbooks". *Japanese Language and Literature*, 42(2): 439-470.

Ellis, R. (2002) "Methodological Options in Grammar Teaching Materials", in E. Hinkel and S. Fotos (Eds.) *New Perspectives on Grammar Teaching in Second Language Classrooms*. Mahwah (NJ), Lawrence Erlbaum Associates: 155-179.

Fernández, C. (2011) "Approaches to Grammar Instruction in Teaching Materials: A Study in Current L2 Beginning-Level Spanish Textbooks". *Hispania*, 94(1): 155-170.

Gabbianelli, G. (2020) "An Analysis of Grammar Instruction in Beginner-Level Foreign Language Chinese Textbooks". *Confuci Acadèmic Journal*, 1: 12-21.

Garton, S. and K. Graves (2014) "Identifying a Research Agenda for Language Teaching Materials". *The Modern Language Journal*, 98(2): 654-657.

Jean, G. and D. Simard (2011) "Grammar Learning in English and French L2: Students' and Teachers' Beliefs and Perceptions". *Foreign Language Annals*, 44(4): 465-492.

Krashen, S. (1985) *The Input Hypothesis: Issues and Implications*. Beverly Hills (CA), Laredo Publishing Company.

Krashen, S.D. and T.D. Terrell (1983) *The Natural Approach: Language Acquisition in the Classroom*. New York, Pergamon Press.

Langé, G. (2018) "L'insegnamento di lingua e cultura cinese nelle scuole italiane", in A. Brezzi and T. Lioi (Eds.) *La didattica del cinese nella scuola secondaria di secondo grado. Esperienze e prospettive*. Roma, Sapienza Università Editrice: 21-32.

Larsen-Freeman, D. (2009) "Prediction or Retrodiction? The Coming Together of Research and Teaching", in K. Losey and C. Pearson (Eds.) *Spotlight on Re-search: A New Beginning. The Selected Proceedings of the 2008 MITESOL Conference*. Raleigh (NC), LuLu Press: 5-16.

Larsen-Freeman, D. (2015) "Research into Practice: Grammar Learning and Teaching". *Language Teaching*, 48(2): 263-280.

Li, C.N. and S.A. Thompson (1981) *Mandarin Chinese. A Functional Reference Grammar*. Berkeley, University of California Press.

Masini, F., C. Romagnoli, T. Zhang, and Y. Chang (2021) *Comunicare in cinese* 我们会说汉语, vol. 1, Milano, Hoepli.

Masini, F., T. Zhang T., G. Gabbianelli, and R. Wang (2018) *Parliamo cinese* 我们说汉语, vol. 1. Milano, Hoepli.

Ministero dell'Istruzione, dell'Università e della Ricerca (2016) Direzione generale per gli ordinamenti scolastici e per la valutazione del sistema nazionale di istruzione, Ufficio VIII - Internazionalizzazione del sistema educativo di istruzione e formazione, *Il sillabo della lingua cinese. Quadro di riferimento unitario per l'insegnamento della lingua cinese nella scuola secondaria di secondo grado.* Roma, MIUR.

Nassaji, H. and S. Fotos (2004) "Current Developments in Research in the Teaching of Grammar". *Annual Review of Applied Linguistics*, 24: 126-145.

Paternicò, L.M., V. Varriano and H. Tian (2021) *Grammatica della lingua cinese*. Novara, Utet.

Scibetta, A. (2018) "Il Sillabo per l'insegnamento del cinese nella scuola secondaria di II grado: un punto di partenza per la didattica". *Lend. Lingua e Nuova Didattica*, 3: 34-44.

Shintani, N., S. Li and R. Ellis (2013) "Comprehension-Based Versus Production-Based Grammar Instruction: A Meta-Analysis of Comparative Studies". *Language Learning*, 63: 296-329.

Swain, M. (1995) "Three Functions of Output in Second Language

Learning", in G. Cook, and B. Seidelhofer (Eds.) *Principle and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson*. Oxford, Oxford University Press: 125-144.

Swain, M. (2005) "The Output Hypothesis: Theory and Research", in E. Hinkel (Ed.) *Handbook of Research in Second Language Teaching and Learning*. Mahwah (NJ), Lawrence Erlbaum Associates: 471-483.

Tomlinson, B. (2012) "Materials Development for Language Learning and Teaching". *Language Teaching*, 45(2): 143-179.

Trentin, G. (2020) *Didattica con e nella rete: dall'emergenza all'uso ordinario*. Milano, Franco Angeli.

VanPatten, B. (1990) "Attending to Content and Form in the Input: An Experiment in Consciousness". *Studies in Second Language Acquisition*, 12: 287-301.

--- (1996) *Input Processing and Grammar Instruction in Second Language Acquisition*. Westport (CT), Greenwood.

--- (2003) *From Input to Output: A Teacher's Guide to Second Language Acquisition*. New York, McGraw-Hill.

--- (Ed.) (2004) *Processing Instruction: Theory, Research, and Commentary*. Mahwah (NJ), Lawrence Erlbaum.

Varriano, V., P. Velardi, M. Cigliano, B. He and T. Xu (2022) *Shangke le!* 上课了!. Milano, Mondadori.

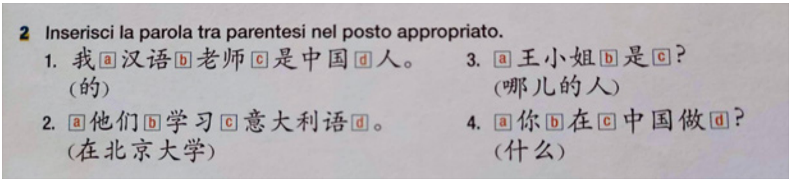
Vold, E. (2017) "Meaningful and Contextualised Grammar Instruction: What Can Foreign Language Textbooks Offer?". *The Language Learning Journal*, 48(1):1-15.

White, J.P. and C. Farinelli (2014) "Crossing the Bridge of Communication: An Analysis of ESL Materials and a Hybrid Solution". *International Journal of Instructional Technology and Distance Learning*, 11(12): 3-18.

Wong, W. and B. VanPatten (2003) "The Evidence is IN: Drills are OUT". *Foreign Language Annals*, 36: 403-418.

Appendix

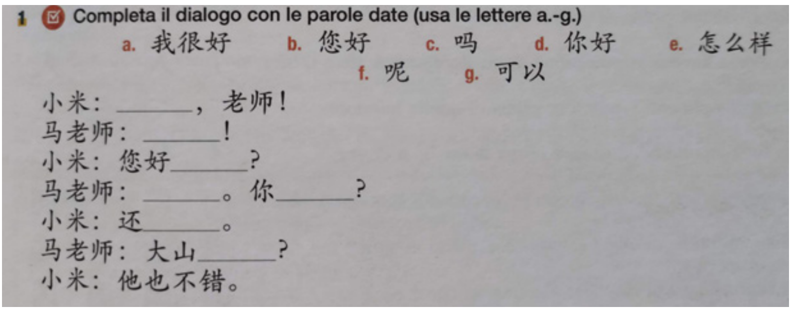
Examples of grammar activities from the four analyzed textbooks.



2 Inserisci la parola tra parentesi nel posto appropriato.

1. 我[a]汉语[b]老师[c]是中国[d]人。
   (的)
2. [a]他们[b]学习[c]意大利语[d]。
   (在北京大学)
3. [a]王小姐[b]是[c]?
   (哪儿的人)
4. [a]你[b]在[c]中国做[d]?
   (什么)

N1. Input-based, selecting activity (*Comunicare in cinese*, 2021: 40).



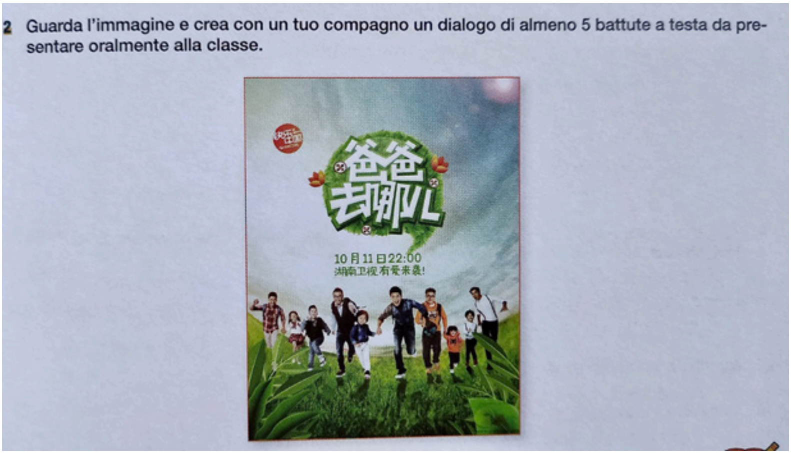
1 Completa il dialogo con le parole date (usa le lettere a.-g.)

a. 我很好 b. 您好 c. 吗 d. 你好 e. 怎么样 f. 呢 g. 可以

小米：______，老师！
马老师：______！
小米：您好______？
马老师：______。你______？
小米：还______。
马老师：大山______？
小米：他也不错。

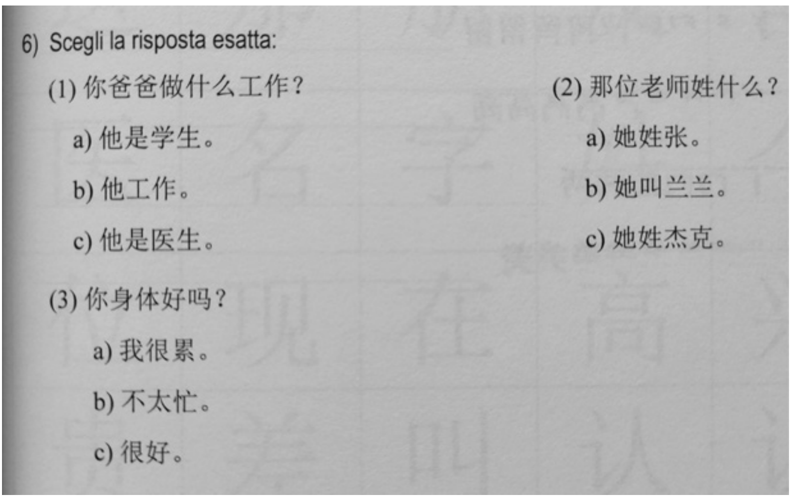
N2. Input-based, selecting activity (*Comunicare in cinese*, 2021: 19).

N3. Output-based, meaningful drill (*Comunicare in cinese*, 2021: 39).



2 Guarda l'immagine e crea con un tuo compagno un dialogo di almeno 5 battute a testa da presentare oralmente alla classe.

N4. Output-based, communicative language practice (*Comunicare in cinese*, 2021: 95).



6) Scegli la risposta esatta:

(1) 你爸爸做什么工作？
a) 他是学生。
b) 他工作。
c) 他是医生。

(2) 那位老师姓什么？
a) 她姓张。
b) 她叫兰兰。
c) 她姓杰克。

(3) 你身体好吗？
a) 我很累。
b) 不太忙。
c) 很好。

N5. Input-based, selecting activity (*Lingua cinese. Corso elementare*, 2019: 59).

8) Scegli tra le frasi A, B, C, D, E quelle che, in base al senso, si collegano alle frasi 1, 2, 3, 4, 5:

A 很晚了，他在哪儿？

B 给我一杯水吧！

C 您还要别的吗？

D 太麻烦您，谢谢！

E 1美元可以换六块三毛五人民币。

(1) 他还在办公室工作。 (A) *esempio*

(2) 请问，您要咖啡还是红茶？ ( )

(3) 不客气，再见！ ( )

(4) 今天美元的牌价是多少？ ( )

(5) 不用了，我就要这些东西。 ( )

N6. Input-based, matching activity (*Lingua cinese. Corso elementare*, 2019: 223).

5) Riscrivi le frasi collocando le parole nella posizione grammaticalmente corretta:

(1) 手续费 换 要 钱 不 要 请问 ？

____

(2) 面包 我 两 个 要 来 一 糖 公斤 再 ， 。

____

(3) 瓶 一 钱 多少 这 啤酒 种 ？

____

(4) 吗 要 您 别的 还 ？

____

(5) 不 喜欢 太 我 这 颜色 种 蓝 。

N7. Output-based, meaningful drill (*Lingua cinese. Corso elementare*, 2019: 222).

6) Traduci in cinese le seguenti frasi (in caratteri e in *pinyin*):

(1) Quanti anni hanno tuo padre e tua madre?

____

____

(2) Dove sei? Perché non mi scrivi un SMS?

____

____

(3) Martedì sarà il compleanno di Davide. Gli comprerò un libro.

____

____

(4) Tutte le sere guardo la televisione. E tu?

____

____

(5) Perché non vieni a trovarmi? Scusa, sono troppo occupato.

____

____

N8. Output-based, meaningful drill (*Lingua cinese. Corso elementare*, 2019: 89).

15. Elimina la parola di troppo in ciascuna frase.

1	你	是	哪	国	人	吗？
2	爸爸	也	是	说	汉语。	
3	我们	都	妹妹	学习	中文。	
4	你们	上	高	几	吗？	
5	你	是	中国	哪	哪里	人？

N9. Input-based, selecting activity (*Shangke le!*, 2022: 67).

20. **Trova l'intruso.**

1. 英语 汉语 意大利语 外语
2. 节 牙 门 口 个
3. 星期 明白 年 月 天
4. 牙刷 怎么 什么 哪 谁
5. 用 从 给 课 到

N10. Input-based, selecting activity (*Shangke le!*, 2022: 302).

25. **Collega il segmento di frase della colonna A con il segmento di frase della colonna B per crearne una completa di senso compiuto, poi leggile ad alta voce.**

A
你打电话
我还
您认识
今天的作业
意大利人

B
也很多。
找谁?
帅不帅?
想加他微信。
不认识马老师?

N11. Input-based, matching activity (*Shangke le!*, 2022: 298).

21. **Parliamo!**

Rivolgi a un tuo compagno le seguenti domande e appunta le risposte. Poi scambiatevi i ruoli.

1. 你是哪儿的人?
2. 你住在哪儿?
3. 你上高几?
4. 你说汉语吗?

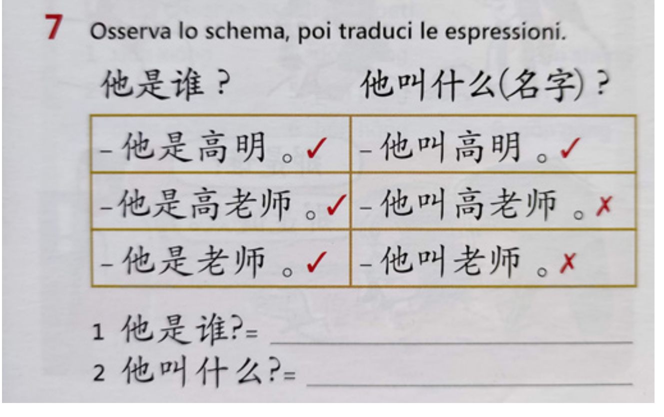
N12. Output-based, meaningful drill (*Shangke le!*, 2022: 68).

3 1.41 Ascolta e abbina le espressioni alla traduzione corretta.

1 ☐ 我叫琳达。
2 ☐ 你姓什么?
3 ☐ 你叫什么?
4 ☐ 大家好!
5 ☐ 我姓陈。
6 ☐ 欢迎!

a Come ti chiami?
b Come ti chiami di cognome?
c Mi chiamo Linda.
d Di cognome mi chiamo Chen.
e Ciao a tutti!
f Benvenuta!

N13. Input-based, matching activity (*Studiamo il cinese!*, 2019: 35).



7 Osserva lo schema, poi traduci le espressioni.

他是谁?	他叫什么(名字)?
-他是高明。✓	-他叫高明。✓
-他是高老师。✓	-他叫高老师。✗
-他是老师。✓	-他叫老师。✗

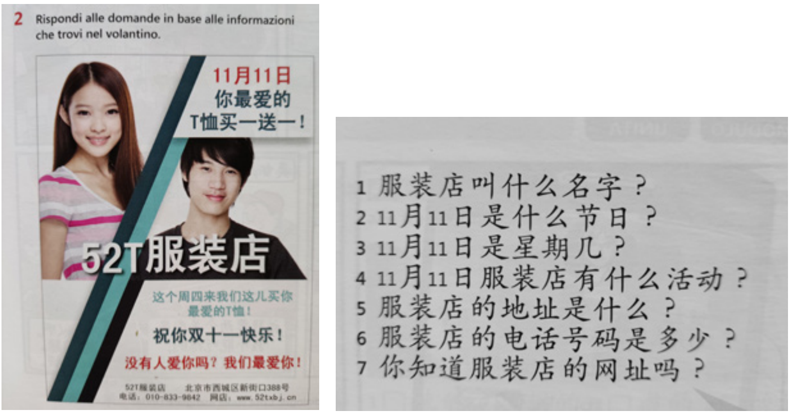
1 他是谁?= __________
2 他叫什么?= __________

N14. Input-based, translation activity (*Studiamo il cinese!*, 2019: 41).

沟通能力 Abilità in gioco

1 Gioco. Tutti i partecipanti assumono un nome e un cognome di fantasia (può essere il nome di un personaggio noto) e, a turno, si presentano usando i verbi 叫 e 姓. Quando tutti si sono presentati, un partecipante estratto a sorte ne sceglie un altro e lo interroga, in cinese, sul nome e il cognome degli altri. Al primo errore, i ruoli cambiano e starà all'interrogato scegliere un altro partecipante a cui fare le domande.

N15. Output-based, meaningful drill (*Studiamo il cinese!*, 2019: 225).



2 Rispondi alle domande in base alle informazioni che trovi nel volantino.

1 服装店叫什么名字？
2 11月11日是什么节日？
3 11月11日是星期几？
4 11月11日服装店有什么活动？
5 服装店的地址是什么？
6 服装店的电话号码是多少？
7 你知道服装店的网址吗？

N16. Output-based, meaningful drill (*Studiamo il cinese!*, 2019: 133).

10 HSK 选出正确的问题或答案。**Collega le domande/risposte con le risposte/domande giuste.**

1. 为什么你今天晚上不去看电影？ ☐
2. 你的电话号码是多少？ ☐
3. 我们一起去踢足球，好吗？ ☐
4. 明天晚上我们一起去游泳，好吗？ ☐
5. 明天晚上你想去哪儿？ ☐

a. 对不起，我不会踢足球。
b. 我和我姐姐一起去看电影。
c. 好，我很想去游泳。
d. 因为我要去看朋友。
e. 我的手机号码是3388210457。

N17. Input-based, matching activity (*Parliamo cinese*, 2018: 233).

6. **Completa il dialogo con le parole date e scrivi in pinyin.**

a. 怎么样 zěnmeyàng b. 很 hěn c. 吗 ma

1. 你好☐?
   Nǐ hǎo ____________________?
2. 还可以。你☐?
   Hái kěyǐ. Nǐ ____________________?
3. 我☐好。
   Wǒ ____________________ hǎo.

N18. Input-based, selecting activity (*Parliamo cinese*, 2018: 11).

会话 Comunicare

6. **In coppia. Formula le domande e completa il programma delle lezioni del tuo compagno.**

▶ 明天第一节是什么课?

Studente A

课表	
你	同学
历史	
地理	
地理	
汉语	
数学	

Studente B

课表	
你	同学
英语	
历史	
法语	
数学	
语文	

N19. Output-based, communicative language practice (*Parliamo cinese*, 2018: 103).


6. In coppia. Fai domande a un compagno usando 哪国, 什么 o 哪儿 per completare la tabella.

Studente A

她叫__________
她是中国人。

她叫爱玛。
她是__________

他叫__________
他是罗马人/
他是意大利人。

Studente B

她叫小云。
她是__________

她叫__________
她是威尼斯人/
她是意大利人。

他叫路卡(Lùkǎ)。
他是__________

N20. Output-based, communicative language practice (*Parliamo cinese*, 2018: 41).

DESIGNING AND COMPILING THE WRITTEN SUB-CORPUS OF THE BIMODAL ITALIAN LEARNER CORPUS OF CHINESE (BILCC): METHODOLOGICAL ISSUES

Alessia Iurato

Ca' Foscari University of Venice; Bremen University

1. Introduction[1]

The definition of what can be considered a 'learner corpus' has been a matter of debate since early studies in Learner Corpus Research (LCR) in the late 1980s (Meunier 2021; Tracy-Ventura and Miles 2015). When the field of LCR emerged, it aligned itself with the methodology and theoretical framework of corpus linguistics. It therefore adopted corpus linguistics' definition of what is intended by a (learner) corpus, i.e., a collection of naturally occurring, authentic, continuous, spontaneous spoken or written (learner) language samples (Callies and Götz 2015; Meunier 2021). Many corpus linguists consequently refused to consider a collection of learner data obtained through

[1] I would like to thank the two anonymous reviewers for their insightful comments and suggestions, and Serena Zuccheri for her patience and commitment as editor of this volume. I would also like to thank Bianca Basciano for reading a draft of this paper; naturally, all mistakes and errors are solely my responsibility.

elicitation methods as a 'learner corpus' because it lacked spontaneity and authenticity (Gilquin and Gries 2009; Lozano and Mendikoetxea 2013; Sinclair 2005). Later, several studies (Gilquin 2021; Norris and Ortega 2003; Tracy-Ventura and Myles 2015) showed that in the field of Second Language Acquisition (SLA), when the object of study is a rare structure, construct underrepresentation is a problematic issue that frequently occurs in general-purpose learner corpora. Tracy-Ventura and Myles (2015: 60) argued that "it is imperative to ensure the corpus contains multiple examples of the feature(s) under investigation" to meet SLA needs. Although the definition of learner corpus remains a hot topic in LCR and SLA, today researchers in both fields agree that two types of learner corpora can exist: 1) corpora as collections of authentic and natural data, i.e., "naturally occurring samples" (Granger 2012: 8); 2) corpora as the result of open-ended tasks (e.g., picture description, role-play) allowing learners to choose their own wording and whatever linguistic resources they want – or are able – to use. This type of learner data is what Granger (2012: 8) calls "clinically elicited samples".

Today, most of the available corpora collect data from L2 English learners (Gráf 2017). Although LCR has also spread to the Chinese context since the 1990s and an increasing number of L2 Chinese corpora are being compiled[2], there is a lack of L2 Chinese corpora that collect data from learners whose L1s are European languages (Iurato 2022a; Zhang and Tao 2018). In the Italian context, for instance, the growing number of students and the widespread interest in Chinese language teaching (Romagnoli and Conti 2021) have not been matched by an equally flourishing research on corpus compilation to support research on the acquisition of L2 Chinese by Italian learners (Iurato 2022a). The compilation of a learner corpus is a challenging issue due to the strict criteria that need to be observed for corpus design and data collection (Castillo Rodríguez *et al.* 2020; Dutra and Gomide 2015; Lozano 2021). This paper addresses these issues and presents the methodological steps necessary for the compilation of a written Italian

[2] The purpose of this paper is not to discuss the development of Chinese LCR, nor to provide an overview of L2 Chinese learner corpora. For an overview of these issues, see Iurato (2022a), Iurato (2022b), Xu (2019), and Zhang and Tao (2018).

learner corpus of L2 Chinese. The aim of this paper is threefold: first, illustrating the methodological stages involved in the compilation of a target-oriented corpus when the object of study is a specific (rare) structure that may be underrepresented in general-purpose learner corpora; second, describing a well-structured methodology grounded in LCR for compiling Italian learner Chinese corpora that can be reproduced in future studies, given the growth of L2 Chinese studies in the Italian context and the lack of Italian learner Chinese corpora; three, promoting the standardization of the proposed methodology, as the compilation of the presented corpus implements rigorous design principles and attempt to address some of the gaps in past LCR studies.

First, the paper will discuss the rigor and transparency required in the compilation process, explaining the criteria to be applied in the corpus design. Second, it will present the case study of a written L2 Chinese corpus specifically designed to explore the pragmalinguistic knowledge of the "*shì* 是*...de* 的 focus proper cleft" construction (Paul and Whitman 2008: 424) in L1 Italian learners' production. Here, the corpus is intended both as the result of open-ended tasks (Gilquin 2021; Tracy-Ventura and Myles 2015) and as a collection of contextualized data produced by L2 learners (Callies and Götz 2015). Corpus features, corpus typology, as well as environment, learner and task variables will be described. The collection procedure will also be explained. This corpus, presented as a case-study in the present work, constitutes the written sub-corpus of a larger project: the 'Bimodal Italian Learner Corpus of Chinese' (BILCC). It is named 'bimodal' because it collects two types of data (written and spoken data) from L1 Italian learners of L2 Chinese. The strength of such a bimodal mode corpus is that it allows us to get a deeper insight into the L1 Italian learners' pragmalinguistic knowledge and acquisition process of L2 Chinese language from different perspectives. Further details on BILCC and the design and collection of the written sub-corpus will be explored in Sections 3 and 4.

2. Design criteria

A random collection of heterogenous learner data is not a learner corpus (Granger 2012). A learner corpus is compiled according to strict design criteria, and the usefulness of a learner corpus is directly proportional to the attention that has been paid to controlling the design criteria. These criteria primarily concern the participants and the task design, i.e., the two specific variables of learner corpora (Gilquin 2015; Meunier 2021). Careful selection, documentation, explanation, and justification of all criteria also increase "the likelihood that the resulting corpus is methodologically-sound" (Bell and Payant 2021: 56).

2.1 Learner corpus typology

Defining the corpus typology is the first step in designing a corpus. The typology of a learner corpus depends on several aspects, i.e., medium, size, text type, time of collection, target language (L2), learners' mother tongue (L1), and scope of collection.

2.1.1 Medium

Learner corpora can consist of written texts or phonetic/prosodic transcriptions of spoken discourse. The number of existing written L2 Chinese learner corpora is significantly higher compared to the number of oral corpora (Iurato 2022a; Iurato 2022b; Xu 2019; Zhang and Tao 2018). New types of corpora are multi-modal corpora (see, for example, Gao and Wang 2017; Huang 2018; Kong 2013), which usually contain collections of photo-pictorial elements, video, and speech recordings accompanied by transcriptions and gesture annotations. Multimodal corpora allow us to study how two or more modalities interface with one another in human communication.

2.1.2 Size

A distinction can be drawn between global and local learner corpora. Global corpora are large-scale projects and collect a vast amount of data from students from multiple universities/research centres; local corpora are small-scale projects and collect a minor amount of data from small groups of learners, who are usually both contributors and users of the corpus (Gilquin 2015). A further type of corpora is 'in-house learner corpora', which lie somewhere between global and local corpora. In this case, the contributors do not correspond to the users, but they come from the same population of learners (generally the same university) (Gilquin 2015).

2.1.3 Text type

Theoretically, any text type, also referred to as 'genre', may be represented in a learner corpus (Bell and Payant 2021). However, in practice, the two most common text types are argumentative essays for writing and informal interviews for speaking (Callies and Götz 2015; Granger 2012). This selection reflects the need to sample the least constrained types of production data (Granger 2012). Nonetheless, diversification in terms of textual genres is desirable. Recently, SLA and LCR researchers have attempted to include a variety of genres (i.e., task types) to ensure a balanced representation of learner interlanguage (see, for example, Campillo Llanos 2014; Lozano 2021).

2.1.4 Time of collection

Learner corpora can be cross-sectional or longitudinal. The former collect samples of learner production from different categories of learners at a particular point in time; the latter include data from the same learners produced at different stages in their development (Granger 2012; Meunier 2021). Quasi-longitudinal learner corpora (sometimes referred to as 'pseudo-longitudinal learner corpora') are also quite common; they contain data collected from learners at different proficiency levels at a single point in time (Granger 2012). In LCR and SLA, cross-sectional and quasi-longitudinal corpora are the most common, as they allow researchers to gather more data in a short period of time (Gilquin 2015). In pseudo-longitudinal corpora, the developmental stages of learners are classified according to external criteria, such as proficiency test or grade level. This can be problematic, as proficiency level is often assessed according to different parameters in different school systems, and it does not always reflect the actual learners' proficiency[3], especially if it is calculated on external unreliable variables (Tono 2003).

2.1.5 Target language

Learner corpora can be classified on the basis of the target language they sample. English is still the predominant target language, as re-

[3] For an in-depth discussion about learners' language proficiency assessment in LCR, see Leclercq *et al.* (2014) and Callies and Götz (2015).

vealed by the *Learner Corpora Around the World* database[4]. However, over the last few years, other L2s have gradually "joined the learner corpus bandwagon" (Granger 2012: 12). Most learner corpora are monolingual, as they contain data from only one target language, such as the *Jinan Chinese Learner Corpus* (JCLC; Wang *et al.*, 2015). On the other hand, a small but increasing percentage of learner corpora is multilingual, like the *Multilingual Corpus of Second Language Speech* (MuSSeL; Rubio *et al.* 2021), which collects texts produced in four languages: Chinese, French, Portuguese, and Spanish.

2.1.6 Learners' mother tongue (L1)

Mono-L1 learner corpora include data from learners from one and the same L1 background, i.e., a single L1 population (Gilquin 2015). Differently, multi-L1 learner corpora include data from learners from different mother-tongue backgrounds (Granger 2012), like the *Guangwai-Lancaster Chinese Learner Corpus*[5]. It is a collection of written and spoken data produced by learners of L2 Chinese from 80 different countries studying at the Guangdong University of Foreign Studies in China. Multi-L1 learner corpora are very useful for investigating the effect of L1 crosslinguistic influence.

2.1.7 Scope of collection

Learner corpora can be distinguished according to the purpose for which they are compiled. Commercial learner corpora, such as the *Cambridge Learner Corpus*[6], are compiled by publishers with the aim of creating learning materials based on learner outputs (Granger *et al.* 2015). Academic learner corpora are generally compiled by researchers interested in exploring learners' language use and interlanguage. However, if existing available corpora do not suit one's research purpose, or if there is a shortage of corpora from which to extract the data needed for a particular research, there remains the option of compiling one's

---

4 https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html (visited 2023/02/20).

5 https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fguangwai (visited 2023/02/20).

6 https://www.sketchengine.eu/cambridge-learner-corpus/ (visited 2023/02/20).

own local learner corpus (Giqluin 2015). The biggest advantage of such a bespoke corpus is that it is fully controllable (Millar and Lehtinen 2008).

2.2 Environment, learner and task variables

Three main variables play a role in the corpus compilation process: the environment in which the data are gathered, the learners whose performances are being collected, and the tasks that participants are asked to complete (Bell and Payant 2021; Gilquin 2015).

As for the environment, a learner corpus can be compiled in different linguistic contexts, and each linguistic context has different implications for the collection process and data analysis. For instance, in second language contexts, learners are exposed to the target language in daily activities, while in foreign language contexts opportunities for interaction in the target language are limited because the context of the common target language use is the classroom (Bell and Payant 2021). Furthermore, a distinction can be made between data collected in educational settings (at school/university) and in natural settings (mundane activities outside school/university). This distinction is particularly significant because second languages can be used in several varieties of contexts, but foreign languages can also be used outside educational settings (Gilquin 2015), for example to send e-mails to colleagues.

As for the learner variables, it is important to stress that the type and number of participants, the criteria, and the rationale for recruiting participants will affect the analyses of the data (Mackey and Gass 2021). Gathering and making available a rich set of metadata is therefore fundamental to increasing the rigor and transparency of learner corpora (Bell and Payant 2021; Tono 2003). This information can be obtained through the "learner profile questionnaire" (Gilquin 2015: 18), which collects: 1) personal information about the learner (e.g. age, gender, nationality, mother tongue, level of education, level of proficiency); 2) information about the learner's knowledge of other languages (e.g. additional language(s) studied and related level of proficiency, extensive experience of living abroad); 3) information about the learner's educational background (e.g. length of time studying the target language, universities and countries where the target language was studied, where he/she went to school and university). The learner profile questionnaire is accompanied by the informed consent form that learners are required to complete if they allow their data to be used for research purposes (Bell and Payant 2021).

As for the tasks[7], they can involve different variables, like timing constraints (the learner may have a limited time to write the text; timing can be controlled while performing computer-based tasks), availability of reference tools (grammar books, dictionaries), intertextuality (allowing or not the consultation of secondary sources such as articles, sample texts), computerization (writing by hand or using a computer) (Gilquin 2015). Topic (complex or sensitive themes to be discussed, for example) can also affect learners' performance (Mackey and Gass 2021). In addition, the researcher should take into account that if the composition that the participants are asked to write is part of an examination, the pressure to perform may alter the final results. Finally, motivation can affect the quality of learner data; motivated learners are more likely to complete the texts carefully and not to leave the paper blank. Participants should therefore be volunteers, and recruitment should be through general online advertisements, rather than through individual solicitations (Mackey and Gass 2021). Selecting the most appropriate tasks for data collection is also a crucial issue. There are innumerable types of tasks that can be created to compile a written corpus[8]. The choice of one task over another is highly dependent on the research questions outlined and may also be related to the theoretical framework within which the research is being developed (Mackey and Gass 2021).

3. The *Bimodal Italian Learner Corpus of Chinese* (BILCC)

The written sub-corpus that I will describe as a case-study in Section 4 is a portion of a larger ongoing corpus project: the compilation

---

[7] Here, I will only focus on variables and the design of written tasks. For an overview of oral tasks, see Faitaki and Murphy (2020), Prior (2018), Rolland *et al.* (2020).

[8] For an overview of tasks for written data elicitation, see Mackey and Gass (2021), Gass (2018).

of the *Bimodal Italian Learner Corpus of Chinese* (BILCC). BILCC, which is methodologically grounded in the LCR framework, corresponds to the concept of learner corpus as a collection of contextualized data produced by L2 learners. 'Bimodal' describes the corpus mode of BILCC. It is defined as 'bimodal' because the medium of the data it collects, i.e., one of the aspects of the corpus typology (see Section 2.1), has a dual nature that allows us to explore L1 Italian learners' pragmalinguistic knowledge of L2 Chinese from two different perspectives (written and spoken production). In fact, similar to the *Arabic Learner Corpus*[9] and the *YKI National Certificate Corpus*[10] listed in the CLARIN digital infrastructure[11] (Hinrichs and Krauwer 2014; Jong *et al.* 2020), BILCC comprises written and spoken data from Italian learners of L2 Chinese. Specifically, the spoken data consist of recordings of speech and their transcriptions. The data collection was conducted from December 2020 to March 2021. The written data were gathered from 103 BA (N=56) and MA (N=47) beginner (N=19), intermediate (N=50), and advanced (N=34) L1 Italian learners of L2 Chinese with an average age of 23, studying at the Ca' Foscari University of Venice. The spoken data were collected from 58 BA (N=30) and MA (N=28) students, divided in beginner (N=16), intermediate (N=21), and advanced (N=21) levels, who had previously completed the written tasks for the compilation of the written corpus. Since in LCR external proficiency measures are considered the only reliable criteria[12], the learners were grouped into three different proficiency levels according to their HSK language proficiency test scores[13]. The written sub-corpus of BILCC includes

---

9 https://www.arabiclearnercorpus.com/about-the-corpus-en (visited 2023/02/20).

10 https://metashare.csc.fi/repository/browse/the-national-certificates-corpus/944099dafccc11e18b49005056be118efc2ef6e1f96241b681c1d9bec0e9033a/ (visited 2023/02/20).

11 https://www.clarin.eu/ (visited 2023/02/20).

12 In LCR, proficiency based on external factors (e.g., institutional level, age) and self-assessment practices are considered unreliable and problematic (see Callies and Götz 2015; Leclerq *et al.*; Tono 2003).

13 Although the HSK language proficiency test has been criticised (Fu *et al.* 2013; Peng *et al.* 2021), it has been adopted to assess learners' Chinese language proficiency for practical reasons and because more reliable criteria have not been found.

53,248 Chinese characters, 38,793 word tokens, and 693 word types. The spoken sub-corpus consists of 25-hour recordings, while the corresponding transcriptions, which were manually performed, consist of 14,321 Chinese characters, 10,414 word tokens, and 285 word types. The bimodal corpus mode is one of BILCC's strengths, since most of the existing (L2 Chinese) corpora are mainly written, and spoken corpora are rare (see Iurato 2022a; Iurato 2022b; Zhang and Tao 2018). Another strength is that it includes multiple sources of data from the same group of learners, a design feature generally absent and thus highly encouraged in LCR (Tracy-Ventura *et al.* 2021). The corpus is accompanied by a control corpus of 30 L1 Chinese speakers for comparative purposes. The written data consist of 19,073 Chinese characters, 10,414 word tokens, and 285 word types. The spoken data consist of 7-hour recordings and the related transcription include 11,872 Chinese characters, 9,048 word tokens, and 113 word types. All participants voluntarily completed the tasks.

Another distinctive feature of BILCC is that it is a specific-purpose learner corpus compiled to explore the (explicit/implicit) pragmalinguistic knowledge of a particular syntactic structure: the Chinese "*shì* ... *de* proper focus cleft" (Paul and Whitman 2008: 424) with [V 的*de* O] order (henceforth the terms '*shì*...*de* cleft construction' and 'proper cleft' will be used interchangeably). It is used to highlight a specific information (agent, time, place, manner, instrument, cause, etc.) of a concluded event that is generally given as presupposition in the discourse (Cui and Sung 2021; Li 2008; Li and Thompson 1981; Lü 1982; Paris 1979; Paul and Whitman 2008), as illustrated in (1):

(1) 我们是昨天去的图书馆。

wǒmen	*shì*	*zuótiān*	*qù-de*	*túshūguǎn*
1PL	SHI	yesterday	go-DE	library

'It was yesterday that we went to the library.'

Descriptively, the *shì*...*de* proper cleft is signalled by two morphemes: *shì* in pre-verbal position marking the clefted element; *de* in post-verbal position between the verb and the object. Unfortunately, in the literature there is no consensus on the syntactic roles of *shì*[14] and *de*[15]. In the analysis of BILCC, following Cheng (2008), Hole (2011), Paul and Whitman (2008), Xu (2014), *shì* is considered a copula serving as focus marker. On the other hand, following Lü (1982), Shi (1994), Paul and Whitman (2008), *de* is identified as an aspect marker that leads to the mandatory past-tense interpretation of the sentence (Paul and Whitman 2008; Cui and Sung 2021). In fact, one of the most striking language-specific features of Chinese clefts is that, unlike English *it*-clefts, they have a default past tense reading. As a matter of fact, material contradicting the past-tense interpretation (e.g., future-oriented temporal adverbials) of the sentence cannot occur in Chinese proper clefts (Cui and Sung 2021; Hole 2011; Li and Thompson 1981; Paul and Whitman 2008).

From a discourse-pragmatic point of view, the *shì...de* proper cleft is generally considered a focalizing device that shows a bipartioning between the focus (i.e., the clefted constituent) and the presupposed content (Jing-Schmidt 2017). Similarly to English *it*-clefts, the Chinese cleft includes a narrow focus signaled by the copula. The syntactic constituents that can be clefted, and thus occupy the post-copular position, are subjects and adjuncts (Paul and Whitman 2008), whereas, due to Chinese word order constraints, post-verbal elements such as objects and verbal complements cannot[16] (see Luo 2009).

Following Li (2008), Xu (2014), and Cui and Sung (2021), we argue

---

14 Variously analyzed as a copula by Paris (1979), Ross (1983), Li and Thompson (1981), Cheng (2008), Paul and Whitman (2008), as a copula marking the focus by Li (2008), Xu (2014), as an intensifier adverb functioning as an emphasis marker by Shi (1994).

15 Variously categorized as a nominalizer of a headless relative clause by Paris (1979), Li and Thompson (1981), Cheng (2008), Li (2008) and Xu (2014), as an aspect marker by Zhao (1979), Lü (1982), Shi (1994) and Cui and Sung (2021), as a head of an aspectual phrase (AspP) projection by Paul and Whitman (2008), as an enclitic past tense marker by Simpson and Wu (2002).

16 It must be pointed out that post-verbal constituents can receive focus by means of phonological prominence (Lü 1982; Cheng 2008; Cui and Sung 2021), as in the case of "object focus clefts" (Paul and Whitman 2008: 424; Hole 2011: 1712), where the object is not the "cleft focus", but the prosodically "marked focus" (Hole 2011: 1712).

that the *shì…de* proper cleft has both contrastive and non-contrastive discourse-pragmatic functions. We can distinguish two types of *shì…de* proper clefts: contrastive clefts and non-contrastive clefts. In contrastive *shì…de* proper clefts, the focal element conveys contrast because the information it expresses is opposed to other relevant information in the discourse. Such sentences are used to correct, expand, or clarify the listener's assumptions (see Berretta 1994). Therefore, sentences such as (2) contain a corrective contrastive focus (Jing-Schmidt 2017; Cui and Sung 2021).

(2) 他不是开车来的，是坐火车来的。(Zhao 1979: 62)

tā	*bú*	*shì*	*kāi-chē*	*lái-de*
3SG	NEG	SHI	drive-car	come-DE
shì		*zuò-huǒchē*		*lái-de*
SHI		by-train		come-DE

'It was not by car that he came, but by train.'

Conversely, in non-contrastive clefts, although the clefted constituent is syntactically focalized, it does not convey contrast because the information it expresses is not opposed to other information already given in the discourse (Berretta 1994; Xu 2014; Garassino 2014). For example, *wh*-interrogative cleft sentences such as (3A) are mainly used non-contrastively (Li 2008; Cui and Sung 2021). Here, the non-contrastive focus has the full "original" focus marking function (Korzen 2014: 232), as it is used to direct the listener's attention to a specific piece of information of a concluded event, without any intention to create contrast (Berretta 1994; Cui and Sung 2021). This also applies to corresponding responses such as (3B).

(3) A: 你是怎么去的中国？(Zhao 1979: 61)

nǐ	*shì*	*zěnme*	*qù-de*	*Zhōngguó*
2SG	SHI	how	go-DE	China

'How did you go to China?'

B: 我是坐飞机去的。

wǒ	*shì*	*zuò-fēijī*	*qù-de*
1SG	SHI	by-plane	go-DE

'I went by plane.'

Furthermore, non-contrastive clefts have a textual function, serving anaphoric recovery: the clefted constituent recalls or summarizes what has been said before (Berretta 1994). In other words, the cleft sentence brings an element from the "background" to the "foreground" of the text (Prince 1978: 891). For example, in sentence (4), the *shì...de* cleft served to draw the listener's attention to a detail that had remained in the background, i.e., *lái táonàn* 来逃难 ('come to be a refugee'), which had already appeared as a complete predicate. The example in (4) and the corresponding translation are taken from Xu (2014: 174).

(4) 我来逃难我都什么也不管。头发也不管，[. . .] 衣服没有买过一件[. . .]，先天上就觉得说我是来逃难的，我就应该很吃苦耐劳。

wǒ	*lái*	*táonàn*		*wǒ*	*dōu*	*shénme*
1SG	come	take.refugee		1SG	all	whatever
yě	*bù*	*guǎn*	*tóufǎ*	*yě*	*bù*	
also	NEG	regard	hair	also	NEG	
guǎn		*yīfú*	*méiyǒu*	*mǎi-guo*		*yī*
regard		clothes	NEG	buy-EXP		one
jiàn	*xiāntiān-shàng*			*jiù*	*juédé*	
CLF	in.nature-upon			just	feel	
shuō	*wǒ*	*shì*	*lái*	*táonàn-de*		
say	1SG	SHI	come	take.refugee-DE		
wǒ	*jiù*	*yīnggāi*	*hěn*	*chī-kǔ-nàiláo*		
1SG	just	should	very	be.able.to.bear.hardships		

'I came to be a refugee, so I have regards for nothing. I don't care about my hair, [...] and I have not bought one piece of clothing [...]. Internally I feel that I am here to be a refugee, and I should be able to bear all hardships.'

Based on the assumption that the *shì...de* proper cleft has both corrective contrastive and non-contrastive functions, BILCC, inspired by the working models adopted by Callies (2009), identifies two different pragmatic functions for annotating the corpus data at the pragmatic level: 'intensification' and 'corrective contrast'. Intensification refers to *shì...de* cleft sentences that do not convey a contrastive focus, such as the non-contrastive clefts in (3)-(4), where the focus has the original

function of highlighting a piece of information. Corrective contrast refers to sentences in which the *shì...de* pattern signals a corrective contrastive focus, such as the contrastive clefts in (2).
To summarize, the data collected for the compilation of BILCC have been annotated at the grammatical and pragmatic levels to explore L1 Italian learners' knowledge of the syntactic and pragmatic properties of the *shì...de* cleft construction[17]. The corpus annotation process started in 2021 and is still ongoing. BILCC will be made freely available once the compilation and annotation are completed.

4. A case study: The written sub-corpus of BILCC

In this section, I will address the methodological issues relating to the design and collection of the written sub-corpus of BILCC. It was assembled according to strict specific design criteria and is the result of "theoretically motivated" (Tracy-Ventura and Paquot 2021: 4) open ended tasks. In what follows, I will describe: a) general corpus design principles and SLA-motivated features; b) the corpus typology; c) environment, learner, and task variables; d) the data collection procedure.

4.1 Corpus design features

Based on the design criteria recommended by Tracy-Ventura *et al.* (2021) to fill current gaps in corpus compilation in LCR, the written sub-corpus of BILCC shows the following features:

1. *It focuses on L2s other than English*. Since most learner corpora collect data on L2 English, and in LCR there is a general shortage of corpora collecting data from other L2s (Gráf 2017; Tracy-Ventura *et al.* 2021), this corpus starts to fill the gap in LCR by collecting data on L2 Chinese, an underexplored language variety in LCR (Iurato 2022a; Iurato 2022b).

[17] The analysis of the *shì...de* cleft construction on syntactic, semantic and discourse-pragmatic levels goes beyond the scope of this paper. For a literature overview of the topic, see Cheng (2008), Hole (2011), Iurato (in preparation), Jing-Schmidt (2017), Li (2008), Lü (1982), Paul and Whitman (2008) and Xu (2014).

2. *It includes data from learners at all proficiency levels*. Unlike most corpora that collect data from intermediate and advanced learners (Tracy-Ventura *et al.* 2021), this corpus includes data from beginner, intermediate, and advanced learners. The inclusion of beginner learner data is important, as the purpose of SLA research is to explain acquisitional development from beginning to end (Tracy-Ventura *et al.* 2021).
3. *It includes a control corpus of L1 speakers for comparative purposes*. Based on the Integrated Contrastive Model (Granger 1996) for the study of cross-linguistic influence, the learner corpus is accompanied by a control corpus of 30 Chinese native speakers as a benchmark of the (variety of) language learners are exposed to (Lozano 2021)[18]. Moreover, following one of the most important corpus design criteria outlined by Sinclair (2005), the two corpora are comparable because the tasks that were administered to the learners and the control group were identical. In other words, the same design across the two corpora ensures comparability, a key issue particularly emphasized by Lozano (2021).
4. *It contains a rich set of metadata on learner variables*, as it is important to document learner variables accurately both to support data interpretation (Bell and Payant 2021) and to increase reliable comparability across studies (Tracy-Ventura *et al.* 2021). Adhering to another corpus design principle defined by Sinclair (2005) on the documentation of variables, the present corpus contains systematically collected metadata of learners' and Chinese native speakers' sociolinguistic variables.
5. *It includes a pilot study of the data collection*. Since piloting is not a common practice in learner corpus design and should rather become part of it (Tracy-Ventura *et al.* 2021; Bell and Payant 2021), this corpus went through piloting to a) check the effectiveness of the tasks; b) ensure that the instructions and tasks were understandable for the participants; c) measure the time that the participants needed to complete the tasks; d) check whether the expected

[18] Control corpora are justified in LCR; see the 'comparative fallacy' vs. 'comparative hypocrisy' debate (Meunier 2021; Tracy-Ventura *et al.* 2021).

findings could emerge from the collected data; e) correct/eliminate mistakes left in the instruments (Dörney and Csizér 2012).

6. *It is freely available to the research community*. Making corpora freely available is a highly encouraged practice in LCR. This is especially the case for (error)-annotated corpora, which are unfortunately rarely shared, but definitely necessary for the development of NLP tools (Tracy-Ventura *et al.* 2021). To fill this gap in the LCR literature, the present corpus will be made freely available to the research community once completed. Furthermore, it will be accompanied by the documentation on metadata and effects on piloting, as this practice will help researchers understand the necessary decisions to be made when compiling a corpus (Bell and Payant 2021).

4.2 Corpus typology

Following the learner corpus typology dimensions outlined by Bell and Payant (2021), Gilquin (2015), Granger (2012), and Meunier (2021), the written sub-corpus of BILCC is:

1. *In-house*. The contributors and the users are not the same students, but they belong to the same population of learners, i.e., L1 Italian learners of L2 Chinese studying in Italy.
2. *Pseudo-longitudinal*. It collects data at a specific point in time (December 2020-March 2021) from different learners at different stages in their development (beginner, intermediate, and advanced learners).
3. *Mono-L1*. It contains data produced by a single L1 population.
4. *Academic*. It is compiled for research purposes.
5. *Specific purpose designed*. Due to the lack of data from Italian learners in existing L2 Chinese learner corpora, the corpus was specifically designed to analyze the use of the *shì...de* cleft construction in L1 Italian learners' production.
6. *Representative*. The language samples in the corpus are representative of learners' contextualized language use at three different proficiency levels, as the data are produced through open-ended tasks that allow learners to choose their own wording (Callies and Götz 2015). The feature of representativeness distinguishes the corpus from common data collections (Meunier 2021).

7. *Error and pragmatically tagged*. A target-oriented error taxonomy and an error tagset with 20 labels for the annotation at the grammatical level were designed to spot learners' errors in the use of the *shì...de* cleft construction. A pragmatic annotation was also added to detect the misuse of the *shì…de* proper cleft construction at the discourse-pragmatic level. Following Díez-Bedmar (2015), the identification of errors was carried out simultaneously by a bilingual team consisting of two expert Chinese native speakers and the researcher, whose L1 is the same as that of the learners.

4.3 Environment, learner and task variables

The written section of BILCC was compiled in a foreign language context in an "educational setting" (Gilquin 2015: 16), i.e., Chinese language courses at the Ca' Foscari University of Venice.
The metadata of the written corpus collected all information about learner variables. In order to increase the rigor and transparency of the corpus (Bell and Payant 2021), the metadata will be published alongside the learner corpus. Based on the learner metadata scheme proposed by Wang *et al.* (2015), a detailed metadata set that gathers information on the learner profile[19] was collected. Information on learners' educational background[20], knowledge of further foreign language(s) other than Chinese (plus related proficiency level), was also included, as this is crucial for interpreting the role of L1 in learners' interlanguage (Tracy-Ventura *et al.* 2021). These variables were collected through a learner profile questionnaire and a language background questionnaire. Participants completed these two questionnaires before completing the tasks. Students were also asked to complete an informed consent form[21].

---

[19] Age, gender, nationality, L1(s), parents' L1(s), partner's L1(s), current program of study.

[20] Highest level of education, languages officially used at primary, high school and university, countries where the learner attended school and university, periods of Chinese language study in China, purpose of stay in China, experience of living in Chinese communities.

[21] The template of the informed consent form was provided by the Ca' Foscari University of Venice. The specific research objective, i.e., the knowledge of the *shì...de* cleft construction, was not made explicit in order not to influence the participants' output.

As for the design of open-ended tasks, it is important to bear in mind that different types of tasks affect the learner output, as learners' language use varies across tasks. The variety of lexical-grammatical aspect combinations used by learners may also be influenced by other variables such as task topic, prompts, and type of narrative. Moreover, different tasks tap different knowledge (explicit or implicit) (Tracy-Ventura and Myles 2015). Following Callies (2009) and Tracy-Ventura and Myles (2015), one way to get around the above-mentioned issues in creating the written tasks for the compilation of BILCC was to carefully design tasks that naturally create contexts for the features under investigation. As the corpus collects data to explore whether learners are aware of the two pragmatic functions of the *shì…de* cleft construction, specific open-ended tasks were created to elicit the data in a definite discourse context in which the use of the perfective verbal aspect could also emerge. Thus, the data production was contextualized in a scenario in which it was necessary to refer to concluded events. First, the tasks provided a background that allowed students to highlight particular details of a concluded action (approach used to explore learners' knowledge of the pragmatic meaning of intensification), and then a background that allowed them to clarify/correct incorrect information/assumptions related to a concluded event (approach used to explore learners' knowledge of the pragmatic function of the corrective contrastive focus). Four purpose-designed, theoretically motivated, open-ended written tasks were designed: two discourse completion tasks (DCTs) and two picture-based narratives. The tasks provided contexts for the time reference and the aspect as authentically as possible, by implementing the principles for ensuring the task effectiveness defined by Tracy-Ventura and Myles (2015). The tasks were also rich in background and foreground.

Based on Callies (2009), the DCTs consisted of two sections introduced by different situational descriptions designed to create communicative contexts in which specific information needed to be highlighted for reasons of intensification (situation 1) and corrective contrast (situation 2). Each item was contextualized by a short passage extracted from the narrative text that the participants had read earlier. This text passage was followed by a semi-structured dialogue sequence that participants were asked to complete.

In section 1, participants were asked to provide utterances focusing on a detail of information highlighted in the preceding text passage, as exemplified in (5):

(5) 中午，我们在餐厅吃了午饭。
zhōngwǔ wǒmen zài cāntīng chī le wǔfàn
(At noon, we had lunch at the restaurant)
A (your friend's question): ...
B (your response): ...

In section 2, the items involved obvious cases of misunderstanding between two interlocutors: the informant (interlocutor A) and a fictitious fellow student (interlocutor B). Participants were asked to create a dialogue in which interlocutor A corrects or clarifies incorrect assumptions made by interlocutor B concerning a situation that occurred in the context described in the previous text, as illustrated in (6):

(6) 午饭过后，我们在电影院旁边的超市买了一些饮料。
wǔfàn guòhòu wǒmen zài diànyǐngyuàn pángbiān de chāoshì mǎi le yì xiē yǐnliào.
(After lunch, we bought some drinks from the supermarket next to the cinema)
A (your friend's wrong assumption): ...
B (your corrective response/clarification): ...

The DCTs contained ten items, five of which appeared in an intensifying, and five in a contrastive context. Participants were asked to highlight (situation 1) or correct/clarify (situation 2) information about specific details (time, place, manner, agent, etc.) of a concluded event described in the narrative text. No distractors were included in the DCTs. The instructions were highly detailed, and the communicative goal was explicitly stated, thus there was no need to conceal this by including distractors.
The first picture-based narrative also consisted of two sections. In section 1, participants were given a picture preceded by a context with instructions. Participants were asked to answer the question provided in order to highlight a detail of the event described in the picture, see Figure 1.

Tom has just bought a new book. He shows it to his friends Hannah and John. Hannah and John are curious to know where and when he bought that book. How could they ask Luca for such information? And how could Tom reply to them to provide this specific information?

Figure 1. Example of picture-based narrative, section 1.

In section 2, informants were asked to create a dialogue between the portrayed interlocutors, in which interlocutor A was required to correct/clarify incorrect assumptions made by interlocutor B about specific details of a concluded event, see Figure 2.

Layla received a bouquet of flowers. Her friend Nicole thinks it was Oliver who gave her those flowers, but Layla corrects this wrong assumption, clarifying that it was Ismael who actually gave them. Write the dialogue between Layla and Nicole.

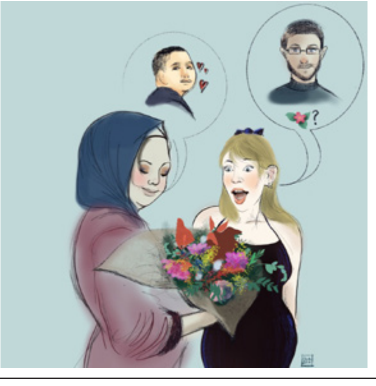


Figure 2. Example of picture-based narrative, section 2.

The second picture-based narrative, 'My trip to China', consisted of an open role-play (Mackey and Gass 2021). Students were given a picture (see Figure 3) representing two interlocutors accompanied by an attack line (*wǒ qù le Zhōngguó* 我去了中国, 'I went to China') from which to create two dialogues in two different contexts. In the first dialogue, interlocutor B was required to obtain specific information from interlocutor A about the concluded trip. In the second dialogue, interlocutor A was required to clarify/correct incorrect information provided by interlocutor B about the concluded event.

Figure 3. Illustration used to complete the second picture-based narrative, open role-play.

An artist[22] was commissioned to draw the pictures included in the tasks. Instructions were given in Italian for the learners and in Chinese for the control group. The use of teaching materials and dictionaries was not allowed. Since the research did not focus on vocabulary knowledge, an entry level vocabulary based on the official HSK2 vocabulary list was included in the tasks; those words that the students might not know were provided. Time on tasks was not controlled, as completion of the tasks was the priority. Before administration, all

[22] I would like to thank Francesca Biundo for drawing the illustrations used for data collection, some of which appear in this paper.

Chinese sentences in the tasks were checked by three native speakers who teach Chinese language at the Ca' Foscari University of Venice. Moreover, the tasks were piloted to ensure that they were manageable for all participants, that the vocabulary was appropriate, and that the instructions were clear (Bell and Payant 2021). The tasks were administered from the least to the most structured (1. picture-based narratives; 2. DCTs) to avoid the completion of the most guided tasks affecting the production (Mackey and Gass 2021).

4.4 Data collection and corpus description

The learner questionnaires, the consent form, and the tasks were administered via Google Form[23]. Three out of the 103 learners who had completed the written tasks were ruled out because their L1 was Chinese. As for the control group, 30 Chinese native speakers (Chinese language teachers: N=24; students: N=6) who live in Italy, Germany, USA, and China completed the tasks. Their average age was 36. Data and metadata from learners and native speakers were downloaded and stored in electronic format on Excel files, so that they could be retrieved and used with different software. The data of each learner and each native speaker were labelled with a univocal code (e.g., L1, NS1, etc.), since personal information revealing names and surnames of students was eliminated to preserve their privacy (Castillo Rodríguez *et al.* 2020). Data cleaning and character counting were carried out using Regex in Nisus Writer Pro[24]. SegmentAnt (Anthony 2017) was used for basic word segmentation (Chinese Jieba Engine); AntConc (Anthony 2019) for word tokens and word types counting. Basic information on the learner corpus and the control corpus size are illustrated in Table 1.

	Learner corpus (100 Ls)	**Control corpus (30 NSs)**
Sentences	4,985	1,504
Chinese characters	53,248	14,321
Word tokens	38,793	10,414
Word types	693	285

Table 1. Size of the written sub-corpora of BILCC.

[23] The selection of data collection tools was dictated by the restrictions caused by the Covid-19 emergency during the lockdown period in Italy.

[24] https://www.nisus.com/pro/ (visited 2023/02/20).

The tasks proved to be effective as the expected results were achieved. Inferential analyses of the learner and the native speaker corpora reveal and confirm the initial hypothesis that the *shì…de* cleft construction is used much more frequently by native speakers than by learners (χ^2=143.307, df=1, *p*=.00001). Moreover, significant differences in the proportion of the *shì…de* cleft construction are observed between advanced and beginner (χ^2=104.637, df=1, *p*=.00001), between advanced and intermediate (χ^2=16.52625, df=1, *p*=.000047), and between elementary and intermediate learners (χ^2=62.8262, df=1, *p*=.00001). Therefore, statistics show that the *shì...de* proper cleft construction is used significantly more frequently by intermediate and advanced than by beginner learners (see Table 2).

	Beginner (16)	**Intermediate (50)**	**Advanced (34)**	**Total Ls**	**NSs (30)**
Frequency of correct *shì...de*	120	824	729	1,673	920
Total sentences in the corpus	815	2,462	1,708	4,985	1,504
Proportion of *shì...de* (%)	14.7	33.46	42.68	33.56	61.17

Table 2. Frequency rate of the *shì…de* cleft construction in learner and native speaker data.

Statistics also reveal that the *shì…de* cleft construction conveying the pragmatic meaning of intensification is more frequently used by native speakers than by learners (χ^2=213.1123, df =1, *p*=.00001), and that it is significantly more frequently used by advanced than by intermediate and beginner learners. Similarly, the *shì…de* proper cleft construction conveying the pragmatic meaning of corrective contrast is more frequently used by native speakers than by learners (χ^2=25.1626, df=1 *p*=.00001), and it is used significantly more frequently by advanced than by intermediate and beginner learners. However, it is interesting to note that the frequency rate of contrast in learners' written production is closer to that in native speakers' production, compared to the frequency rate of intensification in learner data which drastically deviates from the frequency rate of intensification in native speaker data (see Table 3). The analysis therefore reveals that the learners use the *shì...de* cleft construction more to correct/clarify wrong assumptions than to highlight the speaker's evaluation or explanation regarding a concluded event.

Intensification			Contrast	
	Native Speakers (30)	**Learners (100)**	**Native Speakers (30)**	**Learners (100)**
Frequency of correct *shì...de*	645	913	335	776
Total sentences in the corpus	1,504	4,985	1,504	4,985
Proportion of *shì...de* (%)	42.88	18.31	22.27	15.56
Significant difference between NSs and Ls: χ^2=213.1123, df =1, *p*<.05			Significant difference between NSs and LSs: χ^2=25.1626, df =1 *p*<.05	

Table 3. Frequency rate of the *shì...de* cleft construction conveying intensification and contrast in learner and native speakers data.

Since the focus of this article is outlining the methodological steps in corpus design and data collection, no further information on the accuracy rate and error rate of the *shì...de* cleft construction in learner and native speaker data will be provided. This information, as well as details on corpus annotation, corpus exploitation, data analysis, and data interpretation will be presented in future research.

5. Conclusions

This paper defined the methodological steps to be followed when compiling a written learner corpus specifically designed to analyse the morpho-syntactic and pragmatic knowledge of Chinese syntactic features in L1 Italian learners' output. The compilation of the written sub-corpus of BILCC was presented as a case study. The paper demonstrated the importance of methodological decisions in defining the corpus typology, learner and environment variables, and task types, as these aspects impact on the learner output, the validity, and the generalizability of findings emanating from the corpus (Bell *et al.* 2021). The rationale of the corpus compilation and the importance of methodological transparency in documenting all steps in the preparation of the learner corpus were also discussed. The purpose of this paper was presenting a protocolized methodology for the compilation of a learner corpus, which may also be suitable for future corpus projects, given the continuous expansion of L2 Chinese studies in Italy and the current lack of Italian corpora of L2 Chinese. If the methodology is to be replicated for future studies, the methodological steps must be well-structured and as flexible as possible. Only in this way can they be used for different research purposes (Bell *et al.* 2021). The compilation process presented here can be repurposed for future research on L2 Chinese acquisition, since it is grounded on the main methodological stages of LCR (Granger 2012), and it is based on specific design principles (Tracy-Ventura *et al.* 2021). In addition, the corpus size and statistical results demonstrated the effectiveness of the compilation process of BILCC. The standardization of the methodology for compiling a corpus, especially in an expanding field in which few studies have been carried out (as in the case of L2 Chinese corpora collecting data from Italian learners) (Iurato 2022a; Iurato 2022b), is essential to avoid different studies on the same topic generating different, sometimes even contradictory, results. Thus, since so far in Chinese LCR "there is no synthesis of research findings, making it difficult to outline a full picture of [L2 Chinese] learners' development" (Zhang and Tao: 58), the aim of this contribution is encouraging the replicability of the methodology presented here to support future studies on the acquisition of L2 Chinese by L1 Italian learners.

References

Anthony, L. (2017) SegmentAnt (Version 1.1.3) [Computer Software]. Tokyo: Waseda University. https://www.laurenceanthony.net/software (visited 2022/04/09).

Anthony, L. (2019) AntConc (Version 3.5.8) [Computer Software]. Tokyo: Waseda University. https://www.laurenceanthony.net/software (visited 2022/04/09).

Bell, P., L. Collins and E. Marsden (2021) "Building an Oral and Written Learner Corpus of a School Programme: Methodological Issues", in B. Le Bruyn and M. Paquot (Eds.) *Learner Corpus Research Meets Second Language Acquisition*. Cambridge, Cambridge University Press: 214-242.

Bell, P. and C. Payant (2021) "Designing Learner Corpora. Collection, Transcription, and Annotation", in N. Tracy-Ventura and M. Paquot (Eds.) *The Routledge Handbook of Second Language Acquisition and Corpora*. London/New York, Routledge: 53-67.

Berretta, M. (1994) "Ordini marcati dei costituenti di frase in italiano. La frase scissa". *Vox Romanica*, 53: 79-105.

Callies, M. (2009) *Information Highlighting in Advanced Learner English. The syntax–pragmatics interface in second language acquisition*. Amsterdam, John Benjamins.

Callies, M. and S. Götz (2015) "Learner Corpora in Language Testing and Assessment. Prospect and Challenges", in M. Callies and S. Götz (Eds.) *Learner Corpora in Language Testing and Assessment*. Amsterdam, John Benjamins: 1-9.

Campillo Llanos, L. (2014) "A Spanish Learner Oral Corpus for Computer-Aided Error Analysis". *Corpora*, 9(2): 207-238.

Castillo Rodríguez, C., J.M. Díaz Lage and B. Rubio Martínez (2020) "Compiling and Analyzing a Tagged Learner Corpus: A Corpus-Based Study of Adjective Uses". *Círculo de Lingüística Aplicada a la Comunicación*, 81: 115-136. http://dx.doi.org/10.5209/CLAC.67932.

Cheng, L. (2008) "Deconstructing the *shi...de* Construction". *The Linguistic Review*, 25: 235-266.

Cui, S. and K. Sung (2021) *A Reference Grammar for Teaching Chinese. Syntax and Discourse*. Singapore, Springer.

Díez-Bedmar, M.B. (2015) "Dealing with Errors in Learner Corpora to Describe, Teach and Assess EFL Writing: Focus on Article Use", in E. Castello, K. Ackerley and F. Coccetta (Eds.) *Studies in Learner Corpus Linguistics. Research and Applications for Foreign Language Teaching and Assessment*. Bern, Peter Lang: 37-69.

Dörney, Z. and K. Csizér (2012) "How to Design and Analyze Surveys in Second Language Acquisition Research", in A. Mackey and S.M. Gass (Eds.) *Research Methods in Second Language Acquisition. A Practical Guide*. Oxford, Wiley-Blackwell: 74-94.

Dutra, D.P. and A.R. Gomide (2015) “Compilation of a University Learner Corpus”. *BELT, Brazilian English Language Teaching Journal*, 6: 21-33. http://dx.doi.org/10.15448/2178-3640.2015.s.21311.

Faitaki, F. and V.A. Murphy (2020) “Oral Language Elicitation Tasks in Applied Linguistics Research”, in J. McKinley and H. Rose (Eds.) *The Routledge Handbook of Research Methods in Applied Linguistics*. London/New York, Routledge: 360-369.

Fu, J. 符华均, P. Zhang 张晋军, Y. Li 李亚男, P. Li 李佩泽 and T. Zhang 张铁英 (2013) “新汉语水平考试HSK（五级）效度研究”. *Kaoshi Yanjiu*, 3: 65-69.

Gao, F. and B. Wang (2017) “A Multimodal Corpus Approach to Dialogue Interpreting Studies in the Chinese Context: Towards a Multi-Layer Analytic Framework”. *The Interpreters’ Newsletter*, 22: 17-38.

Garassino, D. (2014) “Cleft Sentences. Italian-English in Contrast”, in A.M. De Cesare (Ed.) *Frequency, Forms and Functions of Cleft Constructions in Romance and Germanic: Contrastive, Corpus-Based Studies*. Berlin, De Gruyter: 101-138.

Gass, S.M. (2018) “SLA Elicitation Tasks”, in K. Phakiti, P. De Costa, L. Plonky and S. Starfield (Eds.) *The Palgrave Handbook of Applied Linguistics Research Methodology.* London, Palgrave Macmillan: 313-338.

Gilquin, G. (2015) “From Design to Collection of Learner Corpora”, in S. Granger, G. Gilquin and F. Meunier (Eds.) *The Cambridge Handbook of Learner Corpus Research.* Cambridge, Cambridge University Press: 9-34.

--- (2021) “Combining Learner Corpora and Experimental Methods”, in N. Tracy-Ventura and M. Paquot (Eds.) *The Routledge Handbook of Second Language Acquisition and Corpora*. London, Routledge: 133-144.

Gilquin, G. and S.T. Gries, (2009) “Corpora and Experimental Methods: A State-of-the-Art Review”. *Corpus Linguistics and Linguistic Theory*, 5(1): 1-26. https://doi.org/10.1515/CLLT.2009.001.

Gráf, T. (2017) “The Story of the Learner Corpus LINDSEI_CZ”. *Studie z Aplikovane Lingvistiky*, 8(2): 22-35.

Granger, S. (1996) “From CA to CIA and Back: An Integrated Approach to Computerized Bilingual and Learner Corpora”, in K.

Ajimer, B. Altenberg and M. Johansson (Eds.) *Languages in Contrast. Text-Based Cross-Linguistic Studies*. Lund, Lund University Press: 37-51.
--- (2012) "How to Use Foreign and Second Language Learner Corpora", in A. Mackey and S. M. Gass (Eds.) *Research Methods in Second Language Acquisition. A Practical Guide.* Oxford, Wiley-Blackwell: 7-29.
Granger, S., G. Gilquin and F. Meunier (Eds.) (2015) *The Cambridge Handbook of Learner Corpus Research*. Cambridge, Cambridge University Press.
Hinrichs, E. and S. Krauwer (2014) "The CLARIN Research Infrastructure: Resources and Tools for E-Humanities Scholars", in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, May 2014: 1525-1531.
Hole, D. (2011) "The deconstruction of Chinese shi...de clefts revisited". *Lingua*, 121: 1707-1733.
Huang, L. (2018) "Issues on Multimodal Corpus of Chinese Speech Acts: A case in Multimodal Pragmatics". *Digital Scholarship in the Humanities*, 33(2): 316-326.
Iurato, A. (2022a) "Learner Corpus Research Meets Chinese as a Second Language Acquisition: Achievements and Challenges". *Annali di Ca' Foscari. Serie Orientale*, 58(1): 709-742.
--- (2022b) "Analyzing Chinese Learner Corpus Research and Chinese learner corpora: Main Features, Critical Issues and Future Pathways". *Kervan. International Journal of African and Asian Studies*, 26(2): 533-564.
--- (in preparation) *The Acquisition of the Chinese 是shì...的de Cleft Construction by L1 Italian Learners*: *Triangulating Learner Corpus and Experimental Data*. PhD Dissertation. Venice, Ca' Foscari University of Venice; Bremen, University of Bremen.
Jing-Schmidt, Z. (2017) "Grammatical Construction and Chinese Discourse", in C. Shei (Ed.) *Routledge Handbook of Chinese Discourse Analysis*. London, Routledge: 209-230.
Jong, F. de, B. Maegaard, D. Fišer, U. Dieter Van and A. Witt (2020) "Interoperability in an Infrastructure Enabling Multidisciplinary Research: The case of CLARIN", in *Proceedings LREC 2020, 12th International Conference on Language Resources and Evaluation*. ELRA. https://www.aclweb.org/anthology/2020.lrec-1.417/ (visited 2023/02/20)

Kong, K.C. (2013) "A Corpus-Based Study in Comparing the Multimodality of Chinese- and English- Language Newspapers". *Visual Communication*, 12(2): 173-196. https://doi.org/10.1177/1470357212471594.

Korzen, I. (2014) "Cleft sentences. Italian-Danish in contrast", in A.M. De Cesare (Ed.) *Frequency, Forms and Functions of Cleft Constructions in Romance and Germanic: Contrastive, Corpus-Based Studies*. Berlin, De Gruyter: 217-275.

Leclercq, P., A. Edmonds and H. Hilton (2014) (Eds.) *Measuring L2 Proficiency: Perspectives from SLA*. Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781783092291.

Li, C.N. and S.A. Thompson (1981) *Mandarin Chinese: a functional reference grammar*. Berkeley, University of California Press.

Li, K. (2008) "Contrastive Focus Construction in Mandarin Chinese", in M.K.M. Chan and H. Kang (Eds.) *Proceedings of the 20th North American Conference on Chinese Linguistics* (NACCL-20), vol. 2. Columbus (Ohio), The Ohio State University: 759-774.

Lozano, C. (2021) "CEDEL 2: Design, Compilation, and Web Interface of an Online Corpus of L2 Spanish Acquisition Research". *Second Language Research*, 1-19. https://doi.org/10.1177/02676583211050522.

Lozano, C. and A. Mendikoetxea (2013) "Corpus and Experimental Data: Subjects in Second Language Research", in S. Granger, G. Gilquin and F. Meunier (Eds.) *Twenty Years of Learner Corpus Research: Looking Back, Moving Ahead*. Louvain-la-Neuve, Presses Universitaires de Louvain: 313-323.

Lü, B. 吕必松 (1982) "关于 "是……的"结构的介个问题". *Yuyan Jiaoxue yu Yanjiu*, 4: 21-37.

Mackey, S. and S.M. Gass (2021) *Second Language Research. Methodology and Design*. New York/London: Routledge.

Mai, Z. and B. Yuan (2016) "Uneven Reassembly of Tense, Telicity and Discourse Features in L2 Acquisition of the Chinese shì…de Cleft Construction by Adult English Speakers". *Second Language Research*, 32 (2): 247-276.

Meunier, F. (2021) "Introduction to Learner Corpus Research", in N. Tracy-Ventura and M. Paquot (Eds.) *The Routledge Handbook of Second Language Acquisition and Corpora*. London, Routledge: 23-36.

Millar, N. and B. Lehtinen (2008) "DIY Local Learner Corpora: Bridging Gaps between Theory and Practice". *JALT CALL Journal*, 4(2): 61-72.

Norris, J. and L. Ortega (2003) "Defining and measuring SLA", in C.J. Doughty and M.H. Long (Eds.) *The Handbook of Second Language Acquisition*. MA-US, Blackwell Publishing Ltd: 717-761.

Paris, M.-C. (1979) *Nominalization in Mandarin Chinese. The Morpheme 'de' and the 'shi…de' Constructions*. Paris, Université de Paris VII.

Paul, W. and J. Whitman (2008) "Shi...de Focus Clefts in Mandarin Chinese". *The Linguistic Review*, 25(3/4): 413-451.

Peng, Y., W. Yan and L. Cheng (2021) "Hanyu Shuiping Kaoshi (HSK): A Multi-Level, Multi-Purpose Proficiency Test". *Language Testing*, 38(2): 326-337. https://doi.org/10.1177/0265532220957298.

Prince, E.F. (1978) "A comparison of wh-clefts and it-clefts in discourse". *Language*, 54: 883-906.

Prior, M.T. (2018) "Interviews and Focus Froups", in K. Phakiti, P. De Costa, L. Plonky and S. Starfield (Eds.) *The Palgrave Handbook of Applied Linguistics Research Methodology*. London, Palgrave Macmillan: 225-248.

Rolland, L., J. Dewaele and B. Costa (2020) "Planning and Conducting Ethical Interviews: Power, Language and Emotions", in J. McKinley and H. Rose (Eds.) *The Routledge Handbook of Research Methods in Applied Linguistics*. London/New York, Routledge: 279-289.

Romagnoli, C. and S. Conti (Eds.) (2021) *La lingua cinese in Italia. Studi su didattica e acquisizione*. Roma, Roma TrE Press.

Ross, C. (1983) "On the Functions of Mandarin 'de'". *Journal of Chinese Linguistics*, 11(2): 214-246.

Rubio, F., E. Kia, E. Schnur and J. Hacking (2021) *Multilingual Corpus of Second Language Speech* (MuSSeL). https://l2trec.utah.edu/learner-corpora/mussel/ (visited 2023/02/20).

Shan, C. 杉村博文 (1999) ""的"字结构与分类", in L. Jiang 江蓝生 and J. Hou 侯精一 (Eds.) 汉语现状与历史的研究. Beijing, China Social Sciences Press.

Shi, D. (1994) "The Nature of Chinese Emphatic Sentences". *Journal of East Asian Linguistics*, 3: 81-100.

Simpson, A. and Z. Wu (2002) “From D to T- Determiner Incorporation and the Creation of Tense”. *Journal of East Asian Linguistics*, 11: 169-209.

Sinclair, J. (2005) “Corpus and Text – Basic Principles”, in M. Wynne (Ed.) *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford, Oxbow Book Company: 1-16.

Tono, Y. (2003) “Learner Corpora: Design, Development and Applications”, in D. Archer, P. Rayson, A. Wilson and T. McEnery (Eds.) *Proceedings of the Corpus Linguistics 2003 Conference*. UCREL 16, Lancaster University: 800-809.

--- (2016) “What is Missing in Learner Corpus Design?”, in M. Alonso-Ramos (Ed.) *Spanish Learner Corpus Research: Current Trends and Future Perspectives*. Amsterdam, John Benjamins: 33-52.

Tracy-Ventura, N. and F. Myles (2015) “The Importance of Task Variability in the Design of Learner Corpora for SLA Research”. *International Journal of Learner Corpus Research*, 1(1): 58-95.

Tracy-Ventura, N. and M. Paquot (Eds.) (2021a) *The Routledge Handbook of Second Language Acquisition and Corpora*. London/New York, Routledge.

--- (2021b) “Second Language Acquisition and Corpora. An Overview”, in N. Tracy-Ventura and M. Paquot (Eds.) *The Routledge Handbook of Second Language Acquisition and Corpora*. London/New York, Routledge: 1-8.

Tracy-Ventura, N., M. Paquot and F. Myles (2021) “The Future of Corpora in SLA”, in N. Tracy-Ventura and M. Paquot (Eds.) *The Routledge Handbook of Second Language Acquisition and Corpora*. London/New York, Routledge: 409-424

Xu, J. (2019) “The Corpus Approach to the Teaching and Learning of Chinese as an L1 and an L2 in Retrospect”, in X. Lu and B. Chen (Eds.) *Computational and Corpus Approaches to Chinese Language Learning*. Singapore, Springer: 33-53.

Xu, Y. (2014) “A Corpus-Based Functional Study of *shi…de* Constructions”. *Chinese Language and Discourse*, 5(2): 146-184.

Wang, M., S. Malmasi and M. Huang (2015) “The Jinan Chinese Learner Corpus”. *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*. Denver (Colorado), Association for Computational Linguistics: 118-123.

Zhang, J. and H. Tao (2018) "Corpus-Based Research in Chinese as a Second Language", in C. Ke (Ed.) *The Routledge Handbook of Chinese Second Language Acquisition*. New York, Routledge: 48-62.

Zhao, S. 赵淑华 (1979) "关于"是……的"句". *Yuyan Jiaoxue yu Yanjiu*, 1: 57-66.

PRAGMATIC MARKERS AND THE RIGHT PERIPHERY IN MANDARIN CHINESE: A SYSTEMATIC REVIEW OF TYPES, FUNCTIONS AND CO-OCCURRENCE

Carmen Lepadat

Roma Tre University

1. Pragmatic markers: an introduction

In the past few decades, *pragmatic markers* have become one of the most prolific research topics in linguistics. Triggered by Schourup's (1982) and Schiffrin's (1987) seminal works on linguistic expressions occurring at the periphery of certain speech units, at the turn of the century the field had already become one in which "it is almost impossible to find one's way through the jungle of publications" (Fischer 2006: 1).

A plethora of labels, definitions and classifications flourished likewise (Dér 2010). Among these, the most frequently employed and quoted in the literature is Schiffrin's (1987: 31) category of *discourse markers*, defined as "sequentially dependent elements which bracket units of talk". Fraser's (1990: 386) category of *pragmatic markers*, which he defines broadly as "signals of the speaker's communicative intentions", is generally considered as being superordinate with respect to the former

(Traugott 2015). Other terms employed with partially overlapping meanings are discourse and pragmatic particles (Brinton 2017), which nonetheless carry a stronger association with short and monosyllabic terms – thus non covering all the expressions compatible with the functional properties of the class – and with an emergent use in connection with retrospective expressions (Haselow 2012).

A number of different criteria have been adopted to limit and define pragmatic markers, most of which rely on functional rather than formal features. Crucially, the defining function of pragmatic markers also appears to shift according to the approach taken in their study. Among the most important ones, Schiffrin's (1987) initial approach is in terms of discourse coherence, with discourse markers operating on different planes of discourse – frequently at the same time – in order to contextualize the utterance they belong to by referring indexically either backward or forward. Fraser (1990, 1996) adopts a pragmatic perspective putting the accent on expressions that are devoid of propositional content, i.e., being procedural in nature[1]. According to Fraser, (the absence of) truth-conditionality serves as the defining factor of a very heterogeneous group of expressions which may either signal the illocutionary force of an utterance, comment on the message therein conveyed, convey an entire message parallel to the former, or specify its relation to the foregoing discourse.

A slightly different approach is adopted by Blakemore (1987, 2002), who is concerned with expressions that are able to impose constraints on the "pragmatic inferences involved in the recovery of implicit content", i.e., with expressions that are *procedural* in nature rather than *conceptual* (Blakemore 2002: 4). From the perspective of diachronic change and semantic reanalysis, Traugott (2010) focuses on the processes leading linguistic expression to acquire new meanings over time, which may be characterised by either an increase of the *subjective* component conveying speaker attitude, feelings or viewpoint, or by

[1] The term adopted by Fraser is argued to follow Blakemore's (1987) initial sense of 'non-truth conditional' (Fraser 1996). However, Blakemore will argue in later works (e.g., Blakemore 2002) that procedural cannot be completely equated to non-truth conditional, since some pragmatic markers may play a role in affecting the truth value of an utterance.

a distinctive expression on the part of the speaking subject of his/her awareness and attention towards the addressee's social image, beliefs and feelings, i.e., *intersubjectivity*.

Latest accounts are provided in Traugott (2022) and Heine *et al.* (2021): the former, by referring to *discourse structuring markers* from the perspective of construction grammar (Goldberg 2006), puts the accent on "the fact that they are used not merely to reflect intended relationships but to signal and even shape such relationships" (Traugott 2022: 3); the latter, in discussing the rise mechanisms of *discourse markers*, identifies next to the process of grammaticalization that of co-optation – a "cognitive-communicative operation whereby a text segment […] is transferred from the domain or level of sentence grammar and deployed for use on the level of discourse organization" (Heine *et al.* 2021: 26) – in order to address both their grammatical functions and the role they play in processing linguistic discourse.

In the following sections, a comprehensive discussion of the scholarly research that has been produced on pragmatic markers from the pragmatic and functionalist perspectives will be carried out, while also duly pointing to studies taking different approaches when deemed necessary.

2. Formal features of pragmatic markers

As the above paragraphs clearly show, there is an evident divergence of opinions concerning the functional definition of pragmatic markers. This is also true as far as the formal features put forth by different scholars are concerned. According to Aijmer (2013), relevant formal features of pragmatic markers include positional, prosodic, lexical and stylistic aspects. Nonetheless, positional and prosodic aspects have been by far the most investigated.

Regarding the pragmatic markers' position, it has been often pointed out that they tend to occur in specific structural slots in the utterance or the conversational turn. In fact, both utterance and turn peripheries have been argued to represent ideal places where information can be provided on how to process upcoming or previous stretches of talk, clarify the latter's structural relation with the context, take over the floor or yield the turn to another interlocutor, etc. (Detges and Waltereit 2014). Interestingly, the position that pragmatic markers occupy may have an influence on the meaning they express (Bazzanella 2001). To give an example, parenthetical expressions such as the French *moi je trouve* 'I think' have been argued to perform an opposite function at the two peripheries, with the left periphery favouring a booster or strengthening interpretation and the right periphery calling for a mitigating role (Detges and Waltereit 2014). Taking this observation one step further, Detges and Waltereit (2014) conclude that left-peripheral elements are mainly concerned with coherence-oriented elements – including mainly referential and turn-taking items, whereas right-peripheral elements are more likely to be modality/stance-oriented – including mostly elements re-negotiating the illocutionary force of an utterance or expressing an intersubjective stance[2]. Similarly, Wang and Tao (2020) address the functions of the pragmatic marker *wǒ júede* 我觉得 'I think', arguing for the recent development of a turn-expansion function when usen in right-peripheral position, in addition to the already existing evaluative and epistemic functions (Wang and Tao 2020: 2). The treatise on the (a)symmetries between elements occurring at the left and right periphery is nonetheless still far from being comprehensive, since different elements appear to behave (slightly) differently when their position in the utterance is taken into account (Traugott 2014: 89).

The prosodic aspects of pragmatic markers have also been addressed to some extent, at least as far as English is concerned. Earlier works describe pragmatic markers as being characterized by "comma intonation" (Fraser 1996; Brinton 1996; Rouchota 1998; Kaltenböck *et al.* 2011), i.e. parenthetical intonation involving prosodic independence and a downstep which is normally found after a comma (Samek-Lodovici 2015: 139). However, experimental studies investigating the actual prosodic contour with which pragmatic markers are produced in naturally-occurring language are still few and show a great variability of results (Wichmann *et al.* 2010; De Cristofaro *et al.* 2022). In her

2 The argument is made in support to Beetching and Detges (2014)'s claim concerning the functional asymmetry existing between left and right periphery. For a less deterministic view on the disparities between the two positions, see Traugott (2014).

1996 volume, based on the review of scholarly works circulating in that period, Brinton argued that pragmatic markers are "short items, often phonologically reduced or unstressed" (Brinton 1996: 33). Although this statement might need to be toned down with more and more new studies appearing, there is nonetheless evidence that "pragmatic markers in present-day English are typically 'small' monosyllabic or disyllabic words", and that sometimes – especially in correspondence of a particularly frequent use of the expression – they might undergo phonological reduction (Brinton 2017: 4). This is what happens with *y'know*, *kinda*, and *sorta* in American English (Brinton 2017: 4). In other cases, however, instead of presenting phonological reduction and being part of a larger tone group, pragmatic markers were shown to be phonologically stressed and to be followed by a pause marking an independent tone group (Brinton 2017: 5).
From the above, it is clear that any list of features attempting to define pragmatic markers can only be applicable if conceived of as representing the (fuzzy) core of this category, with some members representing prototypical cases and others being located in more peripheral positions.

3. Defining the right periphery

The remaining part of this paper will be specifically concerned with right-peripheral pragmatic makers. Nonetheless, before moving to the description of these elements, a few words are in order as far as the exact definition of the right periphery is concerned. Scholarly works tend to be in general rather vague on this point, considering the small consensus around the units with respect to which pragmatic markers should be considered peripheral. Traugott (2015) has pointed out that the terms 'initial' and 'final' are usually employed in connection with units such as the clause, utterance or sentence, and, when conversational data is taken into account, also the turn (e.g., Wang and Tao 2020). To complicate things, however, is the fact that in addition to pragmatic markers, sentence peripheries can also host elements of marked information structure, including left and right dislocations, as well as instances of focus fronting. The two types of elements can be observed in (1a-c) and (2) respectively: the first two examples illustrate that the left periphery of the sentence may contain either a (topic) left dislocation which is coreferential with the sentence-internal pronoun *it* (1a), or a focus fronting that is not resumed by any overt coreferential form (1b), while in (1c) an instance of topic right dislocation is produced at the sentence right periphery; (2) shows the production of two pragmatic markers, *well* and *you know*, occurring in succession in utterance-initial position.

(1) a. *Your book* you should give it to Kim.
(Adapted from Traugott 2015: 119)
b. *YOUR BOOK* you should give *t* to Kim.
(Adapted from Traugott 2015: 119)
c. 来了吗, 你哥哥?

Lái-le	*ma,*	*nǐ*	*gēge?*
Come-PERF	SFP	2SG	older-brother

'Has (he) come, your older brother ?'
(Guo 1999: 1109)

(2) *Well you know*, I was really interested in biofeedback.
(Traugott 2015: 119)

While there is no doubt that information-structuring elements as those in (1a-c) are syntactically dependent upon the main clause (Lombardi Vallauri 2009; Frascarelli 2017; Badan 2015), the status of pragmatic markers is more controversial. Under the functionalist approach, the general view is endorsed that at least some pragmatic markers are clause-external – i.e. they are syntactically independent (Haselow 2012) – while (recent) studies in syntax tend to consider pragmatic markers as occupying a position that is part of the main clause (Munaro and Poletto 2002; Spean and Tenny 2003; Giorgi 2009; Badan 2020).
With this regard, Traugott (2015) argues that the distinction between the clausal and peripheral slots should be considered as blurred rather than discrete, for it has been shown that many pragmatic markers originate from clause-internal elements which eventually acquire parenthetical meanings trough intermediate stages of syntactic detachment, i.e. they undergo a grammaticalization process (Mulac and Thompson 1991; Traugott and Dasher 2002; Brinton 2017). This is for instance

the case of the English epistemic marker *I think*, which developed by means of the semantic reanalysis of a main verb taking a whole clause as its complement. Through syntactic detachment, the complex initially introducing an opinion by the speaker and followed by the complementizer *that*, has further acquired an epistemic meaning when combining with peripheral position and that-deletion (Mulac and Thompson 1991):

(3) a. *I think that* we're definitely moving towards being more technological.
b. *I think Ø* exercise is really beneficial, to anybody.
c. It's just your point of view you know what you like to do in your spare time *I think*. (Brinton 2017: 17)

Traugott (2015) explicitly recognizes that defining the unit with respect to which pragmatic markers are peripheral is complex. The solution she proposes to deal with this is to consider the existence of a gradient and permeable relationship between an inner periphery (type I elements in Traugott's terminology) surrounding the verb argument structure – including adjunct phrases tending to occur either at the beginning or at the end of the clause – and an outer periphery – including pragmatic markers which can be found either in left or right position (type II elements in Traugott's terminology). According to Traugott (2015: 127), the gradient account is supported by diachronic observations, since elements of the inner periphery have been often reinterpreted and made available as elements belonging to the outer periphery, as in the case of general extenders (e.g., *and stuff*).

For reasons of convenience, this paper will adopt the term utterance-final to refer to right-peripheral pragmatic markers rather than clause- or sentence-final, for its use appears closer to a context-based, pragmatic definition of the relevant syntactic unit, which is nonetheless inclusive of both spoken and written texts, unlike the term turn.

4. Utterance-final pragmatic markers

Different models have been proposed to address the discourse functions of the sentence peripheries (e.g., Onodera 2014, Beeching and Detges 2014, Degand 2014), some of which consider them as symmetric and some arguing for the existence of syntactic and functional differences between them (Traugott 2015). However, utterance-final pragmatic markers (henceforth UFPMs) have been relatively neglected up until recently, with most of the scholarly work focusing on those occurring at the left periphery (Traugott 2015: 119)[3]. An exception to this trend is represented by studies on Asian languages, the majority of which traditionally employ a more well-defined class of words at the end of the utterance to express functions such as illocutionary force, speaker attitude, epistemic modality and other (inter)subjective meanings (Simpson 2014: 157)[4]. This is also captured in the typological-oriented account that Hancil and others (2015) offer of what they term *final particles* – i.e. "elements that have little or no lexical or conceptual, but predominantly procedural meaning" (Hancil *et al.* 2015: 4). In Hancil and others, particles used in Asian languages represent one of the five categories of final particles that can be found across languages worldwide[5]. The remaining four types are classified according to their lexical source:

a) final particles of the conjunction type (e.g., English *but*, Finnish *mutta* 'but' and *ja* 'and', Japanese *kara* 'because') have interaction-oriented functions such as turn completion and/or turn-yielding and are employed to signal the kind of link that the utterance has with respect to an implicit proposition (Hancil *et al.* 2015: 10);
b) final particles of the conjunct/adverbial connector type (e.g., French *alors* 'thus, then', English *then*, Spanish *pues* 'so') are argued to have

[3] Recent works on (left and) right periphery do exist within the functionalist approach (e.g., Beeching and Detges 2014, Hancil *et al.* 2015, Van Olmen and Šinkuniene 2021).

[4] These will be the treated in detail in par. 5.1.

[5] However, as one anonymous reviewer of this paper suggests, it is important to highlight that in spite of the functional affinity between pragmatic markers and sentence-final particles (Degand *et al.* 2013), the two should be kept separate at least as far as their syntactic behaviour is concerned.

derived their textual or discourse-internal functions from former roles in marking structural relations within syntactic units (Hancil *et al.* 2015: 12);
c) final particles of the adverbial type (e.g., English *actually* and *anyway*, French *déja* 'already', German *jetzt* 'now') are mainly time, place or manner adverbs used to "refer to the temporal or segmental (or, metaphorically speaking, 'local') structure of ongoing discourse" (Hancil *et al.* 2015: 13);
d) final particles of the focus particle type (e.g., English *even*, Dutch *aleen* 'only', Cantonese *ye* 'only', 'merely') are used to signal that a speech contribution in ongoing discourse is either noteworthy or unplanned, i.e. a post-factum realisation or an afterthought (Hancil *et al.* 2015: 13).

Other taxonomies based chiefly on the English language include among the pragmatic markers occurring primarily in utterance-final position general extenders (e.g., *and stuff*, *or something*, *or whatever*), question tags (e.g., *isn't it?*) and what Haselow (2012: 183) calls *final particles* in a sense rather close to that of Hancil and others (2015) (e.g., *then*, *though*, *anyway*, *actually*, *even*). In addition, markers that can be found both at the left and at the right periphery include epistemic adverbs (e.g., *surely*, *no doubt*), comment clauses (e.g., *I think*, *you know*, *see*) and vocatives (e.g., *proper names*) (Traugott 2015; Haselow 2016). According to the account offered by Haselow (2016), elements occurring at the right periphery are loosely connected to the previous utterance form a syntactic point of view and are not part of its propositional content. On the other hand, they are crucial from a pragmatic point of view, for their functions relate to "speech planning, processibility, textual coherence, speaker-listener relationship, and contextual embeddedness" (Haselow 2016: 387). Furthermore, great importance is given to the fact that they occur precisely at the end of an utterance, that is to say at "the moment at which the tasks they serve becomes relevant in the real-time emergence of a structural unit" (Haselow 2016: 387). Unlike sentence-initial elements, cognitive tasks emerging in the temporal slot characterizing the end of an utterance are more likely to be related to needs such as closing up projections and adjusting several aspects of the utterance before its final reception on the part of the addressee, including illocutionary force, link or preciseness of equation with the preceding discourse, epistemic status and turn-yielding.

Despite the diverse pool of elements occurring at the right periphery, Haselow (2016) summarizes a number of core features that can be applied to all the six types of elements:

> (i) are used predominantly in spoken discourse;
> (ii) are not potentially turn-constitutive as they are backwards-oriented and require a 'host structure';
> (iii) make no contribution to the propositional content of an utterance (but can modify it, e.g. in terms of epistemic certainty);
> (iv) have procedural rather than conceptual meanings in the sense that they provide an interpretive cue;
> (v) have various functions on the metatextual and interpersonal level;
> (vi) are not integrated into the morphosyntactic dependency relations of the unit they follow;
> (vii) are morphologically invariant and tend to be conventionalized units;
> (viii) are functionally variant when produced at other points in time in utterance production or outside the specific construction. (Haselow 2016: 391-392)

5. Utterance-final pragmatic markers in Chinese

Among the linguistic devices that fit the above description in Mandarin Chinese, sentence-final particles (SFPs) have received the higher amount of attention in the literature. Traditionally described as expressing *yúqì* 语气 'modality' (Lü 1942; Zhu 1982; Wang 1985), more recently these "phonologically small elements, most frequently monosyllabic" (Simpson 2014: 157) have been investigated in connection to the wider group of pragmatic markers, with which they seem to share a number of properties, most and foremost the versatility in terms of functions performed (Lee-Wong 1998; Chu 2009; Shei 2014; Lepadat 2017). Much more recent appear to be the studies investigating the use of the remaining types of UFPMs in Mandarin, constituting Hancil and others' (2015) a) to d) groups. Following Miracle's (1991) seminal volume on discourse markers in Mandarin Chinese, in fact, an increasing number of works have been produced to investigate Mandarin's use of pragmatic markers similar to those of Indo-European languages (Fang 2000; Tao 2003; Feng 2008, 2010; Liu 2009). Those specifically or even indirectly focusing on markers occurring at the right periphery, however, are only a handful. The same can be said concerning the studies that acknowledge the existence of and tackle the relationship between SFPs and other UFPMs.

5.1 Sentence-final particles

The body of scholarly work produced in connection to SFPs is rather extensive and diversified in terms of perspectives adopted, and its exhaustive discussion is beyond the purpose of this paper[6]. Instead, the current discussion will focus on the set of functions or meanings that can be conveyed through their use, for a more fruitful comparison with the other types of UFMPs.

The most frequently studied and commonly accepted modal particles are *de* 的, *le* 了, *ma* 吗, *ba* 吧, *ne* 呢 and *a* 啊, but more recent studies also tackled less frequent and much more orality-oriented devices such as *ou* 呕, *ei* 欸, *la* 啦, etc. Described as "hallmarks of natural conversation" (Luke 1990: 11), the difficulty of grasping each particle's exact meaning(s) (Li and Thompson 1981) is probably connected to their taking on a whole array of nuances connected to the pragmatic environment in which they occur (Chu 2009). Overall, their functions have been argued to pertain to several distinct domains in the area of pragmatics, encompassing both discourse-structural, subjective and intersubjective uses.

Studies such as Zhu (1982) and more recently Paul and Pan (2017) argue that while (some uses of the) particles *ma*, *ne* and *ba* are primarily connected to the expression of illocutivity, SFPs such as *a*, *ou*, *ei*, etc., are instead specialized in expressing the speaker's attitude. In the former case, the use of the SFP is not optional, since it serves to codify a specific sentence type, as the yes/no question realized by *ma* in (4):

6 For a more extensive account of the SFP inventory and their functions in Mandarin, see e.g., Qi 2002.

(4) 你想吃点什么吗？
Nǐ xiǎng chī diǎn shénme ma ?
2SG want eat a.bit what SFP
'Do you want to eat a little something?' (Paul and Pan 2017: 5)

On the other hand, when expressing speaker attitude, SFPs are argued to be syntactically optional, but pragmatically indispensable in order to express the intended meaning. Speaker attitude or even modality are often used in a broad way to include both the degree of the speaker's epistemic commitment towards the truth of the proposition (e.g., Chu (1998) regarding *ba*, Li (2006) for *ba* and *ma*) and other evaluative notions such as impatience (Paul and Pan (2017) concerning *ou*), or the signalling of unusual or extraordinary content (Li 2006 for *ne*). According to Liu *et al.* (2001), when used in interrogatives, *ba* expresses the speaker's epistemic uncertainty towards the proposition, as shown in (5) below:

(5) 这座楼可能是你们的宿舍吧？
Zhè zuò lóu kěnéng shì nǐmen de sùshè ba?
This CL building maybe be 2PL ATTR dormitory SFP
'Perhaps this is your dormitory, isn't it?' (Liu *et al.* 2001: 424)

Quite differently, the use of *ou* 喔 in (6) is argued by Shei (2014: 264) to stress whatever the speaker's emotions might be, including the portrayal of a situation as undesirable or unfavourable.

(6) 安钧璨：你们好令我失望喔！
Ān Jūncàn: nǐmen hǎo lìng wǒ shīwàng o!
An Juncan: 2PL so make me disappoint SFP
'Male Guest: I am so disappointed with you all!' (Shei 2014: 264)

Other functions performed by SFPs belong to the area of intersubjectivity (Nuyts 2006; Tantucci 2013). Lee-Wong's (1998: 388) seminal study stresses the role of SFPs such as *ba*, *a* and *ne* as "mitigators in a context where face threat is implicit". In the same direction seem to be pointing Tantucci and Wang (2018: 64), who treat SFPs as intersubjective "operators of rapport maintenance, as they are employed to overtly account for H[earer]'s potential reactions to S[peaker]'s utterance". This can be observed in (7) below, where both *a* and *ba* are employed to maintain the faces of both speaker and hearer, the former signalling the casual tone of the question, and the latter further softening the forcefulness of the invitation (Lee-Wong 1998: 396):

(7) A male colleague asks a female colleague for a date.
这个星期六你有没有时间啊？我们去看电影吧。

Zhè-ge	*xīngqiliù*	*nǐ*	*yǒu*	*méi*	*yǒu*
This-CL	Saturday	2SG	have	not	have
kòng	*a?*	*Wǒmen*	*qù*	*kàn*	
free.time	SFP	1PL	go	watch	
diànyǐng	*ba.*				
movie	SFP				

'Are you free this Saturday? Let's go to the movies.' (Adapted from Lee-Wong 1998: 396)

Among the intersubjective uses of SFPs can be included also evidentiality, i.e., the domain encoding the source of information for a given statement (Willett 1988; Aikhenwald 2004), whenever it implies such information to be shared by the speaker with a wider group of people, possibly (but not necessarily) including the hearer (Nuyts 2006: 14). Evidential uses have been identified for a number of Cantonese SFPs, including *ge5* (Li 2006; Sybesma and Li 2007) and *wo3* (Yap *et al.* 2014), and more recently for the Mandarin SFP *ma* 嘛. In Lepadat (2017), I argue – following previous analyses of the particle in terms of marker of states of affairs that are "highly evident in nature" (Chappell and Peyraube 2016: 323) – that it presents a certain evaluation as information shared among the members of a society, i.e., it represents a marker of interpersonal evidentiality (Tantucci 2013). This can be observed in (8) below.

(8) 他们的城堡一般都是依在山坡高的地方而建。[…] A: 易守难攻嘛。

Tāmen		*de*	*chéngbǎo*	*yībān*	*dōu*	*shì*	*yī*
3PL		POSS	castle	normally	all	be	near
zài	*shānpō*		*gāo*	*de*	*dìfāng*	*ér*	*jiàn.*
at	slope		high	ATTR	place	to	build

Yì-shǒu-nán-gong	*ma.*
Easy-defend-difficult-attack	SFP

'Their castles are usually built on high slopes. Easy to defend and hard to attack, of course.'
(Adapted from Lepadat 2017: 258-259)

Interestingly, it has been shown that from the perspective of diachronic development new intersubjective functions that SFPs might acquire in time follow a unidirectional path of change, i.e., new polisemies increasingly oriented towards the expression of the awareness of the addressee develop later than subjective or propositional meanings (Traugott and Dasher 2002; Tantucci and Wang 2020a). Similarly, the same unidirectional path of diachronic development has been argued to be followed in the acquisition process, for children (and arguably learners in general) tend to acquire in first instance literal meanings or functions and only successively those grounded in social cognition (Tantucci and Wang 2020b; Tantucci 2021; Lepadat 2022).
From the angle of discourse organisation, in addition to conveying specific illocutionary forces, it has been argued that SFPs may either express relationships between clauses (Chu 1998) or signal a specific information-structural distribution (Qiang 2011). As far as the latter function is concerned, Qiang (2011) argues that SFPs are used to mark a topic-comment structure and are therefore disallowed in right dislocations and other marked orders of information structure, as can be seen in (9), involving a comment-topic order[7].

(9) A: 这电影你看过吗？

Zhè-ge	*diànyǐng*	*nǐ*	*kàn-guò*	*ma?*
This-CL	movie	2SG	see-ASP	SFP

'Have you seen this movie?'

[7] It is beyond the purpose of this paper to assess the feasibility of this important claim. On this occasion, the discussion will be limited to reporting the array of functions connected with the SFPs that have been identified in the literature.

B: 我看过，这电影 (*呢/吧/么/啊)。
Wǒ *kàn-guò,* *zhè* *diànyǐng* (**ne/ba/me/a*).
1SG see-ASP this movie SFP
'I've seen it, this movie.' (Adapted from Qiang 2011: 192)

Lastly, Chu (1998; 2006; 2009) argues that SFPs can serve the purpose of increasing the relevance of the utterance they mark in the context, thus giving coherence to the discourse. Such is the case of *ne* in (10), wherein it is argued to signal that the speaker needs to look back for contrast (Chu 2006: 18).

(10) 他们还不看电视，还不如我呢！
Tāmen *hái* *bú* *kàn* *diànshì,* *hái*
3PL still not watch television still
bù *rú* *wǒ* *ne!*
not equal 1SG SFP
'They still don't watch TV. So, they are not as [up to the times] as I am!' (Chu 2006: 18)

The different functions attributed to the SFPs in the literature are summarised in Table 1 below.

<table>
<tr><th>Domain</th><th>Functions</th><th>Source</th></tr>
<tr><td rowspan="3">Discourse structure</td><td>Illocutionary force/ sentence type</td><td>Zhu 1982; Paul and Pan 2017; Qi 2002</td></tr>
<tr><td>Discourse organization</td><td>Chu 1998; Chu 2009; Li 2006</td></tr>
<tr><td>Information structure</td><td>Qiang 2011; Lepadat 2017</td></tr>
<tr><td rowspan="2">Subjectivity</td><td>Speaker attitude</td><td>Liu et al. 2001; Paul and Pan 2017</td></tr>
<tr><td>(Epistemic) modality</td><td>Chu 1998; Liu et al. 2001</td></tr>
<tr><td rowspan="2">Intersubjectivity</td><td>Politeness/Rapport management</td><td>Lee-Wong 1998; Tantucci and Wang 2018</td></tr>
<tr><td>Evidentiality</td><td>Yap et al. 2014; Lepadat 2017</td></tr>
</table>

Table 1. Functions of SFPs in the literature.

5.2 Other utterance-final pragmatic markers

Studies focusing on UFPMs in Mandarin are very rare, despite the abundance of such expressions in spoken language (Lepadat 2021). Since Miracle (1991), an increasing number of works has been devoted to pragmatic markers in general, with some *passim* information that can be retrieved concerning their use in utterance-final position. In addition to *wǒ juédé* (Liu 2009; Lim 2011; Endo 2013) and (*Nǐ*) *zhīdào* (*ba/ma*) 你知道吧/吗 'you know' (Tao 2003; Liu 2006) – two among the most frequent UFPMs (Lepadat 2021) – other utterance-final uses that have been mentioned in the literature are concerned with pragmatic markers such as *jiùshì* (*shuō*) 就是说 'that is (to say)' (Biq 2001), *zhēnshì* 真是 'really (is)' and *shízài shì* 实在是 'indeed (is)' (Wu and Biq 2011), *fǎnzhèng* 反正 'anyway' (Zhou and Bao 2014), *ránhòu* 然后 'then' (Wang 2018), *jiùshìle* 就是了 'that's it/all' (Shi 2019), *kǒngpà* 恐怕 'I'm afraid' and *búguò* 不过 'though' (Yap *et al.* 2014).

Among the very few studies specifically tackling UFPMs are those by Yap and her research group (Yap *et al.* 2010; Yap *et al.* 2014), in which common pathways of grammaticalization are identified for both SFPs and other UFPMs (*utterance tags* in their terminology). This is of particular relevance for the ongoing discussion, since it points into the direction of a strong affinity between different types of pragmatic markers found at the right periphery. As an illustration, it is possible for both SFPs and other UFPMs to have developed through a process of clausal integration or clause-combining (Givón 1985). SFPs such as *éryǐyǐ* 而已矣, *bàle* 罢了 and *hǎole* 好了, are argued to have emerged via the syntactic restructuring of a bi-clausal unit into a mono-clausal unit integrating the original evaluative terminal clause (Yap *et al.* 2014: 190). Similarly, UFPMs such as the mitigative *jiùshìle* are also argued to have originated from bi-clausal constructions, with the terminal result clause becoming syntactically and prosodically integrated into the initial clause (Shi 2019).

Additional sources of information are represented by two lines of research who address UFPMs indirectly, i.e., those on right dislocations (Bourgerie 1991, 1998; Song 2018; Lepadat 2021, *i.a.*) and increments (Luke and Zhang 2007; Lim 2014, *i.a.*). Although both are concerned with elements that are not necessarily (and fully) grammaticalized in utterance-final position but simply respond to specific contingencies of the ongoing discourse, they offer valuable insight concerning the array of elements that are likely to have acquired or to be in the process of developing (inter)subjective meanings at the utterance right periphery.

Song (2018) addressed adverbial elements occurring at the right periphery, providing a rather extensive corpus-based list of elements. Among these are adverbs that can be classified as pertaining to Hancil and other's (2015) b), c), and d) types. Furthermore, Lepadat (2021) also provides an extensive discussion of the elements occurring at the utterance right periphery, encompassing elements belonging to Hancil and other's (2015) a) to d) types. Moreover, comment clauses occurring in utterance-final position similar to those discussed for English in Traugott (2015) and Haselow (2016) have been discussed in Lim (2014) and Lepadat (2021), *i.a.*

Based on the consulted literature, Table 2 puts together several examples of elements pertaining to all the four types of final particles identified by Hancil and other (2015) (but it disregards the fifth type corresponding to SFPs and discussed in Section 5.1 above), in addition to the comment clauses tackled by Traugott (2015) and Haselow (2016), to give a comprehensive view of the UFPMs that can be used at the utterance right periphery in Mandarin Chinese[8].

[8] Question tags, general extenders and vocatives identified by Haselow (2016) as part and parcel of the right periphery have been left out from the discussion, since it is not yet completely clear to what extent they can be conflated with pragmatic markers, in spite of the many overlaps existing between their uses (Carter 2004). In particular, question tags are frequently endowed with an independent illocutionary force (Huddleston and Pullum 2002), whereas the use of vocatives and general extenders as pragmatic markers has been rarely addressed in the literature on Chinese.

Type	**Examples**
Adverb-type markers (Hancil *et al.* 2015)	*hǎoxiàng* 好像 'it seems', *qíshí* 其实 'actually', *jiǎnzhí* 简直 'simply', *jūrán* 居然 'unexpectedly', *nándào* 难道 'is it possible that', *dàgài* 大概 'probably', *dàodǐ* 到底 'in the end', *zhōngyú* 终于 'finally', *háishì* 还是 'after all', *yuánlái* 原来 'as a matter of fact', *sìhū* 似乎 'seemingly', *yěxǔ* 也许 'perhaps', *zhǐhǎo* 只好 'have no choice but', *jiūjìng* 究竟 'actually; after all', *gēnběn* 根本 'at all', *dāngrán* 当然 'certainly', *dàyuè* 大约 'probably', *kǒngpà* 恐怕 'I fear' (Yap *et al.* 2014; Shi 2018)
Focus particle-type markers (Hancil *et al.* 2015)	*jiù* 就 'just', *cái* 才 'only', *yòu* 又 'again', *hái* 还 'in addition', *yě* 也 'also', *zài* 再 'again', *dōu* 都 'all' (Shi 2018)
Conjunction-type and conjunct/ adverbial connector-type markers (Hancil *et al.* 2015)[9]	*fǎnzhèng* 反正 'anyway/nevertheless', *búguò* 不过 'but/however/though', *érqiě* (*shì*)而且(是) 'in addition', *jiùshì*(*le*) 就是(了) 'that's it/all', *rúguǒ*(*shì*) 如果(是) 'if (that's the case)' (Lim 2014; Shi 2019; Lepadat 2021)
Comment clauses (Traugott 2015; Haselow 2016)	*wǒ júede* 我觉得 'I think', *wǒ xiǎng* 我想 'I think', *wǒ gǎnjúe* 我感觉 'I feel/think', *nǐ shuō* 你说 'you tell me', *nǐ rènwéi* 你认为 'you think', *nǐ yào zhīdào* 你要知道 'you must know', *nǐ zhīdào* 你知道 'you know', *wǒ tīng shuō* 我听说 'I heard' (Lu 1980; Liu 2006; Xi and Zhang 2008; Lim 2014; Yap *et al.* 2014; Lepadat 2021)

Table 2. Types of expressions found at the right periphery in the literature.

As can be observed from the table, elements of the adverbial type present the largest inventory, including UFPMs that can fulfil several (inter)subjective functions. Starting from subjectivity, epistemic evaluations of a state of affairs can be conveyed through adverbs such as *dàgài* 大概, *sìhū* 似乎, and *kǒngpà*, as well as by means of 1st person comment clauses such as *wǒ júede* or *wǒ xiǎng* 我想. An example is given below in (11), wherein *kǒngpà* is argued to serve an epistemic function expressing a low degree of certainty:

[9] The two groups are treated together here because of the fuzzy boundary between conjunctions and linking adverbs in Mandarin (Liu 2016), in addition to the low frequency of such elements appearing at the utterance right periphery.

(11) 他不喜歡我了，恐怕。

Tā *bù* *xǐhuān* *wǒ* *le,* *kǒngpà.*
3SG NEG like 1SG SFP fear
'He doesn't like me anymore, probably (< I'm afraid).'
(Yap *et al* 2014: 195)

Further subjective uses include the expression of a particular stance on behalf of the speaker, which can be conveyed through several adverbial-type UFPMs such as *jiǎnzhí* 简直 or *jūrán* 居然, but also through focus particles such as *jiù* 就, *cái* 才, *dōu* 都 or *yòu* 又. In (12) below, *yòu* is argued to convey a sense of disapproval:

(12) B: 不是后来小江也换单位了吗? 'Didn't Xiao Jiang also change jobs afterwards?'
A: 上哪儿了, 又?
Shàng *nǎ'r* *le,* *yòu?*
Go.up where SFP again
'Where did she go, this time?' (Lepadat 2021: 255)

As far as interpersonality is concerned, both the evidential and the rapport management functions identified for SFPs can be also performed through other UFPMs. In (13), the reportative expression *wǒ tīng shuō* 我听说 is employed to mark the utterance as information the hearer has come to know based on what (s)he heard from someone else:

(13) 他回来了，我听说。

Tā *huílái-le,* *wǒ* *tīng-shuō.*
3SG.M return-PERF 1SG hear-say
'He came back, I heard.' (Lu, 1980: 33)

On the other hand, the UFPM *wǒ júede* in (14) is argued in Endo (2013) not just to perform an epistemic function – connected more with the utterance-initial uses of the expression – but also to solicit the hearer's agreement and alignment of views.

(14) Ming:这-太不公平了我觉得。

Zhè-	*tài*	*bù*	*góngpíng*	*le*	*wǒ*	*juéde.*
This	too	NEG	fair	SFP	1SG	think

'This- is too unfair I think.'

Li:对啊我们应该考自己的语言嘛。'Right, we should be tested on our own language.'

(Adapted from Endo 2013: 28)

UFPMs that can be employed for discourse structuring purposes include those specifying relationships between clauses (e.g., *búguò*, *fǎnzhèng*, *érqiě shì*), those imposing specific information-structural readings (e.g., *jiù*, *jiùshì(le)*), as well as those managing turn alternation. In example (15) below, *érqiě shì* 而且是 modifies the logical relation between the current and the preceding clause by instantiating a sequential relationship between the two (Lim 2014).

(15) Matt: 但是我觉得你那个是买得挺好的。

Dànshì	*wǒ*	*juéde*	*nǐ*	*nà-gè*	*shì*	*mǎi*	*de*	*tǐng*	*hǎo*	*de.*
But	1SG	think	2SG	that-CLF	be	buy	DE	quite	good	SFP

'But I think the one you bought is quite a good buy.'

[…]

Matt: 你比我早一年买，而且是。

Nǐ	*bǐ*	*wǒ*	*zǎo*	*yī*	*nián*	*mǎi,*	*érqiě*	*shì.*
2SG	COMP	1SG	early	one	year	buy,	furthermore	be

'You bought it one year earlier than me, as well.'

(Adapted from Lim 2014: 228-229)

Moreover, a foregrounding function has been called upon for several UFPMs, including *nǐ zhīdào ma* 你知道吗 (Hu 2015) and *jiùshì*(*le*) (Shi 2019), in addition to markers of the focus-particle type. According to Shi (2019), *jiùshìle* in (16) below is unstressed and represents background information, despite its utterance-final position, while the preceding clause represents foregrounded – i.e. focal – information.

(16) 你只想如何把工作搞好就是了。

Nǐ	*zhǐ*	*xiǎng*	*rúhé*	*bǎ*	*gōngzuò*	*gǎo*
2SG	only	think	how	DISP	job	do
hǎo	*jiùshì*		*le.*			
well	just.be		SFP			

'You just (need to) think about how to get the job done, that's all.'
(Adapted from Shi 2019: 43)

Lastly, UFPMs that can perform turn-yielding or response-inviting functions are several and include *wǒ júede*, *nǐ zhīdào ma* and *nǐ shuō* 你说. In example (17) below, *nǐ rènwéi* 你认为 is used in turn-final position to invite the addressee to give his/her opinion on the matter under discussion.

(17) 达到预期效果了吗? 你认为。

Dádào	*yùqí*	*xiàoguǒ*	*le*	*ma?*	*Nǐ*	*rènwéi.*
Reach	expected	effect	PERF	SFP	2SG	think

'Have we reached the expected result, you think?'
(Xi and Zhang 2008: 12)

Overall, what can be concluded from the above discussion is that Mandarin Chinese possesses a wide inventory of expressions that can be used at the utterance right periphery to express several (inter)subjective functions in addition to SFPs. Although different expressions might be positioned at different points along the grammaticalization path, it is undoubtedly true that Chinese too can make use of different types of UFPMs that are available in other languages of the world (Hancil *et al.* 2015).

5.3 Co-occurrence

The co-occurrence of elements pursuing similar functions in the utterance has been taken into account with respect to a limited number of linguistic phenomena. One case is represented by Tantucci and Wang (2018), who identify different dimensions of what they call 'illocutional concurrences'. These encompass modal expressions, usage-based instantiations of face and SFPs used in evaluative speech acts (Tantucci and Wang 2018: 65).

Despite SFPs and the other UFPMs discussed above showing clear areas of overlap, their co-occurrence has hardly ever been taken into account. One exception is Lepadat (2021), who dedicates one specific section to tackle the intersection between SFPs and what she terms sentence-final expressions to include different elements occurring at the utterance right periphery, including UFPMs.
From the discussion therein contained it can be observed that UF-PMs can co-occur with SFPs to further strengthen or mitigate the illocutionary force of the utterance, thus the two types of markers appear to act as (possibly redundant) strategies aiming to fulfil similar roles. The example in (18) shows the UFPM *wǒ jiù júede* 我就觉得 'I really think' being used to further reinforce the function of the SFP *a*, which conveys a strong personal involvement on the part of the speaker.

(18) Context: Speaker A is talking with B about her sister's love affairs and explains that her mother also asked her to talk to her.
A: 反正得慎重啊, 我就觉得, 对吧.

Fǎnzhèng	*děi*	*shènzhòng*	*a,*	*wǒ*	*jiù*
Anyway	must	be.discreet	SFP,	1SG	just
juéde,	*duì*	*ba.*			
think	right	SFP			

'Anyway [she] needs to be discreet, I really think, isn't it so?'
(Adapted from Lepadat 2021: 251)

An opposite function appears to be performed by the UFPM *dàgài* expressing approximation. In (18), it occurs together with the SFP *ba*, which has been argued to convey a roughly similar sense of speculation on behalf of the speaker.

(19) 走了吧，大概。

Zǒu-le	*ba,*	*dàgài.*
Leave-PERF	SFP	probably

'(He) has probably left.'
(Bourgerie 1998:141)

6. Conclusions

This paper has addressed pragmatic markers occurring at the utterance right periphery in Mandarin Chinese, a topic that has been showed to deserve special attention, on a par with what is occurring for other Indo-European languages (Traugott 2015). After introducing the major studies concerning the right periphery from a typological perspective or in relation to other languages – mostly English –, it emerged that different types of pragmatic markers are employed for turn management, illocutionary force and epistemic status negotiation, expression of politeness, etc. (Hancil *et al*. 2015; Haselow 2016). As far as Asian languages – and in particular Chinese – are concerned, it is generally acknowledged that they dispose of a well-defined world class of sentence-final particles to express several discourse-structural, subjective and intersubjective meanings, which have been the focus of most of the scholarly work on right-peripheral markers.

After careful examination of the relevant literature, however, it has been pointed out that in addition to sentence-final particles, Mandarin Chinese also makes use of a number of other pragmatic markers at the end of the utterance in order to perform roughly equivalent meanings. What is more, studies such as Lepadat (2021) have additionally tackled the co-occurrence of the different types of expressions at the right periphery, showing how (inter)subjectivity can be expressed – sometimes redundantly – by means of several comparable devices. This is an important fact to be acknowledged both from a language-specific and from a typological perspective.

In the final analysis, despite the prominent use of utterance-final pragmatic markers in spoken Mandarin (Lepadat 2021), their features and functions have been rarely investigated outside the area of sentence-final particles. Nonetheless, it has been shown that more extensive and in-depth analyses of the elements (co-)occurring at the utterance right periphery are necessary in order to reach a better understanding of their functions, including from the perspective of the (a)symmetries with left-peripheral elements.

References

Aijmer, K. (2013) *Understanding Pragmatic markers. A Variational Pragmatic Approach*. Edinburgh, Edinburgh University Press.

Aikhenvald, A. (2004) *Evidentiality*. Oxford, Oxford University Press.

Badan, L. (2015) "Left and right peripheries in Mandarin". *Faits de Langues*, 46(1): 23-38.

--- (2020) "Italian Discourse Markers: The Case of Guarda te". *Studia Linguistica. A Journal of General Linguistics*, 74(2): 303-336.

Bazzanella, C. (2001) "Segnali discorsivi e contesto", in W. Heinrich and C. Heiss (Eds.) *Modalità e Substandard*. Bologna, CLUEB: 41-64.

Beeching, K. and U. Detges (2014) "Introduction", in K. Beetching and U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden, Brill: 1-23.

Biq, Y. (2001) "The Grammaticalization of Jiushi and Jiushishuo in Mandarin Chinese". *Concentric: Studies in English Literature and Linguistics*, 27(2): 53-74.

Blakemore, D. (1987) *Semantic Constraints on Relevance*. Oxford: Blackwell.

--- (2002) *Relevance and Linguistic Meaning*. Cambridge/London, Cambridge University Press.

Bourgerie, D.S. (1991) "Postposed Modals in Chinese: Evidence for Restructuring". Paper presented at the 1991 Meeting of the Chinese Language Teachers Association (Washington D.C., November).

--- (1998) "Expanding the Scope of the Sentence-Final Position: Postposed Modals in Cantonese", in S. Matthews (Ed.) *Studies in Cantonese Linguistics*. Hong Kong, Linguistic Society of Hong Kong: 133-146.

Brinton, L.J. (1996) *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/New York, De Gruyter.

--- (2017) *The Evolution of Pragmatic Markers in English: Pathways of Change.* Cambridge, Cambridge University Press.

Chappell, H.M. and A. Peyraube (2016) "Mood and Modality in Sinitic Languages", in J. Nuyts and J. van der Auwera (Eds.) *The Ox-*

ford Handbook of Modality and Mood. Oxford, Oxford University Press: 296-329.

Chu, C. (2009) "Relevance and the Discourse Functions of Mandarin Utterance-Final Modality Particles". *Language and Linguistic Compass*, 3(1): 282-199.

--- (1998) *A Discourse Grammar of Mandarin Chinese*. New York, Peter Lang Publishing.

--- (2006) "A Contrastive Approach to Discourse Particles: A Case Study of the Mandarin UFP Ne". *Journal of Foreign Languages*, 3(3): 7-29.

De Cristofaro, E., C. Crocco, L. Badan and K. Plevoets (2022) "Discourse Markers and Turn-Planning at the Pragmatics-Prosody Interface: The Case of Allora in Spoken Italian". *Journal of Pragmatics*, online preview.

Degand, L. (2014) "'So Very Fast Then' Discourse Markers at Left and Right Periphery in Spoken French", in K. Beetching and U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden, Brill: 151-178.

Degand, L., B. Cornillie and P. Pietrandrea (2013) "Modal Particles and Discourse Markers: Two Sides of the Same Coin?", in L. Degand, B. Cornillie and P. Pietrandrea (Eds.) *Discourse Markers and Modal Particles. Categorization and description*. Amsterdam, John Benjamins: 1-18.

Dér, C.I. (2010) "On the Status of Discourse Markers". *Acta Linguistica Hungarica*, 57(1): 3-28.

Detges, U. and R. Waltereit (2014) "Moi Je ne Sais pas vs. Je ne Sais pas Moi: French Disjoint Pronouns in the Left vs. Right Periphery", in K. Beeching and U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery. Crosslinguistic Investigations of Language Use and Language Change*. Leiden, Brill: 24-46.

Endo, T. (2013) "Epistemic Stance in Mandarin Conversation: The Positions and Functions of Wo Juede (I Feel/Think)", in Y. Pan and D.Z. Kádár (Eds.) *Chinese Discourse and Interaction: Theory and Practice*. Briston, Equinox: 12-34.

Fang, M. 方梅 (2000) "自然口语中弱化连词的话语标记功能". *Zhongguo Yuwen*, 5: 459-470.

Feng, G. (2008) "Pragmatic Markers in Chinese". *Journal of Pragmatics*, 40(10): 1687-1718.

--- (2010) *A Theory of Conventional Implicature and Pragmatic Markers in Chinese*. United Kingdom, Emerald Group Publishing.

Fischer, K. (2006) "Towards an Understanding of the Spectrum of Approaches to Discourse Particles: Introduction to the Volume", in K. Fischer (Ed.) *Approaches to Discourse Particles*. Amsterdam, Elsevier: 1-20.

Frascarelli, M. (2017) "Dislocations and Framings", in A. Dufter and E. Stark (Eds.) *Manual of Romance Morphosyntax and Syntax*. Berlin/Boston, De Gruyter: 472-501.

Fraser, B. (1990) "An Approach to Discourse Markers". *Journal of Pragmatics*, 14(3): 383-398.

--- (1996) "Pragmatic Markers". *Pragmatics*, 6(2): 167-190.

Giorgi, A. (2009) *About the Speaker: Towards a Syntax of Indexicality*. Oxford, Oxford University Press.

Givón, T. (1985) "Iconicity, Isomorphism and Non-Arbitrary Coding in Syntax", in J. Haiman (Ed.) *Iconicity in Syntax*. Amsterdam, John Benjamins: 187-220.

Goldberg, A.E. (2006) *Constructions at Work: The Nature of Generalization in Language*. Oxford, Oxford University Press.

Guo, J. (1999) "From Information to Emotion: The Affective Function of Right-Dislocation in Mandarin Chinese". *Journal of Pragmatics*, 31(9): 1103-1128.

Hancil, S., A. Haselow and M. Post (Eds.) (2015) *Final Particles*. Berlin/München/Boston, De Gruyter.

Haselow, A. (2012) "Subjectivity, Intersubjectivity and the Negotiation of Common Ground in Spoken Discourse: Final Particles in English". *Language and Communication*, 32(3): 182-204.

--- (2016) "Intensifying Adverbs 'Outside the Clause': A Cognitive Analysis", in G. Kaltenböck, E. Keizer and A. Lohmann (Eds.) *Outside the Clause*. Amsterdam/Philadelphia, John Benjamins Publishing Company: 379-416.

Heine, B., G. Kaltenböck, T. Kuteva and H. Long (2021) *The Rise of Discourse Markers*. Cambridge, Cambridge University Press.

Hu, J. 胡建锋 (2015) "前景化与'知道吗'的功能". *Yuyan Kexue*, 14(2): 194-205.

Kaltenböck, G., B. Heine and T. Kuteva (2011) “On Thetical Grammar”. *Studies in Language*, 35: 852-897.
Lee-Wong, S.M. (1998) “Face Support – Chinese Particles as Mitigators: A Study of Ba A/Ya and Ne”. *Pragmatics*, 8(3): 387-404.
Lepadat, C. (2017) “The Modal Particle *Ma* 嘛: Theoretical Frames, Analysis and Interpretive Perspectives”. *Quaderni di Linguistica e Studi Orientali*, 3: 243-70.
--- (2021) *Not Just Postposed Topics: An Integrated Pragmatic Account of the Sentence-Final Slot in Mandarin Chinese*. PhD Dissertation, Sapienza University of Rome.
--- (2022) “Vittorio Tantucci: Language and Social Minds: The Semantics and Pragmatics of Intersubjectivity. Cambridge University Press, 2021”. *Applied Linguistics*, amac008: 1-4.
Li, B. (2006) *Chinese Final Particles and the Syntax of the Periphery*. PhD Dissertation, Leiden University.
Li, C.N. and S.A. Thompson (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley/Los Angeles/London, University of California Press.
Lim, N.-E. (2011) “From Subjectivity to Intersubjectivity: Epistemic Marker Wo Juede in Chinese”, in Y. Xiao, L. Ta and H.L. Soh (Eds.) *Current Issues in Chinese Linguistics*. New Castle, Cambridge Scholars Press: 265-300.
--- (2014) *Retroactive Operations: On ‘increments’ in Mandarin Chinese conversation*. PhD. Dissertation, University of California.
Liu, B. (2009) “Chinese Discourse Markers in Oral Speech of Mainland Mandarin Speakers”, in Y. Xiao (Ed.) *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21)*. Smithfield (Rhode Island), Bryant University: 358-374.
Liu, L. 刘丽艳 (2006) “汉语标记 ‘你知道’”. *Zhongguo Yuwen*, 5: 423-432.
Liu, Y. 刘月华, W. Pan 潘文娱 and W. Gu 故伟 (Eds.) (2001) 使用现代汉语语法. Beijing, Shangwu Yinshuguan.
Lombardi Vallauri, E. (2009) *La struttura informativa: forma e funzione negli enunciati linguistici*. Roma, Carocci editore.
Lu, J. 陆俭明 (1980) “汉语口语句法里的易位现象”. *Zhongguo Yuwen*, 1: 28-41.

Lü, S. 吕叔湘 (1942) 中国文发要略. Beijing, Shangwu Yinshuguan.

Luke, K.K. (1990) *Utterance Particles in Cantonese Conversation*. Amsterdam, John Benjamins Publishing Company.

Luke, K.K. and W. Zhang (2007) "Retrospective Turn Continuations in Mandarin Chinese Conversation". *Pragmatics*, 17(4): 605-635.

Miracle, W.C. (1991) *Discourse Markers in Mandarin Chinese*. PhD Dissertation, The Ohio State University.

Munaro, N. and C. Poletto (2002) "Ways of Clausal Typing". *Rivista di Grammatica Generativa*, 27: 87-105.

Nuyts, J. (2006) "Modality: Overview and Linguistic Issues", in W. Frawley (Ed.) *The Expression of Modality*. Berlin/New York, De Gruyter: 1-26.

Onodera, N. (2014) "Setting Up a Mental Space: A Function of Discourse Markers at the Left Periphery (LP) and Some Observations about LP and RP in Japanese", in K. Beetching and U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery.* Leiden/Boston, Brill: 92-116.

Paul, W. and V. Pan (2017) "What you See is What you Get: Chinese Sentence-Final Particles as Head-Final Complementisers", in V. Struckmeier and J. Bayer (Eds.) *Discourse Particles – Formal Approaches to their Syntax and Semantics*. Berlin/Boston, De Gruyter: 49-77.

Qi, H. 齐沪扬 (2002) 语气词与语气系统. Hefei, Anhui Jiaoyu Chubanshe.

Qiang, X. 强星娜 (2011) "话题标记与句类限制". *Yuyan Kexue*, 10(2): 190-204.

Rouchota, V. (1998) "Procedural Meaning and Parenthetical Discourse Markers", in A.H. Jucker and Y. Ziv (Eds.) *Discourse Markers*. Amsterdam, John Benjamins Publishing Company: 97-126.

Rühlemann, C. (2019) *Corpus Linguistics for Pragmatics. A guide for Research*. London/New York, Routledge.

Samek-Lodovici, V. (2015) *The Interaction of Focus, Givenness, and Prosody*. Oxford, Oxford University Press.

Schiffrin, D. (1987) *Discourse Markers*. Cambridge, Cambridge University Press.

Schourup, L. (1982) "Common Discourse Particles in English Conversation", in *Working Papers in Linguistics 28.* The Ohio State University.

Shei, C. (2014) *Understanding the Chinese Language*, London/New York, Routledge.
Shi, F. 石飞 (2019) "句末'就是了'的话语立场与话语功能". *Chinese Language Learning*, 6: 39-46.
Simpson, A. (2014) "Sentence-Final Particles", in C.T.J. Huang, A.Y. Li and A. Simpson (Eds.) *The Handbook of Chinese Linguistics*. West Sussex, Wiley B: 156-179.
Song, S. 宋少萌 (2018) "汉语口语中的副词后置现象及生成动因探讨". *Zhongguo Yuwen Yanjiu*, 74: 105-125.
Speas, P. and C. Tenny (2003) "Configurational Properties of Point of View Roles", in A.M. Di Scullo (Ed.) *Asymmetry in Grammar*. Amsterdam, John Benjamins Publishing Company: 315-344.
Sybesma, R. and B. Li (2007) "The Dissection and Structural Mapping of Cantonese Sentence Final Particles". *Lingua*, 117: 1739-1783.
Tantucci, V. (2013) "Interpersonal Evidentiality: The Mandarin V-过 guo Construction and Other Evidential Systems Beyond the 'Source of Information'". *Journal of Pragmatics*, 57: 210-230.
--- (2021) *Language and Social Minds: The Semantics and Pragmatics of Intersubjectivity*. Cambridge, Cambridge University Press.
Tantucci, V. and A. Wang (2018) "Illocutional Concurrences: The Case of Evaluative Speech Acts and Face-Work in Spoken Mandarin and American English". *Journal of* Pragmatics, 138: 60-76.
--- (2020a) "Diachronic Change of Rapport Orientation and Sentence-Periphery in Mandarin". *Discourse Studies*, 22(2): 146-173.
--- (2020b) "From Co-Actions to Intersubjectivity throughout Chinese Ontogeny: A Usage-Based Analysis of Knowledge Ascription and Expected Agreement". *Journal of Pragmatics*, 167: 98-115.
Tao, H. 陶红印 (2003) "从语音，语法和话语特征看'知道'格式在谈话中的演化". *Zhongguo Yuwen*, 4: 291-302.
Thompson, S.A. and A. Mulac (1991) "A Quantitative Perspective on the Grammaticization of Epistemic Parentheticals in English", in E.C. Traugott and B. Heine (Eds.) *Approaches to Grammaticalization.* Amsterdam, John Benjamins Publishing Company: 313-329.
Traugott, E.C. (2010) "Subjectification, Intersubjectification and Grammaticalization", in K. Davids, L. Vandelanotte and H. Cuy-

ckens (Eds.) *(Inter)subjectivity and (Inter)subjectification: A Reassessment*. Berlin/New York, De Gruyter: 29-70.
--- (2014) "On the Function of the Epistemic Adverbs Surely and No Doubt at the Left and Right Peripheries of the Clause", in K. Beeching and U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery.* Leiden, Brill: 72-91.
--- (2015) "Investigating 'Periphery' from a Functionalist Perspective". *Linguistics Vanguard*, 1(1): 119-130.
--- (2022) *Discourse Structuring Markers in English: A Historical Constructionalist Perspective on Pragmatics*. Amsterdam, John Benjamins Publishing Company.
Traugott, E.C. and R.B. Dasher (2002) *Regularity in Semantic Change.* Cambridge, Cambridge University Press.
Van Olmen, D. and J. Šinkuniene (Eds.) (2021) *Pragmatic Markers and Peripheries*. Amsterdam/Philadelphia, John Benjamins Publishing Company: 143-170.
Wang, L. 王力 (1985) 中国现代语法. Beijing, Shangwu Yinshuguan.
Wang, W. (2018) "Discourse Uses and Prosodic Properties of Ranhou in Spontaneous Mandarin Conversation". *Chinese Language and Discourse*, 9(1): 1-25.
Wang, W. and H. Tao (2020) "From Matrix cClause to Turn Eexpansion: The Emergence of Wo Juede 'I Feel/Think' in Mandarin Conversational Interaction", in Y. Maschler, S. Doehler Pekarek, J. Lindström and L. Keevallik (Eds.) *Emergent Syntax for Conversation*. Amsterdam/Philadelphia: John Benjamins Publishing Company: 151-182.
Wichmann, A., A.-M. Simon-Vandenbergen and K. Aijmer (2010) "How Prosody Reflects Semantic Change: A Synchronic Case Study of of Course", in K. Davidse, L. Vandelanotte and H. Cuyckens (Eds.) *Subjectification, Intersubjectification and Grammaticalization*, Berlin/New York, De Gruyter: 103-154.
Willett, T.L. (1988) "A Cross-Linguistic Survey of the Grammaticalization of Evidentiality". *Studies in Language*, 12(1): 51-97.
Wu, A.Y.-R. and Y.-O. Biq (2011) "Lexicalization of Intensifiers: Two X- shi Constructions in Spoken Mandarin". *Chinese Language and Discourse*, 2(2): 168-197.
Yap, F.H., J. Wang, and C.T. Lam (2010) "Clausal Integration and the

Emergence of Mitigative and Adhoratative Sentence Final Particles in Chinese". *Taiwan Journal of Linguistics*, 8(2): 63-84.

Yap, F.H., Y. Yang and T.-S. Wong (2014) "On the Development of Sentence Final Particles (and Utterance Tags) in Chinese", in K. Beeching and U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden, Brill: 1-257.

Zhou, Y. 周玉 and L. Bao 暴丽颖 (2014) "汉语标记语'反正'语用新说". *Jixi Daxue Xuebao*, 11: 135-138.

Zhu, D. 朱德熙 (1982) 语法讲义. Beijing, Shangwu Yinshuguan.

CANTONESE TONES AND TONE MARKS: HOW PAST STUDIES CAN HELP PRESENT LEARNERS

Luisa M. Paternicò

"L'Orientale" University of Naples

The history of Cantonese phonology, and in particular that of Cantonese tones, departs very early from Mandarin and needs to be told separately. The earliest materials for Cantonese language teaching report a variable total of twelve, eleven, ten, nine or seven tones with different, quite industrious ways to mark them around the characters or transcribed syllables. Nowadays it is quite accepted that learning six (or seven) tones is enough to pronounce Cantonese syllables properly. However, since there is no definitive, official phonetic transcription of Cantonese sounds, different tone-marking systems still coexist.

After an analysis of the main works used for the didactics of Cantonese in China and abroad in the 19th and 20th centuries, this paper will try to answer the following questions: how many tones were recognized in Cantonese in the language materials (dictionaries, grammars, primers, phrasebooks etc.) compiled by Western missionaries and officials? How were they described? How were they marked? Can these works

help to devise a better system to mark Cantonese tones for foreign learners today?

The present study, first of all, will sketch the history of Cantonese tones to the present day. Afterwards, it will present an overview of early Cantonese tone descriptions and tone-mark systems devised by Westerners to better teach and acquire them. Finally, it will show how good use of these past studies can be made, by proposing a more suitable, iconic way to mark the tones that, according to a recent survey, could benefit second language learners. The preliminary results obtained can be used for further studies on the new tone-marks and their application in classes of Cantonese as a second/foreign language[1].

1. Brief history of Cantonese tones and traditional tone-marks

Sinitic languages are tonal languages, where different tone inflections or contours are fundamental to convey different meanings. As Sagart (1999: 2) states: "Chinese tones are complex suprasegmental events distinguishing (otherwise) homophonous syllables, which are often – but not always – meaning-bearing units, or morphemes". A tone contour is defined by its pitch level or register (which can be high, medium or low) and by its pitch movement, from the beginning of the tone until its end. The movement can be level, rising or falling (Bauer and Benedict 1997: 111).

The historian Shen Yue 沈約 (441-513) is believed to be the first to have recognized and described the Chinese tonal system in his *Sìshēngpǔ* 四聲譜, labelling the tones as: *píngshēng* 平聲 "level tone", *shǎngshēng* 上聲 "rising tone", *qùshēng* 去聲 "departing tone", and *rùshēng* 入聲 "entering tone". Only the first three categories were contrasting tones, while the last one was used only to refer to syllables ending with a glottal consonant stop *(-p*, *-t*, or *-k*) (Bauer and Benedict 1997: 154). This tonal system was immediately used to arrange rhyme dictionaries as in Lu Fayan's 陸法言 *Qie Yun* 切韻 of 601 (Sagart 1999: 3-4). Starting from the Song dynasty, the traditional method to mark the tones was the *sìshēng quāndiǎn* 四聲圈點 or *sìjiǎo biāoquānfǎ* 四角標圈法

[1] For space and time reason, the phenomenon of tone-sandhi in Cantonese will not be dealt with in the present study.

'four tone/corner circle marking', where the four tones were indicated adding by hand a circle or a semicircle in one of the four corners of a sinogram (Branner 1997: 251; Chung 2016: 757), for example:

°上　　去°

。平　　入。

Also, in order to remember the position of the tone diacritics, the later *Kāngxī zìdiǎn* 康熙字典 (1716) would use the image below:

Figure 1. How to remember the four tones in *Kāngxī zìdiǎn*.

However, already in the eighth century, the northern and southern tonal systems had begun to part. In 880, the Japanese Buddhist monk Annen 安然 (ca. 841-915), in his treatise on Sanskrit phonology, wrote about the existence of a northern Chinese tone system, where the *píng* and *shǎng* tones had split, while words with *shǎng* tone and voiced initial stops and affricates or fricative had merged with words with *qù* tone. In 847, two monks, Master Yuishō 惟正 and in 877 Master Chisō 智聰, returned to Japan reporting that in southern China all the four tones had split into two. This presumably happened because the originally voiced stop and affricate initial consonants had lost their voicing, becoming voiceless (Chang 1975: 637-639; Bauer and Benedict 1997: 154-157).

Therefore, at this point in history, Cantonese, as well as other southern dialects, had seen its tone number more than doubled in a *yīn* 陰 set and in a *yáng* 陽 set. The word *yīn* was used in traditional Chinese linguistics to mark the high and middle registers, while *yáng* was used for the low registers (Bauer and Benedict 1997: 121). In Cantonese, in particular, the entering tone had a further subdivision: not only did it split into *yīnrù* 陰入 and *yángrù* 陽入, but *yīnrù* also split into two. This further division presumably happened because of vowel length: syllables with phonetically short vowels moved into the "high" entering tone category, and syllables with long vowels into the "mid" entering tone category (Bauer and Benedict 1997: 155). Table 1 below shows modern Cantonese tones as they developed from ancient Chinese:

		平聲	上聲	去聲	入聲
陰	上 中	**55/52 (53, 51)**	**35 (25)**	 **33**	**5 33**
陽	下	**21 (31)**	**23 (13, 24, 25)**	**22**	**2/22**

Table 1. Cantonese tones developed from ancient Chinese.

In the table above, the values outside the parentheses are those indicated by Y.R. Chao 趙元任 (1892-1982) in his *Cantonese Primer* (Chao 1947: 24), that are mainly confirmed by Yue-Hashimoto (1972: 92) and others. Bauer and Benedict (1997: 116) and Matthews and Yip (2011: 28) both report the high rising as being changed in recent years to 25. Other studies set the tone contours slightly higher or lower (see values in parentheses): Yue-Hashimoto (1972: 92), for example, defines the low rising as 24; Cheung (2007: 59) has the low falling as 31 and the low rising as 13. Recent studies on Hong Kong Cantonese have also proved a trend to merge (in production) or to not distinguish (in reception) the high and the low rising tones, the mid level and the low level tones, or the low falling and the low level tones (Mok *et al.* 2013).

2. The tones in modern Cantonese: variety, number, and tone marks

Lexical tones are a very important contrastive feature in modern Cantonese phonology. They provide a "second phonetic dimension – in addition to consonants and vowels – for contrasting words" (Bauer and Benedict 1997: 109). It is already quite accepted nowadays that modern Cantonese has a system of six basic tones. The 'discount' from the previous acknowledged ten/nine/seven tone systems (see Section 3) is due to mainly two factors: 1. in the past, different tones were assigned to syllables with finals in *-p*, *-t* or *-k* (the entering tone, *rùshēng*, of Middle Chinese), whereas this distinction is not made anymore because it is not contrasting: the vowels, and therefore the tones, are simply shorter in those syllables, but the tone contour is the same; 2. several studies have proved that the high-level and the high-falling tones are not distinctive anymore in the majority of Cantonese speakers (Bauer and Benedict 1997: 111; Lau 2010: 1; Matthews and Yip 2011: 27-29; see also Chu and Taft 2011: 58-61). Thus, the current situation can be synthetized as in Table 2 below:

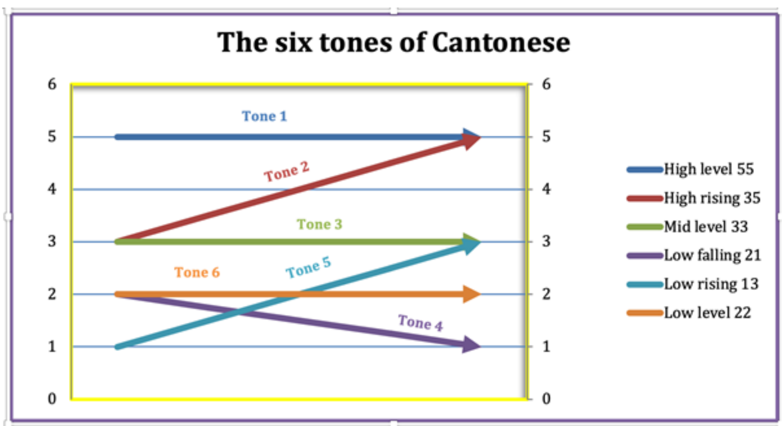


Table 2. Modern Hong Kong Cantonese tones.

As the graphic shows, we can state that in modern Hong Kong Cantonese there are three level tones: high (55), mid (33) and low (22), which are the reference for the other three: the high rising (35), the low rising (13) and the low falling (21). These values mainly follow Cheung (2007: 59) – except for reporting 21 instead of 31 for the low falling tone – and Lam and Paternicò (2017: 10).

Concerning the tone marks, according to the Yale transcription system – judged by the author the best available, so far, for teaching Cantonese as second/foreign language (Lam-Paternicò 2017: XVIII) – the tone is graphically signalled on the vowel of a syllable or on the first vowel of a diphthong, for example: *fā* (high level → first tone), *séui* (high falling → second tone). In the case of the nasal syllables *m* and *ng* the tone is exceptionally placed on the consonant, ex. ḿ, or on the first one of two consonants, ex. *ńg*. The mid level → third tone is the only one that is not graphically marked. The three lower tones are signalled by adding letter *h* after the vowel or the diphthong. Below, a description of the six tones with tone marks in Yale.

First tone or high level (55):
The first tone is a flat, high-pitched sound emission which does not change its course. It is marked with a macron on top of the vowel, for example: *wāi* (威 'powerful').

Second tone or high rising (35):
The second tone is a high rising tone. It starts at a lower pitch and ends at a slightly higher point. It is graphically marked with an acute accent on top of the vowel, for example: *wái* (毀, 'to destroy').

Third tone or mid level (33):
The third tone is a flat, medium-pitched sound emission which does not change its course. It is lower than the first tone. It is not marked because it is similar to natural pitch in non-tonal languages, for example: *wai* (慰, 'to console').

Fourth tone or low falling (21):
The fourth tone has a falling course. It starts quite low and descends even lower, like a sigh. It is marked by a grave accent on the vowel, and by letter *h* after the vowel or after the second vowel of a diphthong, for example: *wàih* (唯, 'only').

Fifth tone or low rising (13):
The fifth tone has a rising course, like the second, but starts lower and does not end quite as high. It is marked by an acute accent on the vowel, and by letter *h* after the vowel or after the second vowel of a diphthong, for example: *wáih* (偉, 'great').

Sixth tone or low level (22):
The sixth tone is a flat, low-pitched sound emission which does not change its course. The only graphic mark for this tone is letter *h* after the vowel or after the second vowel of a diphthong, for example: *waih* (胃, 'stomach').

As the next paragraph will show, the earliest romanization systems were devised by Westerners in southern China in the 19th century. Nowadays, there is not one univocally recognized standard, and several systems coexist (Cheng and Tang 2016). The above-mentioned Yale transcription was devised by Parker Po-fei Huang and Gerald P. Kok in the 1960s for the manual *Speak Cantonese* and other materials produced at Yale University, from which it takes its name (Huang and Kok 1960). From that moment, it has been widely used in Cantonese language teaching materials for foreigners. It has quite a few advantages: it looks similar to the *Pinyin* transcription for Mandarin; it was created for English speakers, so it is simpler for foreigners in general; the tone marks are iconic and easier to learn compared to other systems with numerical or letter notation. Its only arbitrary sign would be letter *h* for the lower range tones. It originally included a seventh tone, the high-falling, which is not precepted as contrastive anymore, as we have already said (Cheng and Tang 2016: 39-41).
Another two systems, created in the 1960s were: the Hong Kong Government Cantonese Romanization and the Sidney Lau system. The first is based on E.J. Eitel's dictionary (I ed. 1877, II ed. 1910) and of J. Dyer Ball's language materials (1883; 1886). However, it merely aimed at transcribing anthroponyms and toponyms to use in street signs. It was not meant to be a teaching tool (although it is the only system many locals know). It cannot be considered a system *stricto sensu*, since it lacks internal coherence, and there is no correspondence from grapheme to phoneme and vice versa. The tones are not marked

(Kataoka and Lee 2008). The second was created by the dean of the Hong Kong Government Language School, Sydney Lau Sek-cheung 劉錫祥, for the radio series "Cantonese by radio", and later on used in his Cantonese primers published in the 1970s. This system is an adaptation of the former romanization devised by the missionaries B.F. Meyer and T.F. Wempe in 1934, in turn inspired – with exception of tone marks – by Chalmers (1870). It quite resembles the Yale system, except for some finals and for the way of marking tones, with superscript numbers (Lau 1972).

However, the most widespread transcription system today, especially in Hong Kong and mainland China, is *Jyutping* 粵拼, a romanization created in 1993 by the Linguistic Society of Hong Kong. Very similar to the Yale system, it differs in the way it transcribes five finals and two initials, and because it marks the tones with numbers (Cheng and Tang 2016: 38-43, 46; Lam and Paternicò 2017: XVII-XVIII).

As it is possible to see from the most common romanization systems listed above, and from others which are less used or no longer used nowadays (Cheng and Tang 2016: 43-36) (see also Section 4), with the exception of Yale, the most common way to mark the tones in the second half of the 20th century was with numbers. Although many native speakers are used to *Jyutping* and its tone marks, the numerical tone notation might result quite arbitrary and hard to distinguish for foreign learners. Is there a better way to mark tones to help Cantonese language teaching? If we look further back in the history of Cantonese tone marks, can we find helpful insights? The following section will present an overview of the early tone descriptions and tone marks devised by and for foreigners for the purpose of learning the language.

3. Early Western descriptions of Cantonese tones and ways to mark them

In the nearly 100 years between the second decade of the 19th and the 20th century, an impressive amount of language materials of all kinds (primers, grammars, dictionaries etc.) were created to teach and learn Cantonese. Most of them were printed in Hong Kong, but also abroad, in Southeast Asian countries or in the Americas. In the following pages, we will make an overview on the way Cantonese tones were described

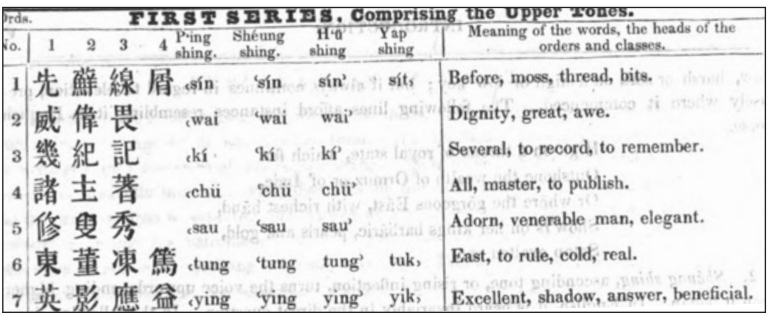
and marked in a selection of sources which were mainly printed in southern China by missionaries and officials. These sources were quite popular among the foreigners who had settled in the area, had several reprints, and influenced later romanization systems for Cantonese.
The earliest dictionary of Cantonese and a foreign language was Robert Morrison's *A Vocabulary of the Canton dialect* (1828). The work, in three volumes, does not present tone marks. We must wait until Elijah Coleman Bridgman (1801-1861), the first American Protestant missionary in China (Paternicò 2020: 176-177), to see the original Chinese tone mark system applied in a Western work. In the introduction to his *Chinese Chrestomathy in the Canton Dialect* (1841: III-IV), Bridgman states that for intonation, his work refers to a small duodecimo tonic dictionary commonly called *Fan-wan*[2], compiled at the beginning of the Qing dynasty by Yu Xuepu 虞學圃 and Wen Qishi 溫岐石. Bridgman describes a set of four main tones (平上去入) that are sub-divided in an upper series (陽) and a lower series (陰):

> *P'ing shing*, or the monotone, is an even sound, and is fairly represented by a horizontal line. It is characterized by an entire absence of all inflection of voice. It may be quick or slow, harsh or soft, on a high or low key; but it always continues in regard to elevation, precisely where it commenced […].
> *Shéung shing*, ascending tone, or rising inflection, turns the voice upwards, ending higher than it began. In English, it is heard invariably in the direct question […].
> *H'ü shing*, the receding tone, is a distinct prolongation of the voice, diminishing, while it is prolonged; in English, it is used in the indirect question, and also in the language of scorn, surprise, authority, and alarm […].
> *Yap shing*, the entering tone, is an abrupt, short sound. In English, it occurs in words like hit, hop, &c […].

[2] The complete is: *Kong-ú ch'ik-tuk, Fan-wan ts'iú hòp tsáp* 江湖尺牘分韻撮要合集 *Collection of Jianghu Correspondence and of Characters Divided According to their Rhymes*. It can be accessed online here: https://ctext.org/wiki.pl?if=gb&res=94354 (visited 2023/02/20).

> Each of these four tones, being subdivided, form the two series, […] one of which may be considered primary, and the other secondary, the two being parallel to each other. The four tones, forming the first series, may be regarded as the fundamental modifications of the voice; and the four forming the second series, […] may be described in precisely the same terms as the first four tones, since they are distinguished from them only by being uttered on a lower key. (Bridgman 1841: IV-V)

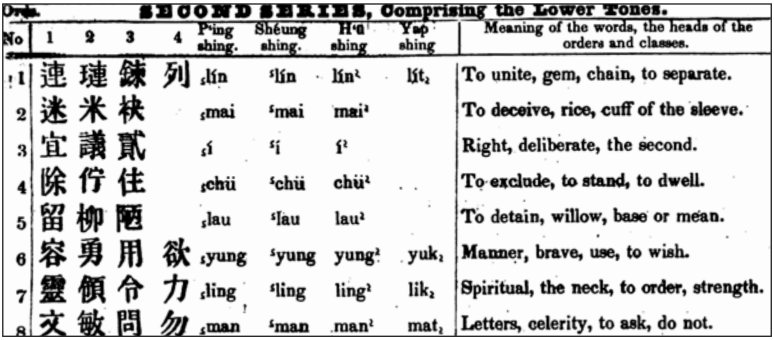
Inspired by the Chinese traditional system marking the tones around the characters, in Bridgman the tones of the upper series are marked with semicircles at the four corners of the Romanised syllables; for those of the lower series, a dash is added beneath the semicircle (see examples in Figures 2-3).



FIRST SERIES, Comprising the Upper Tones.

Ords. No.	1	2	3	4	P'ing shing.	Shéung shing.	Hü shing	Yap shing	Meaning of the words, the heads of the orders and classes.
1	先	蘚	線	屑	꜀sín	꜂sín	sín꜄	sít꜆	Before, moss, thread, bits.
2	威	偉	畏		꜀wai	꜂wai	wai꜄		Dignity, great, awe.
3	幾	紀	記		꜀kí	꜂kí	kí꜄		Several, to record, to remember.
4	諸	主	著		꜀chü	꜂chü	chü꜄		All, master, to publish.
5	修	叟	秀		꜀sau	꜂sau	sau꜄		Adorn, venerable man, elegant.
6	東	董	凍	篤	꜀tung	꜂tung	tung꜄	tuk꜆	East, to rule, cold, real.
7	英	影	應	益	꜀ying	꜂ying	ying꜄	yik꜆	Excellent, shadow, answer, beneficial.

Figure 2. Tone-marks for the upper tones in Bridgman (1841: VI).



SECOND SERIES, Comprising the Lower Tones.

Ords. No	1	2	3	4	P'ing shing.	Shéung shing.	Hü shing	Yap shing	Meaning of the words, the heads of the orders and classes.
1	連	璉	鍊	列	꜁lín	꜃lín	lín꜅	lít꜇	To unite, gem, chain, to separate.
2	迷	米	袂		꜁mai	꜃mai	mai꜅		To deceive, rice, cuff of the sleeve.
3	宜	議	貳		꜁í	꜃í	í꜅		Right, deliberate, the second.
4	除	佇	住		꜁chü	꜃chü	chü꜅		To exclude, to stand, to dwell.
5	留	柳	陋		꜁lau	꜃lau	lau꜅		To detain, willow, base or mean.
6	容	勇	用	欲	꜁yung	꜃yung	yung꜅	yuk꜇	Manner, brave, use, to wish.
7	靈	領	令	力	꜁ling	꜃ling	ling꜅	lik꜇	Spiritual, the neck, to order, strength.
8	文	敏	問	勿	꜁man	꜃man	man꜅	mat꜇	Letters, celerity, to ask, do not.

Figure 3. Tone-marks for the lower tones in Bridgman (1841: VII).

Bridgman's way of numbering, defining, and marking the tones is mirrored by Samuel Wells Williams (1812-1884), an American missionary who had been sent to China in 1833 in order to take care of the typography of the American Board of Commissioners for Foreign Missions in Canton (Paternicò 2020: 177). Both in his primer *Easy Lessons in Chinese* (Wells Williams 1842) and in his *A Tonic Dictionary of the Chinese Language in the Canton Dialect* (Wells Williams 1856), Wells Williams decided to keep the Chinese traditional system. In the introduction to his dictionary, he provides a reason for this choice, distancing himself from the marks adopted in works for Mandarin and other dialects from the time of Matteo Ricci (1552-1610) onward. He writes:

> The eight tones are clearly distinguished by the Cantonese, and every word in this Dictionary has been marked according to its tone by semicircles, something like the mark used by the Chinese, as here exhibited. [...] As the tones in Chinese are totally distinct from accent, in their own or other languages, there are many objections to using the common and generally understood marks (ˊˋˇˆ) on the tops of the vowels to denote them, as has been done in Romanizing some dialects, and in the dictionaries of De Guignes, Medhurst, and others. By taking another sign, there is no mixing of radically different symbols over the same syllable [...] where one of the marks affects the power of the vowels, and the other denotes the tone of the Chinese character. Since diacritical marks in all European languages modify the power of the vowels, it is desirable not to introduce any confusion in writing words, the more so as the tones in Chinese are so entirely different; it is better to adopt a new symbol. (Wells Williams 1856: XXIX)

He also provides a very clear chart for his readers (Figure 4) and a detailed description of the 8 tones.

Figure 4. The tones as in Wells Williams (1856: XXIX).

Bridgman's and Wells Williams' tone systems were also used in the works of the German Lutheran missionary Wilhelm Lobscheid (1822-1893): *Grammar of the Chinese Language* (1864) and *English and Chinese Dictionary* (1866-1868).

The Scottish missionary John Chalmers (1825-1899) made a finer description of the tones and of the regional differences (Paternicò 2017a: 233-234). He also introduced some variations to Williams' system. In the introduction to his dictionary *An English and Cantonese Dictionary* (1870), a very popular text that in 1891 was at its sixth edition (here quoted), he states:

> There are only eight regular tones in the Chinese language, and even these are not all found in many dialects. The upper series of tones is printed in Italics, and the lower series in Roman letters.
>
> 1. Syllables without any tonal mark are in the 'even' tone.
> 2. Those with a tonal mark (‘) on the upper left corner are in the rising tone.
> 3. Those with a tonal mark (’) on the upper right corner are in the going tone.
> 4. And words ending in *k*, *p* or *t* are in the abrupt tone; that is to say, they end in a close consonant which cannot be sung out. (Chalmers 1891ed.: IV-V)

Chalmers also provides a scheme which compares his system with William's (Figure 5).

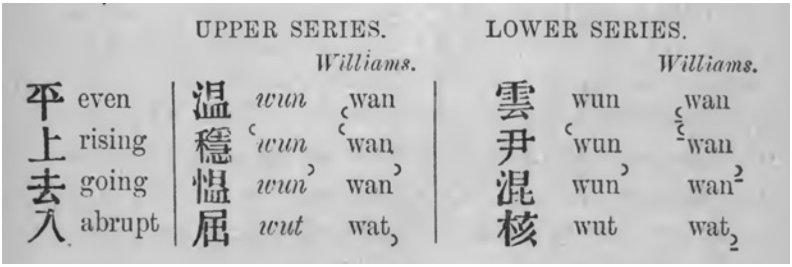


Figure 5. Chalmers tone marks compared to Williams' (Chalmers 1891 ed.: V).

Chalmers was therefore the first in the history of Cantonese tone notation to decide to leave those syllables he considered natural speech (even tones) unmarked. Also, he differentiated the upper and lower tone-series simply writing the former in Italics. The result is a script with less diacritics and less burdening for the eye. However, the use of such a system could prove quite confusing for a learner.
Chalmers was also the first to acknowledge the merit of E.H. Parker (1849-1926) who, in an article written on the *Overland China Mail* (October 27, 1874), drew the attention on "additional tones" that can be heard in Cantonese colloquial, and on tone sandhi, for which he devises new tone marks, that Chalmers synthesized as follows:

> First, there are three even tones instead of two, the uppermost one being a sort of cry. [...] They are distinguished in this Dictionary by a tonal mark in the lower left corner of the syllable aside its being printed in italics. I have been guided entirely by a Cantonese teacher in this and the following distinctions of tones.
> Second, there are three abrupt tones instead of two [...] a middle abrupt tone takes place in connection with long vowels. [it is not marked]"
> Third, all the tones, except perhaps the upper even and the upper abrupt, are liable to be changed into an exag-

> gerated rising tone. [...] This is indicated by an asterisk on the upper left side of the syllable.

In 1877, Ernest Johann Eitel (1838-1908), a Protestant missionary of the Basel Mission (Paternicò 2020: 182), compiled a dictionary which would be considered a fundamental tool for generations: *A Chinese Dictionary in the Canton dialect*, London (1877-1887), in 2 volumes (II ed. 1910-1911). In the *Introduction*, he retraces the history of Chinese and "vernacular" tones and ton-splits (Eitel 1877: XXIV) and then proposes a system of 4 tone classes, each of which is split into two, with the exception of the fourth that is split into three, thus a total of 9 tones are recognized, as in the figure 6 below:

I. Tone class.			
꜀[] Upper even tone (上平)	e. g.	呼	꜀fú.
꜁[] Lower even tone (下平)	„	扶	꜁fú.
II. Tone class.			
꜂[] Upper rising tone (上上)	„	府	꜂fú.
꜃[] Lower rising tone (下上)	„	婦	꜃fú.
III. Tone class.			
[]꜄ Upper departing tone (上去)	„	富	fú꜄.
[]^2 Lower departing tone (下去)	„	父	fú^2.
IV. Tone class.			
[]꜀ Upper entering tone (上入)	„	竹	chuk꜀
[]$_{\circ}$ Medial entering tone (中入)	„	捉	chuk$_{\circ}$
[]$_2$ Lower entering tone (下入)	„	濁	chuk$_2$

Figure 6. Eitel's 9 tone system (1877: XXV-XXVI).

However, he adds that in daily speech, other tones might be heard, which are derived from the above basic tones or simply "modulated under the influence of emphasis or emotion" (Eitel 1877: XXIV).

He decides not to take into consideration the medial even tone 中平, since – according to him – the difference with the other two even tones is not of practical importance. As we can see from figure 5, Eitel preserves the traditional tone marks adding a circle "as in the Siam and Annam tradition" for the medial entering tone 中入.
A thorough study of tones and tone-marks in Cantonese was made by James Dyer Ball (1847-1919), son of a missionary in Canton, who worked in the Hong Kong Civil service for 35 years with different positions. Dyer Ball dedicated several works to Cantonese language. In his primer *Cantonese Made Easy*, he makes a long description of the tones (Dyer Ball 1883: XVIII-XXXII) and adopts the system of 10 tones (as in Chalmers) plus tone sandhi, as in Figure 7:

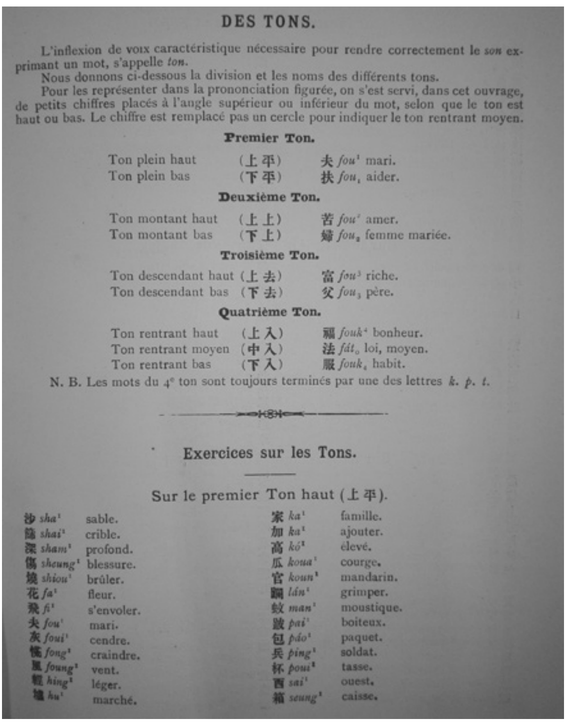
	Upper Series.		*Middle tones.*	*Lower Series.*	
1	上平 Shöng^2 ꜀p'ing.	10	中平 ꜀Chung ꜀p'ing.	下平 Há^2 ꜀p'ing.	5
2	上上 Shöng^2 ꜂shöng.		上聲變音 Shöng^2 ꜀shing pín' ꜀yam.*	下上 Há^2 shöng^2.	6
3	上去 Shöng^2 hui'.			下去 Há^2 hui'.	7
4	上入 Shöng^2 yap$_2$.	9	中入 ꜀Chung yap$_2$.	下入 Há^2 yap$_2$.	8

* This is scarcely a correct name for this tone. It is really a 變音 pín' ꜀yam for the 下平 há^2 ꜀p'ing, 下去 há^2 hui' and any other of the tones which occasionally rise into it.

Figure 7. The tone set and tone marks in Dyer Ball (1883: XXIII).

As for what tone notation is concerned, Dyer Ball makes a review of other systems devised by Westerners. He states that the accent system used for Pekingese or Hakka (with macrons and accents) cannot work for Cantonese because the latter has more tones, there are no typesets available to print them and they are unknown to the locals, with the paradox that a foreign student would need to explain what these marks stand for to his native speaker teacher. He also does not share Chalmers' point of view in leaving some tones unmarked: for Dyer Ball, every syllable needs to have its tone signalled. That is why he prefers to keep the traditional way of marking tones with semicircles and dashes as in Bridgman. For the three newly identified tones, Dyer Ball decides to use a circle at the bottom left corner for 中平, a circle at the bottom right position for 中入 and to signal tone sandhi through an asterisk (Dyer Ball 1883: XXIX-XXXI).

The earliest materials using numerical tone marks for Cantonese (probably under the influence of the Wade-Giles system developed in the second half of the 19th century for Mandarin) were the dictionaries compiled by Louis Aubazac (1871-1919), a MEP missionary who arrived in China in 1894: the 法粵字典 *Dictionnaire Français-Cantonnais* (1902) and the 粵法字典 *Dictionnaire Cantonnais-Français* (1912). Aubazac (1902: VII) describes a system of 4 main tones, divided in high, middle and low registers (for a total of 9 tones), which he marks with superscript or subscript numbers, with the exception for the 中入, which he marks with a subscript circle. He also specifies that the fourth tone is used only in syllables ending with a glottal stop, as Figure 8 below shows:



DES TONS.

L'inflexion de voix caractéristique nécessaire pour rendre correctement le *son* exprimant un mot, s'appelle *ton*.

Nous donnons ci-dessous la division et les noms des différents tons.

Pour les représenter dans la prononciation figurée, on s'est servi, dans cet ouvrage, de petits chiffres placés à l'angle supérieur ou inférieur du mot, selon que le ton est haut ou bas. Le chiffre est remplacé pas un cercle pour indiquer le ton rentrant moyen.

Premier Ton.

Ton plein haut	(上平)	夫 *fou*^{1} mari.
Ton plein bas	(下平)	扶 *fou*$_{1}$ aider.

Deuxième Ton.

Ton montant haut	(上上)	苦 *fou*^{2} amer.
Ton montant bas	(下上)	婦 *fou*$_{2}$ femme mariée.

Troisième Ton.

Ton descendant haut	(上去)	富 *fou*^{3} riche.
Ton descendant bas	(下去)	父 *fou*$_{3}$ père.

Quatrième Ton.

Ton rentrant haut	(上入)	福 *fouk*^{4} bonheur.
Ton rentrant moyen	(中入)	法 *fát*$_{\circ}$ loi, moyen.
Ton rentrant bas	(下入)	服 *fouk*$_{4}$ habit.

N. B. Les mots du 4e ton sont toujours terminés par une des lettres *k. p. t.*

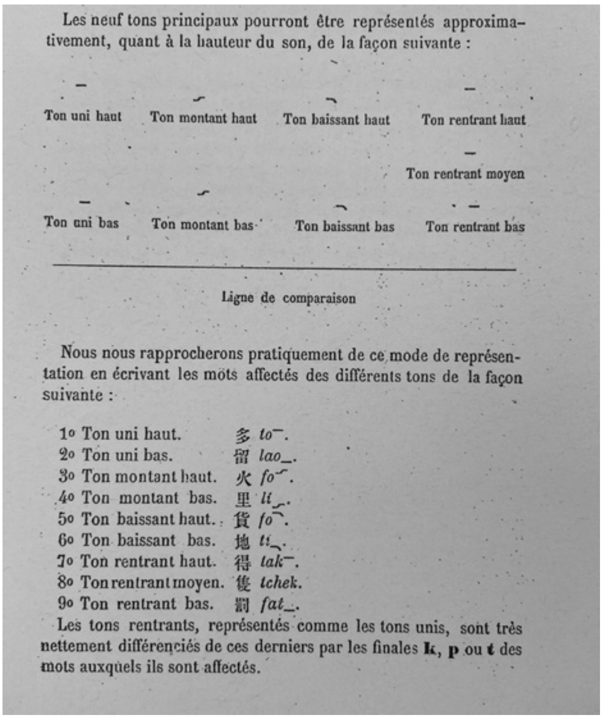
Exercices sur les Tons.

Sur le premier Ton haut (上平).

沙 *sha*^{1}	sable.	家 *ka*^{1}	famille.
篩 *shai*^{1}	crible.	加 *ka*^{1}	ajouter.
深 *sham*^{1}	profond.	高 *kó*^{1}	élevé.
傷 *sheung*^{1}	blessure.	瓜 *koua*^{1}	courge.
燒 *shiou*^{1}	brûler.	官 *koun*^{1}	mandarin.
花 *fa*^{1}	fleur.	攀 *lán*^{1}	grimper.
飛 *fi*^{1}	s'envoler.	蚊 *man*^{1}	moustique.
夫 *fou*^{1}	mari.	跛 *pai*^{1}	boiteux.
灰 *foui*^{1}	cendre.	包 *páo*^{1}	paquet.
慌 *fong*^{1}	craindre.	兵 *ping*^{1}	soldat.
風 *foung*^{1}	vent.	杯 *poui*^{1}	tasse.
輕 *hing*^{1}	léger.	西 *sai*^{1}	ouest.
墟 *hu*^{1}	marché.	箱 *seung*^{1}	caisse.

Figure 8. Aubazac's tone system as in the *Dictionnaire Français-Cantonnais* (1902), p. VII.

To our knowledge, the first to use a more iconic way to mark Cantonese tones was the Lieutenant Leblanc for his *Cours de langue Chinoise Parlée, Dialecte Cantonnais*, published in Hanoi in 1910. Leblanc uses the same 9 tone system as his predecessors but changes the way to mark them, in his words: "*La notation qui nous parait la plus logique est une notation figurative qui représente autant que possible les caractères distinctifs de chaque ton*"[3] (Leblanc 1910: 8). He then presents the set that can be seen in Figure 9:



Les neuf tons principaux pourront être représentés approximativement, quant à la hauteur du son, de la façon suivante :

Ton uni haut — Ton montant haut — Ton baissant haut — Ton rentrant haut

Ton rentrant moyen

Ton uni bas — Ton montant bas — Ton baissant bas — Ton rentrant bas

Ligne de comparaison

Nous nous rapprocherons pratiquement de ce mode de représentation en écrivant les mots affectés des différents tons de la façon suivante :

1o Ton uni haut. 多 *to*‾.
2o Ton uni bas. 留 *lao*_.
3o Ton montant haut. 火 *fo*⌐.
4o Ton montant bas. 里 *li*_.
5o Ton baissant haut. 貨 *fo*¬.
6o Ton baissant bas. 地 *ti*_.
7o Ton rentrant haut. 得 *tak*‾.
8o Ton rentrant moyen. 隻 *tchek*.
9o Ton rentrant bas. 罰 *fat*_.

Les tons rentrants, représentés comme les tons unis, sont très nettement différenciés de ces derniers par les finales **k**, **p** ou **t** des mots auxquels ils sont affectés.

Figure 9. Leblanc's "figurative" tone marks (Leblanc 1910: 9).

It was with the quite popular Meyer-Wempe romanization system (Meyer and Wempe 1935), devised by two Catholic missionaries in

[3] Translation: "The notation we find most logical is a figurative notation, which represents the distinctive traits of each tone as much as possible".

Hong Kong, Bernard F. Meyer and Theodore F. Wempe, that the tone system was reduced to 6 and an attempt was made at marking them with different diacritics that were easily available for printing presses and familiar to a Western eye (see Table 3 below):

Tone	Description	Tone mark	Example
1	high level/high falling	no mark	*ma*
2	high rising	acute accent (´)	*má*
3	mid level	grave accent (`)	*mà*
4	low falling	circumflex (ˆ)	*mâ*
5	low rising	caron (ˇ)	*mǎ*
6	low level	macron (ˉ)	*mā*

Table 3. Meyer-Wempe tones and tone marks.

In this system, the high level and the high falling tones are treated as not distinctive anymore and left unmarked; the macron is used for the low level tone. The grave accent, which one would expect for a falling tone, is instead used for the mid level. All in all, the marks are still not truly iconic.

Another example of the first half of the 20th century comes from Thomas O'Melia's famous primer *First Year Cantonese* (Hong Kong 1939), which was used by the Hong Kong Government as the syllabus for the Final Examination in Cantonese, organized by the Board of Examiners for the so-called Cadet System (Paternicò 2019: 31). O'Melia clearly states that he believes the Cantonese tone system has only 6 tones and decides to mark them slightly differently from Meyer and Wempe see Table 4:

Tone	Description	Tone mark	Example
1	upper even	no mark	*ma*
2	upper rising	acute accent (´)	*má*
3	upper going	grave accent (`)	*mà*
4	lower even	macron (ˉ)	*mā*
5	lower rising	caron (ˇ)	*mǎ*
6	lower going	circumflex (ˆ)	*mâ*

Table 4. O'Melia's tones and tone system (1939: XIX).

It is therefore clear that at the beginning of the 20th century the trend of the previous decades was changing, moving from a conservative attitude, aiming as much as possible at the preservation of the Chinese traditional tone notation, which was familiar to native Chinese teachers, towards a learner-centred approach, which made use of diacritics that could be more familiar to Westerners and easier to reproduce by the press.

4. Later proposals and new trends

In the 1920s-1930s Y.R. Chao began to develop a system of "tone letters" which would be later acknowledged as an option by the IPA and is now universally known (Chao 1930: 25). The system used a vertical line to mark the tone height and a horizontal or slanted line at its left to mark the contour. He also used them to explain the tones in his *Cantonese Primer* (Chao 1947: 24):

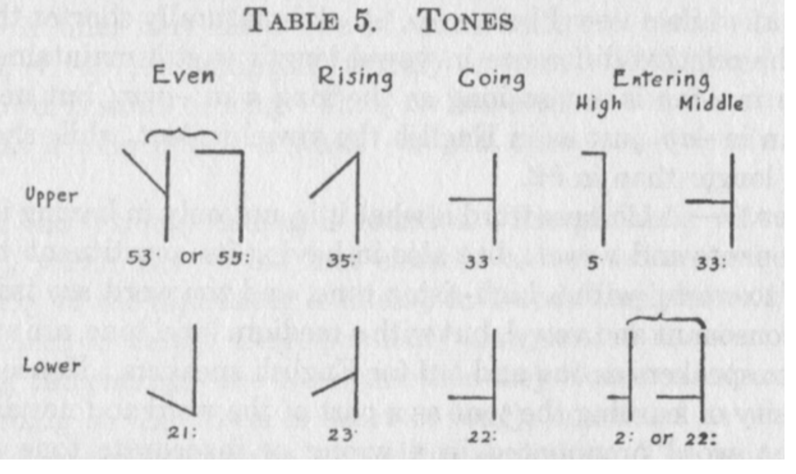


Figure 10. Tone description for Cantonese (Chao 1947: 24).

However, aside from the graphic explanation of the tones, the tone mark system Chao employed in the *Cantonese Primer* is quite a complicated mix of diacritics and other Latin letters, as we can see in the table of comparison with other systems he offers to his reader (Figure 11):

TABLE 13. COMPARISON OF TONES

	Present System	Jones and Woo	Eitel	Ball	Meyer and Wempe
Upper Even	a ai au an ang	ˋa	꜀a	꜀a	a
Upper Rising	ax ae ao anx ag	ˊa	꜂a	꜂a	á
Upper Going	ah ay aw ann aq	¬a	a꜄	a꜄	à
Upper Entering	{High Short vowels	ˋat	at꜆	at꜆	at
	{Low Long vowels	¬a:t	ắt꜆	ắt꜆	aàt
Lower Even	Same as above,	ˏa	꜁a	꜁a	ā
Lower Rising	but with lower-	ˏa	꜃a	꜃a	ǎ
Lower Going	tone initials.	ˍa	a꜅	a꜅	â
Lower Entering		ˍat	at꜇	at꜇	ât
Changed Tones					
Upper Even	a°	ˉˋa	—	꜀a*	—
Others	ha*	ˊa	—	꜂a*	꜂a*

Figure 11. Y.R. Chao's tone marks (Chao 1947: 33).

Convinced that "[…] the constant addition of the necessary diacritical marks and tonal figure makes words and sentences so confusing to the eye that it is not only extremely wasteful of effort, but usually results in the student's inability to gain any clear idea of the sounds of words or to make sure connections between meanings" (Chao 1947: 16), Chao leaves most of the tones unmarked and 'simply' changes the spelling of the syllable's final. Furthermore, he distinguishes the upper and lower series by the initial. The result is a mix of letters that are meant to be pronounced and of other letters that simply indicate the tone, see Figure 12.

TABLE 6 A. EXAMPLES OF UPPER-TONE FORMS

Ending	*Even Tone*	*Rising Tone*	*Going Tone*
zero	ka 'home'	kax 'false'	kah 'frame'
–i	sai 'west'	sae 'wash'	say 'small'
–u	kau 'ditch'	kao 'nine'	kaw 'enough'
–m	kam 'present'	kamx 'so' (manner)	kamm 'so' (degree)
–n	pin 'border'	pinx 'flat, thin'	pinn 'change'
–ng	foang 'square'	foag 'imitate'	foaq 'let go'

TABLE 6 B. EXAMPLES OF LOWER-TONE FORMS

Ending	*Even Tone*	*Rising Tone*	*Going Tone*
zero	ma 'hemp'	max 'horse'	mah 'scold '
–i	lai 'come'	lae 'propriety'	lay 'example'
–u	mou 'have not'	moo 'have not'	mow 'fog'
–m	laam 'blue'	laamx 'behold'	–laamm 'warship'
–n	yün 'round'	yünx 'far'	yünn 'court'
–ng	zeung 'constant'	zeug 'go up'	zeuq 'up'

Figure 12. Upper and Lower tone series in Chao (1947: 26).

The author admits that memorising the charts he proposes is the only way to proceed further in the study. Whether Chao's romanization and tone marks could be useful to teach Cantonese was object of debate in the following years (Barnett 1950; Ulving 1958; Egerod 1958). The scholars, on one hand, felt a deep respect towards Chao and his contribution to linguistics, however, on the other hand, they felt that from a didactic point of view, his romanization could be perfected. In particular, Kenneth M.A. Barnett (1911-1987), a civil servant of the Hong Kong Government, proposed some changes and devised a solution that would be known as the Barnett-Chao romanization system (Barnett 1950). As far as the tones are concerned, this transcription method still uses letters before or after the main vowel or diphthongs to mark the tone: if a letter *h* is placed before the vowel, it means that the start of the tone is high, while after the vowel it indicates that the tone falls; if a letter *r* is placed before the vowel it means that the start of the tone is low, while after the vowel it indicates that the tone rises. The rising or falling tone is also signalled by changing the spelling of the finals (see Table 5 below). The total number of tones is 10, plus one modified tone option. See the example below for syllables *si* and *sin*:

Tone	**Barnett-Chao**		
high-falling	shih	shinn	
mid-rising	sir	sirn	
mid-flat	si	sin	
low-falling	srih	srinn	
low-rising	srir	srirn	
low-flat	sri	srin	
high-entering			shek
mid-entering			sit
low-entering			srek
high-flat	shihx	shinnx	
mid-rising modified	shihv	shinnv	

Table 5. Barnett-Chao Romanization system.

Although it was adopted by the London School of Oriental and African Studies (SOAS) and by the Chinese Language Training Section of the Hong Kong Government from 1963 to 1967, this romanization was considered too complex for teaching Cantonese as a second language (see for example Ulving 1958: 93 and Egerod 1958: 371, both proposing their own 'not less complicated' alternatives) and was, thus, abandoned in favour of Sydney Lau's system, which used superscript numbers for tones and paved the way to the longer lasting fortune of the *Jyutping* romanization (see Section 2).

All in all, the constant trend in the second half of the 20th century was to move towards simplification and against the use of too many diacritics. This has been evident in Hong Kong and southern China in general, where the numerical system prevailed for a while for both Mandarin (with Wade-Giles) and Cantonese. In the following pages we will try to argue that an iconic system would be preferable for foreign learners of Cantonese, in China or abroad, and especially for those who are already familiar with the iconic *pinyin* tone marks adopted for Mandarin.

5. Is it possible to devise a better iconic system to mark Cantonese tones?

The necessity of using or creating a proper transcription system with an accurate way of marking the tones was felt as a strong duty by scholars and teachers of Cantonese in the second half of the 20th century. As Ulving writes:

> The prospect student of Chinese is no doubt aware of some of the difficulties lying ahead of him. He is prepared to meet a sound system with several strange sounds and a wholly unfamiliar phenomenon of significant tone inherent in every syllable, he knows that he will find a grammar unlike anything familiar to him from European languages, and he may shudder at the thought of having to remember thousands of complicated and seemingly arbitrary characters […].

> And while he will resignedly accept the difficulties of pronunciation, grammar and script as inherent in the Chinese language and nothing to be done with, he may feel a justified indignation about unnecessary complexities created by badly devised methods of transcription. (Ulving 1958: 6)

The search history of an unequivocal tone marking system is quite long and controverted for all the languages with lexical tones. As far as Cantonese is concerned, the trend in the last century has been towards simplification in at least four main directions: 1. few simple diacritics, leaving high or mid level tones unmarked; 2. letter tone marks with no diacritics, except for tone sandhi; 3. numerical tone marks; 4. iconic tone marks. We agree with other scholars, that the best way to have an analytical and transparent (for the reader) notation is an iconic tone system (Heselwood 2013: 75), which can describe the tone pitch and contour in a stylised manner. Chao's tone letters, as already explained, are one of the best ways to represent tones iconically, but quite complex to reproduce in a typeset or to use within the phonetic transcription in a handbook to teach Cantonese as a second/foreign/additional language. As a matter of fact, Chao himself did not use them in his primer, preferring letters (Section 4).
The lack of grasp of tone contour can be quite frustrating for those learners whose first language is neither Cantonese nor a tonal language (Li *et al.* 2016: 10-11), however a proper tone marking system could help foreign students learning and remembering Cantonese tones and their contours. In fact, there is evidence that learning Cantonese can be easier to those who have previous literacy in Mandarin and its tones (Li *et al.* 2016: 13-14).
Bearing in mind the above premises and aiming at helping learners of Cantonese as a second/foreign/additional language, we want to propose a new way of marking the 6 tones of Cantonese in a new iconic way. See examples in Figure 13 below:

1st High level tone	ā ē ī ō ū
2nd High rising tone	á é í ó ú
3rd Mid level tone	a e i o u
4th Low falling tone	a̖ e̖ i̖ o̖ u̖
5th Low rising tone	a̗ e̗ i̗ o̗ u̗
6th Low level tone	a̱ e̱ i̱ o̱ u̱

Figure 13. New proposal for Cantonese tone marks.

This new system clearly derives from the Yale Romanization, fixing the total of the tones to 6 (Section 2) and eliminating the arbitrary use of letter *h* to mark the tones of the lower register. This choice has been made mainly because letter *h* is already used in Yale as initial consonant *h-* or as *ch-*, but also because it seems preferable to see a tone mark before or right on top of the sound it belongs to. Drawing a lesson from the past systems devised in the last two centuries, which used superscript and subscript symbols/numbers for the tones, it also occurred to us that placing the diacritics for the lower tones below the vowel would be more appropriate and iconic.

All the tones, in this new system, are marked with diacritics, with exception of the third mid level tone. The tone marks are simple, stylised and transparent to the reader. The high level tone is marked with a macron above the vowel, the high rising tone with an acute accent above the vowel, the low falling tone with a grave accent below the vowel, the low rising with an acute accent below the vowel, the low level with a macron below the vowel.

This system can be especially helpful for foreign learners who are already familiar with Mandarin and its *pinyin* transcription which, having long time ago abandoned the numerical notation of Wade-Giles, makes use of the same tone marks placing them all above the vowel.

A first brief survey recently conducted on 24 Cantonese learners as a foreign language, with prior knowledge of Mandarin, has shown that the majority of them think that this new system might have helped them learning Cantonese tones. In particular, the new system was deemed as 'definitely helpful' by those who had studied Cantonese with Yale transcription (78%), and 'probably helpful' by those who had used *Jyutping* or other romanizations (22%).

This preliminary result is quite encouraging. We therefore intend to carry on further studies, using this new notation system (and others in control groups) directly in teaching Cantonese tones to a larger number of foreign students, in order to prove whether it can be more effective than others in acquiring Cantonese tones.

References

Aubazac, L. (1902) 法粵字典 *Dictionnaire Français-Cantonnais*. Hong Kong, Imprimerie de la Société des Missions Étrangères.

--- (1912) 粵法字典 *Dictionnaire Cantonnais-Français*. Hong Kong, Imprimerie de la Société des Missions Étrangères.

Barnett, K.M.A. (1950) "A Transcription for Cantonese: Notes on Mr. Yuen Ren Chao's 'Cantonese Primer'". *Bulletin of the School of Oriental and African Studies*, 13(3): 725-745.

Bauer, R.S. and P.K. Benedict (1997) *Modern Cantonese Phonology*. Berlin/New York, De Gruyter.

Branner, D. (1997) "Notes on the Beginnings of Systematic Dialect Description and Comparison in Chinese". *Historiographia Linguistica*, 24(3): 235-66.

Bridgman, E.C. (1841) *Chinese Chrestomathy in the Canton Dialect*. Macau: S. Wells Williams.

Chalmers, J. (1870 I ed., 1891 VI ed.) *An English and Cantonese Dictionary*. Hong Kong: Kelly and Walsh.

Chang, K. 張昆 (1975) "Tonal Developments among Chinese Dialects". *The Bulletin of the Institute of History and Philology, Academia Sinica*, XLVI(4): 636-709.

Chao, Y.-R. 趙元任 (1930) "ə sistim əv toun-letəz". *Le Maître Phonétique Troisième Série*, 8(45): 24-27.

--- (1947) *Cantonese Primer*. New York, Greenwood Press.

Cheng, S.P. and S.W. Tang (2016) "Cantonese Romanization", in S-W. Chan (Ed.) *The Routledge Encyclopedia of the Chinese Language*. https://www.routledgehandbooks.com/doi/10.4324/9781315675541.ch3 (visited 2023/02/20).

Cheung, H. 張洪年 (2007 rev. ed.) 香港粵語語法的研究 (增訂版). Hong Kong, The Commercial Press.

Chu, C.K. and M. Taft (2011) "Are There Six or Nine Tones in Cantonese?", in *Proceedings of the Psycholinguistics Representation of Tone Conference*. Hong Kong: Chinese University of Hong Kong: 58-61. https://www.researchgate.net/publication/267544978_ARE_THERE_SIX_OR_NINE_TONES_IN_CANTONESE (visited 2023/02/20).

Chung, K.S. (2016) "Wade-Giles Romanization System", in S-W. Chan (Ed.) *The Routledge Encyclopedia of Chinese Language*, London: Routledge: 756-776.

Dyer Ball, J. (1883) *Cantonese Made Easy*. Hong Kong, China Mail.

--- (I ed. 1886, 1908 II ed) *Cantonese Made Easy Vocabulary*. Hong Kong: China Mail.

Eitel, E.J. (1877-1887) *A Chinese Dictionary in the Canton Dialect*. II vols. London/Hong Kong, Thubner & Co. Lane, Crawford & Co.

Egerod, S. (1958) "The Tonal Spelling of Cantonese". *T'oung Pao*, 46(3-5): 369-375.

Heselwood, B. (2013) *Phonetic Transcription in Theory and Practice*. Edinburgh, Edinburgh University Press.

Huang, P. and G. Kok (1960) *Speak Cantonese*, Book 1. New Haven, Yale University.

Kataoka, S. and C. Lee (2008) "A System Without a System: Cantonese Romanization Used in Hong Kong Place and Personal Names". *Hong Kong Journal of Applied Linguistics*, 11(1): 79-98.

Lam, K.W. and L.M. Paternicò (2017) *Corso di lingua cantonese*. Milano, Hoepli.

Lau, E. (2010) "Tone-Melody Relationship in Cantonese". *Working Papers in Linguistics*, 41(3): 1-12.

Lau, S. (1972) *Elementary Cantonese*, vol. 1. Hong Kong, The Government Printer.

Leblanc, J. (1910) *Cours de langue Chinoise Parlée, Dialecte Cantonnais, 1ere partie, Grammaire*. Hanoi/Haiphong, Imprimerie d'Extreme Orient.

Lobscheid, W. (1864) *Grammar of the Chinese Language*. Hong Kong, Daily Press.

--- (1866-1868) *English and Chinese Dictionary*. Hong Kong, Daily Express Office.

Matthews, S. and V. Yip (2011, 2nd ed.) *Cantonese. A Comprehensive Grammar*. New York, Routledge.

Meyer, B.F. and T.F. Wempe (1935) *The Student's Cantonese English Dictionary*. Hong Kong, St. Louis Industrial School Printing Press.

Mok, P.P.K, D. Zuo and P.W.Y. Wong (2013) "Production and Perception of a Sound Change in Progress: Tone Merging in Hong Kong Cantonese". *Language Variation and Change*, 25: 341-370.

Morrison, R. (1828) *A Vocabulary of the Canton dialect*. Macau, East India Company's Press.

O'Melia, T. (1939) *First Year Cantonese*. Hong Kong, Maryknoll House.

Paternicò, L.M. (2017a) "Studiare il *Punti*: I materiali linguistici per l'apprendimento della lingua cantonese compilati a cavallo tra XIX e XX secolo", in T. Pellin and G. Trentin (Eds.) *Atti del XV Convegno AISC, Macerata, settembre 2015*. Venezia, Cafoscarina: 223-243.

--- (2017b) "I primi occidentali alle prese con la lingua cantonese". *Sulla via del Catai*, 16: 109-119.

--- (2019) "Shaping Cantonese Grammar – Early Western Contribution". *Histoire Epistémologie Langage*, 41(1) :15-37.

--- (2020) "Sinitic Languages in the English Periodical Press of 19th Century China: Focus on *The Chinese Repository* and *The China Review*". *International Journal of China Studies*, 11(2): 347-370.

Sagart, L. (1999) "The Origin of Chinese Tones", in *Proceedings of the Symposium/Cross-Linguistic Studies of Tonal Phenomena/Tonogenesis, Typology and Related Topics*. Tokyo, Tokyo University of Foreign Studies: 91-104.

Ulving, T. (1958) "The Transcription of Cantonese: A Critical Review of Some Current Systems of 'Tonal Spelling', and a Presentation of a New Rational Transcription". *T'oung Pao*, 46(1-2): 81-110.

Wells Williams, S. (1842) *Easy Lessons in Chinese*. Macao, Chinese Repository.

--- (1856) *A Tonic Dictionary of the Chinese Language in the Canton Dialect*. Canton, Office of the Chinese Repository.

Yue-Hashimoto, A. (1972) *Phonology of Cantonese*. Cambridge, Cambridge University Press.

A TWO-TIERED ANALYSIS OF CHINESE POLITICAL DISCOURSE: THE CASE OF XI JINPING'S COMMEMORATIVE SPEECH FOR THE CENTENNIAL OF THE CCP*

Carlotta Sparvoli

Ca' Foscari University of Venice

Chiara Romagnoli

Roma Tre University

1. Introduction

This paper proposes a discourse analysis of the speech delivered by President Xi Jinping during the ceremony celebrating the centenary of the Chinese Communist Party (CCP) at Tian'anmen Square in Beijing on 1 July 2021. The reasons why this speech deserves a separate in-depth analysis are to be found in the representativeness it displays as both official and political discourse as well as in its significance given the unique occasion on which it was delivered.

The goal of this paper is to shed light on the rhetorical devices and lexical resources employed in this seminal speech. More specifically, we intend to highlight the specific persuasion strategies conveyed by the modal expressions and to identify the semantic and distributional

---

* Although this paper was collaboratively produced, Sections and Sub-sections 2.1, *3.1.1*, *3.1.3*, 4.2 and 5 were written by Carlotta Sparvoli only, and Section and Sub-sections 1, 2.2, *3.1.2*, 3.2 and 4.1 by Chiara Romagnoli only.

features of the modal verbs and adverbs employed in this seminal political discourse. The crucial role of the modalized expressions visible in the "nine-musts" articulated in the medial part of the speech points to the relevance of modal disambiguation. However, to single out the specific semantics at work, a further level of analysis is in order, aimed at identifying the features of the lexical environment characterizing this speech.

We thus propose a two-tiered analysis that intersects the disambiguation of the modal expressions on necessity and possibility with the identification and scrutiny of the most frequent lexical items, aimed at identifying their specific usages, semantic features and regularities in terms of collocates and collocations. Particular attention will also be paid to the interaction between modals and other operators, in particular evaluative adverbs marking the speaker's stance.

In order to quantify the occurrences of the vocabulary employed by the speaker, we used the corpus query tool in Sketch Engine and created a corpus consisting of 3606 Chinese words out of the total 7275 Chinese characters comprised in the speech. The small dimension of the corpus justified the type of analysis we have carried out, which is predominantly qualitative. Nevertheless, some quantitative considerations will be also offered to capture the main figures of the words selected. The data so retrieved will then be combined with the modal distribution of 46 modal items across the main rhetoric units (narrative, prescriptive argumentation, conclusion) and further discussed to highlight the pragmatic value of the modalized expressions in the speech.

Section 2 illustrates the theoretical frameworks adopted for the classification of modal expressions and adverbial operators. After a brief synopsis of the speech, Section 3 presents the data, starting from the features of the corpus collected, the target items selected and shows how they are used in political discourse, also including an overview of the modal distribution. Section 4 discusses the data combining the pragmatic role of the modalized proposition in each rhetoric unit with the features of the co-occurring adverbial operators. Section 5 presents the conclusions.

2. Theoretical framework

2.1 The modal taxonomy

Modality denotes the state of affairs as *non-factual, non-actualized* (Lü 1942), is paradigmatically related to the notion of possibility and necessity (Van der Auwera and Plungian 1998) and is closely linked to different strategies adopted for convincing someone to prefer something over something else (Portner 2009). Political discourse typically addresses "what is to come" and engages in exhortations to think of a change as 'possible' or to accept an unpopular policy as 'inevitable'. Semantic modal taxonomies, therefore, constitute an effective tool for detecting methods of persuasion and, ultimately, provide essential indicators for intercepting the passage from the discourse on 'a possible change' to that on 'an inescapable future'. Concerning Chinese, the studies already carried out in this field demonstrate how politicians manipulate discourse in order to express their view, although the few works published in English are mainly based on Taiwan political affairs (Cheng 2019). Less has been done on mainland China political discourse, with the exception of Li (2018) which describes how modality is reconstructed in Chinese-English government press conference interpreting and therefore adopts a comparative perspective.

Compared to the above-mentioned studies, we use a more fine-grained taxonomy based on Sparvoli (2012). The latter adopts the framework of Van der Auwera and Plungian (1998) and spells each modal meaning based on Lü Shuxiang (1942). The modal classes and prototypical markers are shown in Table 1[1];the rationale of the primary divisions is clarified in the ensuing paragraphs.

[1] The labels used for each modal meaning and the list of Chinese modal items is based on Sparvoli (2012: 295), for the necessity domain, and Sparvoli (2017: 41) for that of possibility. Each modal meaning is capitalized.

Modality	Subclass	Domain	Modal meaning	Prototypical items
Participant-internal (Dynamic)		POS	(learned) Ability	*huì* 会 'be able'
		POS	Success	Potential construction, *néng* 能, *nénggòu* 能够 'manage, succeed'
		NEC	Need	*xūyào* 需要 'need'
Participant-external	Deontic	POS	Permission	***kěyǐ*** 可以, ***néng*** 能 **'can'**, ***yǔnxǔ*** 允许 'allow'
			Permission denial	*bùdé* 不得 'not permit'
		NEC	Duty, obligation	*yīnggāi* 应该 **'should'**, ***yào*** 要 'must'
			Prohibition	*bù yīnggāi* 不应该 **'should not'**, ***búyào*** 不要 'must not'
	Anankastic (Circumstantial, Goal-oriented)	POS	Feasibility	*néng* 能， ***kěyǐ*** 可以, ***nénggòu*** 能够 'can'
			Impossibility	*bù néng* 不能 **'cannot'**
		NEC	Only possibility	*bùdébù* 不得不 **'cannot but'**, ***zhǐhǎo*** 只要 **'can only'**, ***bìxū … cái néng*** 必须 … 才能 **'only if A then B'**
			Practical necessity	*bìxū* 必须, ***děi*** 得, ***yào*** 要 'must, have to'
			Exemption	*búyòng* 不用, ***búbì*** 不必 **'no need'**
Epistemic		POS	Inference	*huì* 会 **'might'**
		NEC		*yīnggāi* 应该 'should'
		POS	Probability	*kěnéng* 可能 'probably'
		NEC	Certainty	*yídìng* 一定 **'certainly'**
			Inevitability	*bìjiāng* 必将 **'inevitably'**
Post modal usage			Futurity	*huì* 会, ***yào*** 要 **'will'**

Table 1. Taxonomy and modal classification.

In this study, we rely on the deontic vs. anankastic contrast (Wright 1963) and postulate that modals pertaining to duty and necessity are distributed within a semantic domain having two poles (Sparvoli 2012): namely, the deontic, which expresses an obligation (ancient Greek *déon*) and is related to a moral duty, as in (1a); and the anankastic[2] (from *anánkē*, literally 'rope, wire'), which indicates a practical necessity, linked to a specific purpose, as in (1b). While the deontic is grounded on the speaker's belief system, the anankastic's normative source (Hsieh 2005) is situation-oriented.

1) a. 我们**应该/应当**谦虚谨慎 。 [deontic]

Wǒmen	***yīnggāi/yīngdāng***	*qiānxū*	*jǐnshèn*
we	**should**	be.modest	be.prudent

'We **should** be modest and prudent.'
(Translated from Alleton 1984: 200) [deontic]

b. 去 火车站 **得/必须** 坐 第六六 路 公共汽车。

Qù	*huǒchēzhàn*	***děi/bìxū***	*zuò*	*dìliùliù*	*lù*	*gònggōngqìchē*
go	train.station	**have.to**	sit	66th	CL	bus

'To get to the station you **have to** take bus 66.'
(Van der Auwera and Plungian 1998: 80, Chinese translation lightly modified from Li 2004: 107) [anankastic]

Anchored in the notion of 'inevitability', the anankastic expresses what 'cannot be done otherwise' and allows for identification of a consistent class of expressions often flattened in the 'deontic' category. In a positive context, deontic and anankastic modals are mostly interchangeable. However, the classification into either one of the two classes is based on the different interaction with negation displayed by the given necessity modal (cf. Sparvoli 2012, based on Lü Shuxiang 1942). Namely, the negation of a prominent deontic marker produces a Prohibition (necessity + negation), like *bù yīnggāi* 不应该 'should not', while the negation of an anankastic one conveys an Exemption

2 Rough equivalents of the anankastic modality are found in the "participant-external non-deontic" (Van der Auwera, Plungian 1998), the "goal-oriented or teleological" and in the "neutral" or "circumstantial dynamic" modality.

(negation + necessity), like *búyòng* 不用, *búbì* 不必 'no need to'. In the positive form, the modal *yào* 要 typically refers to practical necessity (anankastic), but can also refer to a duty or an obligation (deontic); in the negative form, it can only express a deontic prohibition (similar to the English 'must').

Chinese makes use of a wide range of anankastic items expressing different binding force. The strongest type is conveyed via modal suppletion, that is, via a possibility marker in double negation or combined with *zhǐ* 只 'only', here spelled as Only Possibility. The most common anankastic modal is *bìxū* 必须 'necessarily', conveying Practical necessity. Notwithstanding its status as an adverb, *bìxū'* is considered the equivalent of the English 'must' (Li 2004) and has the same force and scopal properties as semimodals like 'have to, it takes' (Sparvoli 2012).

In the participant-external possibility domain, the deontic/anankastic contrast appears as either Permission or Feasibility, the latter being a possibility stemming from specific circumstances. Finally, the participant-internal possibility is partitioned into two classes: Ability (or capacity) intrinsic to the participant (typically the subject) and Success, that is, the capacity to actualize the state of affairs notwithstanding a challenging situation or task. In other words, the latter is a possibility "determined by a combination of the inherent properties of the subject and of external factors" (Coates 1983: 93).

2.2 The classification of Chinese adverbs

While seminal studies exist addressing Chinese modal adverbs, such as that of Ernst (2009), the literature on this category as such is less developed. Zhang (2014) is one of the few monographs entirely devoted to Chinese adverbs. According to the author, the semantic and syntactic features of these Chinese words make them a particular lexical class, distinguished from the others both for the internal variety and for the quantity of these items, which greatly outnumber other functional words such as prepositions and conjunctions. The long-standing difficulty in providing a comprehensive definition of Chinese adverbs and the different criteria proposed to classify them, has caused a theoretical standstill on these words, to the point that "even for basic issues such as nature, scope and classification of adverbs it is still difficult to reach a shared view" (Zhang 2014: 3). The three macro categories proposed are those of descriptive, restrictive and evaluative adverbs.

Macrocategory	**Subcategory**	**Items**
1. Descriptive	Manner *fāngshì* 方式	*ruìyì* 锐意 'dauntless' (*zhēnxīn* 真心 'sincerely')[3]
	State *zhuàngtài* 状态	*zhúbù* 逐步 'step by step' (*dāngxià* 当下 'immediately')
	Modality *qíngkuàng* 情状	*kǔkǔ* 苦苦 'persistently' (*yìrán* 毅然 'resolutely')
	Comparison *bǐkuàng* 比况	(*fēisù*飞速 'rapidly', *pīnmìng* 拼命 'at all cost')
2. Evaluative		*bìxū* 必须 'necessarily', *yídìng* 一定 'certainly'
3. Restrictive	Linking *guānlián* 关联	*jiù* 就 'right', *yě* 也'also'
	Negation *fǒudìng* 否定	*bù* 不 'not', *méiyǒu* 没有 'not have'
	Time *shíjiān* 时间	*yǒngyuǎn* 永远 'always', *yǐjīng* 已经 'already'
	Frequency *pínlǜ* 频率	*lúnfān* 轮番 'in turn', *bùduàn* 不断 'continuously'
	Repetition *chóngfù* 重复	*yòu* 又 'again', *yě* 也 'also'
	Degree *chéngdù* 程度	*zuì* 最 'most', *hěn* 很 'very'
	Scope *fànwéi* 范围	*dōu* 都 'all', *zhǐ* 只 'only'
	Expressing collaboration *xiétóng* 协同	*yídào* 一道 'together', *yìqǐ* 一起 'together'

Table 2. Classification of adverbs (Zhang 2014).

[3] The adverbs in brackets are those not included in the text analyzed in this work.

Those of the first type are used as verbal modifiers, their position is fixed and is in the middle of the sentence. They mainly describe the action. The bulk of Chinese adverbs is represented by the second category: they can be used in different positions and are used to restrict and differentiate the state and nature of the action. Restrictive adverbs are further divided into 8 categories: connectives, negation adverbs, time adverbs, frequency adverbs, repetition adverbs, degree adverbs, scope adverbs, and adverbs expressing collaboration.

Evaluative adverbs can have a predicative function, the position is rather flexible and they are mainly used to express the speaker's attitude and evaluation towards the action. Since this category tends to overlap with modal adverbs, we treat them separately, based on the taxonomy outlined in Section 2.1.

Although research has mostly focused on English, the case of Chinese has also been taken into account. In particular, the function of adverbs in expressing speaker's stance has been investigated in a number of studies. Among them, Hsieh (2005) analyses how the notion of "source of involvement" can be applied to distinguish and classify Chinese modal verbs and adverbs. Hsieh (2008) is instead based on newspaper data and verifies how the category of evidentiality is encoded to express stance. The same author also investigates the usage of epistemic adverbs in different categories of news, i.e., local, business and politics, reporting a higher frequency in the first category (Hsieh 2009).

Another attempt to deal with this topic is represented by Ai (2012), which focuses on stance adverbs across different communicative modes, diverse genres, and various media channels. Compared to other studies carried out from the contrastive analysis perspective, Ai (2012) is based on large-scale corpus data, the Academia Sinica Balanced Corpus of Modern Chinese, and it is uniquely focused on the category of adverbs. Following Biber (2006)'s framework, the items have been divided into four main categories: certainty, likelihood, attitude, and style. Research results show how epistemic adverbs (expressing certainty and likelihood) are more frequent than the others and that the distribution of the different adverbs across genres also reveals variety.

3. Data and analysis

3.1 The target speech

3.1.1 Historical significance

The historic relevance of the speech is visible from both the grandiosity of the commemorative setting, which strikingly echoes the mass gatherings of the Maoist era, and the international media reaction. Delivered from the same location of Mao Zedong's historical speech of 1949, in front of an audience comprising 70,000 party members and accompanied by a majestic military parade, the commemoration marked the zenith of modern national celebration. Described by a Chinese internal analyst as "a tough talk" delivered in "response to U.S. and western efforts to 'contain' and 'suppress' China" (cf. Tian and Woo 2021), the speech "articulated the full agenda for China and the path the country will take in the next 100 years" (Kimura 2021). This commemorative address is mentioned in the Chinese media using the expression *jiŭ ge bìxū* 九个必须 'the nine-musts'. This label refers to the nine requirements representing the milestones in the roadmap set by Xi Jinping towards the realization of national rejuvenation. Not surprisingly, the invitation to study, implement and "firmly grasp" (*láo láo bǎwò* 牢牢把握) the nine-musts has been pointed out by the *Zhōngguó Jiàoyù Bào* 中国教育报 (*Chinese Journal of Education*) and disseminated by the Chinese Ministry of Education (CMOE 2021).

3.1.2 Synopsis

Before analysing the usage of modal words in the text, it seems useful to briefly provide a synopsis of the speech. After the initial greetings, the speaker retraces Chinese history quoting the main actors of the key period between the second half of the 19th century and the first half of the 20th century and describes the role of the Chinese Communist Party (CCP) in the realization of national rejuvenation. The speech then mentions the changes brought about since the Party's 18th National Congress, with the election of the new Central Committee of the CCP and continues by reviewing the achievements attained over the past hundred years of struggle and quoting the contributions of past leaders to the cause of China's revolution. After a passage expressing gratitude, the second half of the speech changes perspective, from the past to the future, and clarifies what has to be done in nine parts, each introduced by the evaluative adverb expressing anankastic necessity, *bìxū*. The conclusion of the speech expresses belief in the ability to fulfil the Chinese dream and the standard augural greetings to the Chinese Party and people. The speech displays many of the features of the language labelled as *post-Mao official Chinese* in terms of lexicon and metaphor, grammar, rhythm, moral weight, goal orientation and status of truth (Link 2013).

3.1.3 The rhetoric structure

The text under scrutiny is a speech[4] lasting one hour at a very solemn pace. It consists of 7,260 Chinese characters, equivalent, in the official translation, to 5,200 English words and belongs to the genre of commemorative speeches. The latter are consent-oriented texts typically uttered by high-profile political representatives and are aimed at the formation of public attitudes, opinions and will (Reisigl 2008: 253). Its content is analysed here through the lens of the typical rhetorical macro-structures, namely: (1) introduction, (2) narrative, (3) prescriptive argumentation, and finally (4) conclusion (*ibid.*). Adopting Longacre's (1974) division of genres (Table 3), each rhetoric unit can be further described using the notion of prescriptivity, a feature found only in the argumentative part.

Non-prescriptive	**Prescriptive argumentation**	
Narrative	Hortatory	Procedural
First or third person	Second person	Non-specific person
Actor-oriented	Addressee oriented	Goal oriented
Accomplished time (encoded as past or present)	Commands, suggestions (encoded as imperatives or "soft" commands)	Projected time (encoded as past, present or future)
Chronological linkage	Logical linkage	Chronological linkage

Table 3. Longacre's genres (slightly modified from Van Leeuwen 2008: 347).

[4] For the definition of speech, we rely on Reisigl (2008: 243): "A speech is a structured verbal chain of coherent speech acts uttered on a special social occasion for a specific purpose by a single person, and addressed to a more or less specific audience".

Moreover, the prescriptive argumentation can have either a procedural or a hortatory orientation. The former aims at highlighting the means to reach a given goal (know-how-to-do-it), the latter aims to "conduct influence" to get people "to feel or think or do in certain ways" (Van Leeuwen 2008: 347). Table 4 shows our subdivision of the texts into six parts: introduction, narrative, transition to argumentation, prescriptive argumentation (procedural and hortatory), and conclusion.

Rhetoric unit		**Char. No.**	**%**
Introduction		290	3.99%
Narrative		2679	36.90%
Transition		120	1.65%
Argumentation	Procedural	3228	44.46%
	Hortatory	799	11.01%
Conclusion		144	1.98%
Total		7260	100%

Table 4. Sequence and length of the rhetoric units in the speech.

The rhetorical structure of the text is organized in a very consistent and logical sequence. After the introduction, the speech content is developed in a sequence of nine pragmatic units, as shown in (2). It mentions the goal of building a strong modern and socialist country, the historical need of rescuing the county after the experience of semi-colonialism, the role of socialism as the only possibility for redemption, the roadmap to success based on the nine-musts, future commitments, the exhortation to all Chinese people across the Taiwan Strait, the injunction to external forces, another exhortation to the youth, and finally, the promise that the Great Chinese Dreams can certainly be realized.

2) Goal > Need > Only possibility > Roadmap > Commitment > Exhortation > Injunction > Exhortation > Promise

3.2 Lexical features

As for the lexicon, whereas we do not find many occurrences of those abstract nouns (such as *xíngshì* 形势 'circumstance' and *júmiàn* 局面 'situation') and verbs (such as *jiāyǐ* 加以 'apply' and *jìnxíng* 进行 'conduct') that are typical of bureaucratic language, we find many repetitions of some key lexical words and phrases, such as *Zhōngguó* 中国 'China', *rénmín* 人民 'people', *wěidà* 伟大 'great', *Gòngchǎndǎng* 共产党 'Communist Party' and *Zhōnghuá Mínzú* 中华民族 'the Chinese Nation', i.e., the top five lexical words. To give an example, the nominal modifier *wěidà*, with 53 occurrences, can be found throughout the speech since it mostly precedes the noun *fùxīng* 复兴 'rejuvenation' which is one of the topics frequently mentioned by President Xi since the speech given in November 2012 presenting the Chinese dream. Repetition of words and phrases is aimed at fulfilling different functions: it provides rhythm; it is used as a device to connect fragments of discourse; and it also reinforces the truthfulness and salience of what is repeated. The most frequently repeated phrases are listed in Table 5. To quote Link, it is this repetition that "in combination with rhythm, can enhance the implicit sense that 'this is right, there is no need to think about it'" (2013: 264).

Repeated phrases	**No.**
同志们、朋友们! *Tóngzhìmen, péngyǒumen* "Comrades and friends,"	8
为了实现中华民族伟大复兴 *Wèile shíxiàn Zhōnghuá mínzú wěidà fùxīng* "In order to realize the great rejuvenation of the Chinese nation"	4
中国共产党和中国人民以英勇顽强的奋斗 *Zhōngguó gòngchǎndǎng hé Zhōngguó rénmín yǐ yīngyǒng wánqiáng de fèndòu* "Through tenacious struggle, the Party and the Chinese people…"	4
新的征程上，我们必须/要 *Xīn de zhēngchéng shàng, wǒmen bìxū/yào* "In the journey ahead, we have to/must…"	8+1
以史为鉴、开创未来，必须 *Yǐshǐwèijiàn, kāichuàng wèilái, bìxū* "As we learn from the mirror of history to create our future, it is necessary to…"	9

Table 5. Repeated phrases in the speech.

Concerning the modal distribution, the text includes 46 modal markers. As shown in Table 6, their distribution reveals the predominance of participant-external modalities, and among them a striking majority of goal-oriented (56.52%), over morally connotated necessity

(8.70%), that is, a remarkable predominance of anankastic vs. deontic modals.

Modality	**No.**	**%**	**Chinese**	**Modal meaning**	**No.**
Participant-internal Necessity	2	4.35%	*xūyào* 须要 need	Need	2
Participant-internal Possibility	3	6.52%	*shànyú* 善于 be good at	Ability	2
			nénggòu 能够 manage, succeed	Success	1
Participant-external Possibility	3	6.52%	*nénggòu* 能够 can	Feasibility	2
			kěyǐ 可以 can, be sufficient		1
Participant-external Anankastic Necessity	26	56.52%	*zhǐyǒu …cái néng* 只有…才能 only if A then B	Only possibility	2
			bìxū …cái néng 必须…才能 only if A then B		1
			bìxū 必须 must (have to)	Practical necessity	20
			yào 要 must	Necessity	3
Participant-external Deontic necessity	4	8.70%	*yào* 要 should	Moral duty	2
			búyào 不要 must not	Prohibition	1
			bùyǔnxǔ 不允许 do not allow	Permission denial	1
Epistemic	4	8.70%	*bìjiāng* 必将 inevitably	Inevitability	2
			yídìng 一定 certainly	Certainty	2
Futurity (Post modal usage)	4	8.70%	*yào* 要 should	Commitment	2
			búhuì 不会 will not	Negative futurity	2
	46	100%			46

Table 6. Modal distribution by modal categories.

Given the chameleonic nature of *yào* (expressing both deontic and anankastic necessity, futurity and participant-internal Need) we will comment each occurrence in the speech (and the relevant modal classification) in Sections *4.2.3*, *4.2.4*. Other modalities (participant-internal and participant-external possibility) are also represented, while expressions of epistemic certainty, futurity, and inevitability are scattered throughout the speech. To identify the relevance and usage of each modality, we need to observe their distribution in each rhetoric unit. However, before doing this, we also need to single out the features of the evaluative adverbs disseminated in the speech, often co-occurring with the modals. We will address this task in the next section.

4. Discussion

4.1 Speaker's stance markers: evaluative adverbs

4.1.1 Distribution and features

An initial look at the raw occurrences suggests that the most frequently used lexical words are nouns, verbs, adjectives and adverbs. As for the last two lexical classes, we find more adjectival (296) than adverbial tokens and more adverbial (97) than adjectival (81) lexical types, suggesting that although used less frequently, the internal variety of adverbs is higher than that of adjectives.

Lexical word tokens	
Lexical class	# tokens
Noun	1585
Verb	839
Adjective	296
Adverb	236
Lexical word types	
Lexical class	# types
Noun	495
Verb	379
Adverb	97
Adjective	81

Table 7. Lexical distribution by grammar category.

As for the adverbs, we can see that only a small portion of them belong to those traditionally included in the category of modal adverbs. The speaker rarely expresses obligatoriness, doubt or certainty towards the sentence through modal adverbs. What we more frequently find is a set of different adverbs usually classified as degree, manner, time or scope which nonetheless contribute to express the speaker's engagement towards the statement. To have an idea of the "weight" of this lexical class within the corpus, rank and frequency of the top 10 adverbs are given in Table 8.

The first column indicates the rank of the item within the category of adverbs, the second shows the rank within the entire corpus and the third the number of occurrences: the values for the negation adverb *bù* 不 for instance indicate that this item is the most frequently used within the adverbs and ranks 28 within all words of the corpus occurring 26 times. As we can see, among the top 100 words of the corpus, six are adverbs.

Rank 1	Rank 2	Frequency	Item
1	28	26	*bù* 不
2	31	20	*bìxū* 必须
3	34	14	*quánmiàn* 全面
4	64	10	*bùduàn* 不断
5	66	10	*shǐzhōng* 始终
6	77	8	*jiù* 就
7	94	6	*shēnkè* 深刻
8	105	6	*dōu* 都
9	111	6	*zuì* 最
10	114	6	*gèngwéi* 更为

Table 8. Lexical ranking by frequency.

As mentioned previously, using the function wordlist we obtained the list of 97 different adverbs. Nevertheless, although these words are used in adverbial function, they do not belong to the adverbial lexical class, being either an adjective, verb, noun or even phrase used as verbal modifier and labelled by the query tool as *adverb*. For this reason, we selected only those items that are labelled as adverbs in the most recent edition of the *Xiàndài hànyǔ cídiǎn* 现代汉语词典 and the number dropped from 97 to 43. We then classified them according to the three macrocategories proposed by Zhang to have a quantitative picture of these items.

For those adverbs belonging to more than one category, they were classified according to the function they play in each occurrence. One such case is *jiù* 就, which, according to Zhang, can be labelled as a time, scope, evaluative or connective adverb.

Among the adverbs classified as evaluative by Zhang are: *jué* 绝, *juéduì* 绝对, *yě* 也, *cái* 才, *jiù*, *dōu* 都, *qià* 恰, *yídìng* 一定, *bìxū*, and *gēnběn* 根本. Moreover, *qià* 恰 is included only in the reduplicated form but we will consider it within the other items. The analysis of occurrences will start with those items uniquely labelled as evaluative.

Another adverb, not included in Zhang's classification, is composed of *bì* 必and *jiāng* 将, and we believe that it can also be interpreted as evaluative, as they express a strong epistemic nuance.

4.1.2 Bìxū and the nine-musts

The evaluative adverb most frequently used is *bìxū*, which expresses necessity and could be rendered as "necessarily", although in the English version of the speech it is translated by the verb "must". In six out of 20 occurrences, *bìxū* modifies the verb *jiānchí* 坚持 to redundantly express the mandatory feature of "persist in" upholding the firm leadership of the Party, the overall leadership of the Party, Marxism-Leninism, socialism with Chinese characteristics, the Party's basic theory, great unity and solidarity. In the other occurrences, we find either verbs expressing the meaning of implementing something (usually a theory), such as *guànchè* 贯彻, *tuījìn* 推进 and *jìnxíng* or verbs suggesting the meaning of adding force or speed, such as *jiākuài* 加快, *zēngqiáng* 曾强, and *jiāqiáng* 加强. The necessity expressed by *bìxū* is linked to actions whose features are conveyed by other adverbs which, despite not directly expressing necessity, nonetheless provide the meaning of doing something strictly (*jǐnjǐn* 紧紧), in a comprehensive way (*quánmiàn* 全面) and continuously (*bùduàn* 不断). Among the adverbs most frequently used in this text we also find *gèngwéi* 更为 which, like *gèng* 更, intensifies the degree of a quality or state. We find six occurrences of this item, which is always used to modify an adjective, and twice after the sentence introduced by *bìxū* before the two synonyms adjectives *jiānjù* 艰巨 and *jiānkǔ* 艰苦 'arduous'.

3) (…) 同时必须准备付出更为艰巨、更为艰苦的努力。

Tóngshí	*bìxū*	*zhǔnbèi*	*fùchū*	***gèngwéi***	*jiānjù,*
simultaneously	necessarily	prepare	invest	**even.more**	arduous
gèngwéi	*jiānkǔ*	*de*	*nǔlì*		
even.more	hard	DE	effort		

'At the same time it is necessary to prepare to invest **even more** arduous, **even greater** effort'[5]

The rendering of *bìxū* as the English "must" is justified by the lack of other verbs in the same utterance as in (4).

4) 强国必须强军

Qiáng	*guó*	***bìxū***	*qiáng*	*jūn*
powerful	country	**necessarily**	strong	army

'A strong country **must** have a strong army'

Moreover, *bìxū* is used nine times impersonally with the meaning of "it is necessary to" preceded by the motto *yǐ shǐ wèi jiàn, kāichuàng wèilái* 以史为鉴、开创未来 "put conscious effort into learning from history to create a bright future" and nine times with the first-person plural pronoun preceded by the phrase *xīn de zhēngchéng shàng* 新的征程上 "On the journey ahead" (see Table 5). In this case the last occurrence can also be preceded by another modal, *yào*, used to underline the necessity to remember the argument at issue[6].

The two *incipit* in the alternate pattern repeat nine times as follows:

[5] When there is no reference to the official translation of the speech, namely, Xi (2021a), the translation in English is by the authors.

[6] The replacement of *bìxū* with *yào* might be due to the fact that in this segment another motto is quoted, cited from the Maoist repertoire, and including the modal *bìxū*: "*Xīn de zhēngchéng shàng, wǒmen yào láojì dǎtiě bìxū zìshēn yìng de dàolǐ* 新的征程上，我们要牢记打铁必须自身硬的道理. On the journey ahead, we **must** keep firmly in mind the old adage that **it takes** a good blacksmith to make good steel" (Xi 2021a).

5) a. 以史为鉴、开创未来，必须 (…)

Yǐ shǐ	*wèi*	*jiàn,*	*kāichuàng*	*wèilái,*	***bìxū***
use history as		mirror	create	future	**necessarily**

‘As we learn from the mirror of history to create our future, **it is necessary to**…’

b. 新的征程上，我们必须/要 (…)

Xīn	*de*	*zhēngchéng*	*shàng,*	*wǒmen*	***bìxū/yào***
new	DE	journey	on	we	**necessarily/must**

‘In the journey ahead, we **have to/must**…’

4.1.3 Adverbs conveying certainty

Among the items contributing to conveying the certainty of the utterance, we mention the negative polarity item *jué* ‘absolutely’ and the epistemic adverb *yídìng* ‘certainly’. We find five occurrences of *jué*: in three cases it modifies the negation *bù* intensifying the degree, in two cases it is instead used in conjunction with *duì* 对 but as a nominal modifier, and therefore will not be taken into account. This is also the case of *gēnběn*, always used as a nominal modifier.

As for the adverb *yídìng*, we only find two occurrences at the end of the speech to provide certainty regarding the possibility of achieving the goal specified (6a) and realizing the Chinese dream (6b).

6) a. 全面建成社会主义现代化强国的目标**一定能够**实现

Quánmiàn		*jiànchéng*	*shèhuìzhǔyì*	
in.all.respects		build	socialist	
xiàndàihuà		*qiáng*	*guó* *de*	*mùbiāo*
modernize		strong	country DE	goal
yídìng	***nénggòu***	*shíxiàn*		
certainly	**can**	realize		

‘The goal of building a strong modern and socialist country **can certainly be** realized’

b. 中华民族复兴的中国梦**一定能够**实现！

Zhōnghuá	*mínzú*	*fùxīng*	*de*	*Zhōngguó*
Chinese	nation	rejuvenation	DE	Chinese
mèng	***yídìng***		***nénggòu***	*shíxiàn*
dream	**certainly**		**can**	realize

'The Chinese dream of the rejuvenation of the Chinese nation **can certainly be** realized!'

Another adverb labelled as evaluative is *qià*, which occurs only once at the end of the speech and expresses is the timeliness of the appearance or existence of something, in this case the centennial of the CCP (7).

7) 百年**恰**是风华正茂

Bǎi	*nián*	***qià***	*shì*	*fēng*	*huá*
hundred	year	**exactly**	is	wind	splendid
zhèng	*mào*				
just	luxuriant				

'One hundred years **exactly** is the best time'

4.1.4 Other adverbs

As for the adverbs displaying more than one function, *jiù*, which occurs nine times in the corpus, can be considered as belonging to the evaluative category five times. In each occurrence the adverb signals that what is expressed is deemed to be evident and obvious as in (8).

8) 江山**就**是人民、人民**就**是江山

Jiāngshān	***jiù***	*shì*	*rénmín*	*rénmín*	***jiù***	*shì*	*jiāngshān*
country	**just**	is	people,	people	**just**	is	country

'This country is its people; the people are the country'

Another adverb displaying more than one function is *dōu*, but according to the analysis of its six occurrences, it is always used as a scope adverb. This is the case also when this adverb is used to signal that *all* the attempts[7] to save the nation have failed in contrast with the fight conducted by the Party, the only one which was successful (9).

9) (…) **都**以失败而告终

Dōu	*yǐ*	*shībài*	*ér*	*gàozhōng*
all	be	failure	for	reach.an.end

'(…) but **all** of these ended in failure.'

The same can be said about the four occurrences of *cái*, where the adverb is used always before *néng* 能 to express the conditions under which something can occur, both in the construction with *zhǐyǒu* 只有 and without, as in (10a) and (10b).

10) a. 弄明白未来我们怎样**才能**继续成功 (…).

Nòng	*míngbái*	*wèilái*	*wǒmén*	*zěnyàng*	***cái***
make	clear	future	we	how	**only**
néng	*jìxù*	*chénggōng* (…).			
can	continue	successful			

'We can understand how **we can** continue to succeed in the future.'

b. **只有**社会主义**才能**救中国，只有社会主义**才能**发展中国！

Zhǐyǒu	*shèhuìzhǔyì*	***cái***	***néng***	*jiù*	*Zhōngguó,*
only	socialism	**only**	**can**	save	China,
zhǐyǒu	*shèhuìzhǔyì*	*cái*	*néng*		
only	socialism	only	can		

[7] The preceding passage reads as follows: "*Wèile zhěngjiù mínzú wéiwáng, Zhōngguó rénmín fènqǐ fǎnkàng, rénrénzhìshì bēnzǒu nàhǎn, tàipíngtiānguó yùndòng, wùxū biànfǎ, Yìhétuán yùndòng, Xīnhài gémìng jiēlián ér qǐ, gè zhǒng jiùguó fāng'àn lúnfān chūtái, dàn dōu yǐ shībài ér gàozhōng* 为了拯救民族危亡，中国人民奋起反抗，仁人志士奔走呐喊，太平天国运动、戊戌变法、义和团运动、辛亥革命接连而起，各种救国方案轮番出台，但**都**以失败而告终。To save the nation from peril, the Chinese people put up a courageous fight. As noble-minded patriots sought to pull the nation together, the Taiping Heavenly Kingdom Movement, the Reform Movement of 1898, the Yihetuan Movement, and the Revolution of 1911 rose one after the other, and a variety of plans were devised to ensure national survival" (Xi 2021a).

fāzhǎn	*Zhōngguó!*
develop	China

'**Only** socialism **could** save China, and only socialism could develop China.'

The last adverb displaying more than one function is *yě* 也. As we mentioned in 3.2, repetition is one of the rhetorical devices used in official language to reinforce the speaker's illocutionary force.
It is therefore difficult to separate the function of a restrictive adverb from that of an evaluative adverb, expressing the speaker's stance.
In the five occurrences of *yě* in the text, the repeated item is introduced by *yě* in the second occurrence, either in the construction *bùdàn* 不但…*yě* 也…(11a) or to connect separate sentences which are nonetheless semantically linked (11b).

11) a. 中国人民不但**善于**破坏一个旧世界、也**善于**建设一个新世界 (…).

Zhōngguó	*rénmín*	*bùdàn*	***shànyú***	*pòhuài*
China	people	not only	**be.good.at**	to destroy
yī	*gè*	*jiù*		
one	CL	cold		
shìjiè	*yě*	***shànyú***	*jiànshè*	*yī*
world	also	**be.good.at**	to build	one
gè	*xīn*	*shìjiè* (…).		
CL	new	world		

'the Chinese people **were capable** of not only dismantling the old world, but also building a new one.' (Xi 2021a)

b. 中国人民从来没有欺负、压迫、奴役过其他国家人民 (…)。同时，中国人民**也绝不允**许任何外来势力欺负 (…).

Zhōngguó	*rénmín*	*cónglái*	*méiyǒu*	*qīfù*	*yāpò*	*núyì*
China	people	ever	not	bully	oppress	enslave
guò	*qítā*	*guójiā*	*rénmín*			
GUO	other	country	people			
(…). *Tóngshí,*		*Zhōngguó*		*rénmín*	***yě***	***juébù***

at.the.same.time		China	people	**also absolutely.not**	
yǔnxǔ	*rènhé*	*wàilái*		*shìlì*	*qīfù (…).*
allow	any	external		force	bully

‘We have never bullied, oppressed, or subjugated the people of any other country (…). By the same token, **we will never** allow any foreign force to bully, oppress, or subjugate us.’ (Xi 2021a)

4.2 Modal distribution and rhetoric structure

The distribution of modals in each rhetoric unit of the speech (Table 9) appears to be consistent and uniform. For instance, modals pertaining to practical necessity and moral duty uniformly instantiate in two separate parts of the argumentation: modals expressing practical and goal-oriented necessity are used in the medial part of the speech (Procedural argumentation), while deontic modals occur in the section preceding the conclusion (Hortatory argumentation).

While the semantics of necessity covers the argumentative part of the speech, the semantics of possibility is mainly distributed in the opening narrative and in the conclusion. Participant-internal necessity (need), epistemic certainty and only possibility are encountered much more sporadically, however their inclusion is crucial for the topic introduction in the narrative, the transition to the argumentation and the conclusion.

4.2.1 The semantics of possibility and inevitability

The expressions of possibility, though not numerous, convey a strong message in the opening and closing of the speech. In the narrative section, two occurrences of participant-internal ability (‘be good at’ 善于 *shàn yú*), having the CCP and the Chinese people as agents, are emphasized by the correlation of *bùdàn* ‘not only’ and *yě* ‘but’, as shown in (11a) in Section *4.1.4*. Immediately after, the mention of “only possibility” refers to the role of socialism in saving and developing the country (*zhǐyǒu shèhuìzhǔyì cái néng* 只有社会主义才能 ‘only socialism can …’).

Modality and domain		Modal item	Semantic Notion	No.	Collocation
					Introduction
Part. Int.	Necessity	*xūyào* 须要 need	Need	2	Narrative
	Possibility	*shànyú* 善于 be good at	Ability	2	
Part. Ext.	Necessity [Anankastic]	*zhǐyǒu …cái néng* 只有…才能 only if A then B	Only Possibility	2	
Epistemic	Necessity	*bìjiāng* 必将 inevitably	Inevitability	1	
Part. Ext.	Possibility	*kěyǐ* 可以 can, be sufficient	Feasibility	1	Transition to Argumentation
	Necessity [Anankastic]	*yào* 要 should	Practical Necessity	1	
Part. Int.	Possibility	*nénggòu* 能够 can, be sufficient	Success	1	
Part. Ext.	Necessity [Anankastic]	*…cái néng …*才能 only if A… then B	Only Possibility	1	
		bìxū 必须 must (have to)	Practical Necessity	20	Prescriptive Argumentation (Procedural)
Futurity		*búhuì* 不会 will not	Negative Futurity	2	
Part. Ext.	Possibility [Deontic]	*bùyǔnxǔ* 不允许 do not allow	Permission Denial	1	
Epistemic	Necessity	*bìjiāng* 必将 inevitably	Inevitability	1	
Part. Ext.	Necessity [Anankastic]	*yào* 要 should	Practical Necessity	2	
Futurity		*yào* 要 will	Commitment	2	Prescriptive Argumentation (Hortatory)
Part. Ext.	Necessity [Deontic]	bú*yào* 不要 must not	Prohibition	1	
		yào 要 should	Moral duty	2	
Epistemic	Necessity	*yídìng* 一定 certainly	Certainty	2	Conclusion
Part. Ext.	Possibility	*nénggòu* 能够 can, be sufficient	Feasibility	2	
			Total	46	

Table 9. Modal distribution by occurrence in the text.

On the other hand, the two modalized propositions at the end of the speech (*yídìng nénggòu shíxiàn* 一定能够实现 '… certainly can be achieved') have a non-agentive reading and thus are here classified in the participant-external (feasibility) category. As shown in (6a-b), they are marked by an epistemic adverb and have non animate subjects (the Chinese dream, the building of socialism) getting a passive reading. Their participant-external nuance is also shown by two existential sentences providing the necessary conditions for the achievement of the current goal of China. Such preconditions are, namely, the presence of 'strong leadership' (*jiānqiáng lǐngdǎo* 坚强领导) by the CCP and 'close unity (*jǐnmì tuánjié* 紧密团结) by the nation. In other words, in comparison to participant-internal slogans such as 'Yes, we can', the catchphrases at the end of the speech draw attention to the conditions that make the goal feasible. The optimistic push to the statement (needed for a well-formulated slogan) is given essentially by the stress on the epistemic certainty of achievement.

The narrative subdivision also includes the first modal expressing epistemic inevitability, conveyed by the evaluative adverb *bìjiāng* 必将 'inevitably' (see Section *4.1.1*) and in reference to the path opened by the CCP in the last one hundred years.

12) **必将**载入中华民族发展史册、人类文明发展史册！

Bìjiāng	*zàirù*	*Zhōnghuá mínzú*	*fāzhǎn*	*shǐcè,*
inevitably	record	Chinese nation	development	annals
rénlèi	*wénmíng*	*fāzhǎn*	*shǐcè!*	
mankind	civilization	development	annals	

'(…) will be **inevitably** recorded in the annals of the development of the Chinese nation and in the annals of the development of human civilization.'

In the procedural argumentation, at the end of the sixth must, there is a phrase that suggests the inevitability of failure, expressed as "any attempt to divide the Party from the Chinese people is bound to fail". This negative futurity expression is reinforced by the evaluative adverbs *dōu* and *jué* (*bù*) (see Section *4.1.3*).

13) 任何想把中国共产党同中国人民分割开来、对立起来的企图，都是**绝不会**得逞的！
Rènhé xiǎng bǎ Zhōngguó gòngchǎndǎng
any think BA China CCP
tóng Zhōngguó rénmín fēngē kāi lái,
with China people divide depart come
duìlì qǐlái de qìtú, dōu shì
oppose QILAI DE attempt all SHI
jué bù huì *déchěng de!*
absolutely not will prevail DE
'Any attempt to divide the Party from the Chinese people or to set the people against the Party **will by no means** prevail.'

The reference to inevitability also occurs in a set of noun modifiers distributed in all the main rhetoric units:

14) a. **不可磨灭**的贡献 [Narrative]
Bùkě mómiè *de gòngxiàn*
not.possible obliterate DE contribution
'**Indelible** contributions'
[of China towards world civilization]
b. **不可逆转**的历史进程 [Procedural Argumentation]
Bùkě nìzhuǎn *de lìshǐ jìnchéng*
not.possible reverse DE history course
'**Inevitable** historical course'
[referred to China's national rejuvenation]
c. **不可战胜**的强大精神力量
Bùkě zhànshèng *de qiángdà jīngshén lìliàng*
not.possible win DE formidable spirit strength
'**Invincible** strength of the formidable spirit'
[referring to the CCP]
d. **不可阻挡**的步伐迈向伟大复兴
[Hortatory Argumentation]
Bùkě zǔdǎng *de bùfá màixiàng*
not.possible stop DE pace march.towards

wěidà	*fùxīng*
great	rejuvenation

'**Unstoppable** pace marching toward the great rejuvenation'

Other possibility markers in subordinated position are found in the last of the nine-musts ("We must continue to advance the great new project of Party building") with reference to the achievement of the one hundred years of struggle by the CCP, who succeeded to:

15) 团结一切**可以团结**的力量、调动一切可以调动的积极因素

Tuánjié	*yīqiè*	***kěyǐ***	*tuánjié*	*de*	*lìliàng,*	*diàodòng*
unite	all	**can**	unite	DE	trength	mobilize
yīqiè	*kěyǐ*	*diàodòng*		*de*	*jījí*	*yīnsù*
all	can	mobilize		DE	positive	factor

'United all the forces **that can be** united, mobilized all positive factors **that can be** mobilized'

4.2.2 Semantic of only possibility in the transition to procedural argumentation

We need now to focus briefly on the modal distribution in the transition between the starting narrative and the prodedural argumentation including the nine-musts. While the distribution of the modal in each stage is relatively uniform, the passage leading from the narrative to the roadmap is the one having the highest number of modal items. As seen in the following excerpt, a full four modals are included in 120 characters, that refer to anankastic possibility (*kěyǐ* 可以), anankastic necessity (*yào*), participant-internal ability (*nénggòu* 能够) and only possibility (*cái néng* 才能). This sequence (*feasibility* > *necessity* > *success* > *only possibility*) strongly foreshadows the procedural argumentation based on the nine-musts.

> *Yǐ shǐ wèi jiàn, kěyǐ zhī xīngtì. Wǒmen yào yòng lìshǐ yìngzhào xiànshí, yuǎn guān wèilái, cóng Zhōngguó gòngchǎndǎng de bǎinián fèndòu zhòng kàn qīngchǔ guòqù wǒmen wèishéme nénggòu chénggōng, nòng míngbái wèilái wǒmen zěnyàng cáinéng jìxù chénggōng*
> 以史为鉴，**可以**知兴替。我们**要**用历史映照现

实、远观未来，从中国共产党的百年奋斗中看清楚过去我们为什么**能够**成功、弄明白未来我们怎样**才能**继续成功

'Rise and fall **can be understood** by learning from the mirror of history. We **must** use the mirror of history to find where we currently stand and gain foresight into the future. From the Party's 100-year struggle, we see why we **were able** to succeed in the past, and we figure out the **only possible** way to succeed in the future.'

4.2.3 Commitment expressions opening the hortatory argumentation

The roadmap presenting the practical requirements to success – the nine-musts discussed in Section *4.1.2* – is followed by the hortatory argumentation. The latter is opened with two occurrences of *yào*. Concerning the classification of these tokens, it must be noted that, with first-person subjects, identifying the appropriate reading of *yào* (from among obligation, futurity, promise, and intention) can be challenging (Li 2004: 188). This is because in such instances, *yào* expresses a commitment for the future (a meaning that somehow encompasses the deontic and futurity reading). This "commitment reading" is also seen in the main predicate of the modalized proposition including the commissive verbs *guànchè* 'implement', *luòshí* 落实 'carry out', *wéihù* 维护 'safeguard', *bǎochí* 保持 'maintain', in the first occurrence (16a), and the commissive verb *jiānchí* 'persist', in the second occurrence, shown in (16b).

16) a. 我们**要**全面准确**贯彻**"一国两制"、"港人治港"、"澳人治澳"、高度自治的方针，

Wǒmen	***yào***	*quánmiàn*	*zhǔnquè*	***guànchè***
we	**will**	totally	accurately	**implement**

"yīguóliǎngzhì",	*"gǎngrén*	*zhì*	*gǎng",*
one.country.two.system	Hongkonger	administer	Hong Kong

"àorén	*zhì*	*ào",*	*gāodù*
Macanese	administer	Macao	high-degree

zìzhì	*de*	*fāngzhēn,*
autonomy	DE	policy

‘We **will** completely and accurately **implement** the policy of One Country, Two Systems, the people of Hong Kong administer Hong Kong, the people of Macao administer Macao, and the policy of a high degree of autonomy.’

b. **要坚持**一个中国原则和“九二共识”，推进祖国和平统一进程。

Yào	***jiānchí***	*yīgè*	*Zhōngguó*	*yuánzé*
will	**uphold**	one-CL	China	principle
hé	*“jiǔ’èr*	*gòngshì”,*	*tuījìn*	*zǔguó*
and	’92	consensus	promote	advance
hépíng	*tǒngyī*	*jìnchéng.*		
peace	unification	course		

‘We **will uphold** the one-China principle and the 1992 Consensus and promote the course of peaceful national reunification.’

4.2.4 Instances of deontic necessity

As highlighted in Section *4.1.2*, the most frequent modal is *bìxū* expressing anankastic necessity, that is, conveying a type of necessity that is anchored to the achievement of a given goal in a given situation. The occurrences having a fully-fledged deontic meaning are instead only four: two are negative injunction addressed to the West, while the other two are directed towards two distinct addressees, namely, “all of us” compatriots living in the two shores of the Taiwan strait (17a), and Chinese youth (17b). In the latter, the deontic moral nuance is strengthened by the noun *jǐrèn* 己任 ‘responsibility’ in the modalized sentence and *yào* is used, conveying a prescription typically uttered by someone in a position of power towards a weaker addressee (cf. Alleton 1984: 286ff), as suggested also in the official English version, rendered as “our young people”. The difference between the deontic and the anankastic occurrences of *yào* is visible in the comparison between (17a-b) and (17c). The former are related to a morally oriented duty, whereas in the latter *yào* denotes a practical requirement anchored to a given goal. Finally, the participant-external reading of (17c) is due to the impersonal construction, similar to English ‘it takes’, or Italian ‘ci vuole, occorre’.

17) a. 包括两岸同胞在内的所有中华儿女，**要**和衷共济、团结向前，

Bāokuò	*liǎng'àn*	*tóngbāo*	*zàinèi*	*de*
comprise	both.shores	compatriot	included	DE
suǒyǒu	*Zhònghuá*	*érnǚ,*		
all	Chinese	sons.and.daughters		
yào	*hézhōnggòngjì,*		***tuánjié***	***xiàngqián***
must	work.together.with.one.heart		**unite**	**forward**

'[All of us], compatriots on both sides of the Taiwan Strait, **must** come together and move forward in unison.' (Xi 2021a)[8]

b. 新时代的中国青年**要**以实现中华民族伟大复兴为己任，

Xīnshídài		*de*	*Zhōngguó*		*qīngnián*
new.era		DE	China		youth
yào		*yǐ*	*shíxiàn*	*Zhōnghuá*	*mínzú*
must		take	realize	China	nation
wěidà	*fùxīng*		*wéi*		*jǐrèn,*
great	rejuvenation		make.as		responsibility

'In the new era, our young people **must** make it their mission to contribute to national rejuvenation.' (Xi 2021a)

c. 实现伟大梦想就**要**顽强拼搏、不懈奋斗。

Shíxiàn	*wěidà*	*mèngxiǎng*	*jiù*	***yào***	*wánqiáng*
realise	great	dream	then	**need**	tenacious
pīnbó,	*bùxiè*	*fèndòu.*			
strive	untired	struggle			

'Realizing our great dream **will require** hard work and persistence.' (Xi 2021a)

[8] For these two examples, we used the official English translation (Xi 2021a) wherein the addressee is explicitly mentioned, as in first-person "all of us" in (17a) and the third-person "our young people", in (17b). Similarly, (17c) is impersonal also in the English official translation.

As we saw in (11b), Section *4.1.4*, a very strong negative deontic expression (*jué bù yǔnxǔ* 绝不允许 'absolutely not allow') is addressed to "external forces" (*wàilái shìlì* 外来势力) who might like to oppress the Chinese people. The second instance of deontic prohibition is addressed to whomever might like to prevent the mission of reunification of China, with reference to Taiwan (18); the binding force of the negative injunction is emphasized by the scope adverb *dōu* 'all'.

18) 任何人都**不要**低估中国人民捍卫国家主权和领土完整的坚强决心、坚定意志、强大能力！
Rènhérén dōu ***bùyào*** *dīgū Zhōngguó rénmín*
any.person all **must.not** underestimate China people
hànwèi guójiā zhǔquán hé lǐngtǔ wánzhěng de jiānqiáng
protect country sovereignty and territory integrity DE strong
juéxīn, jiāndìng yìzhì, qiángdà nénglì!
determination firm will formidable ability
No one must underestimate the strong determination, the firm will, and the formidable ability of the Chinese people to defend national sovereignty and territorial integrity.'

Finally, another implicit injunction is conveyed by *jué bù* (presented in Section *4.1.3*) combined with the non-agentive verb *jiēshòu* 接受 'accept'. Without resorting to a fully-fledged modal, a negative command is expressed here by suppletion strategy. Instead of uttering a negative prescription to the addressee, "don't preach to us", the speaker commits to "not accepting any preaching from the addressee". The case in point is (19).

19) 但我们**绝不**接受"教师爷"般颐指气使的说教
Dànwǒmen ***jué*** ***bù*** *jiēshòu "jiàoshīyé"*
but we **absolutely** **not** accept arrogant.lecturer
bān yízhǐqìshǐ de shuōjiào
manner arrogant.and.bossy DE preaching
'But we **absolutely do not** accept the sermons by arrogant and bossy "lecturers".'

Narrative			Argumentation		Conclusion
			Procedural	Hortatory	
Problem > Goal >			Roadmap >	Commitment > Injunction >	Promise of fulfilment
Need >	Ability >	Only possibility >	Practical necessity >	Futurity > Prohibition >	Certainty > Feasibility
xūyào	*shàn yú*	*zhǐyǒu.. cái néng*	*bìxū*/*yào*	*yào/búhuì> (bú)yào*	*yídìng néngòu*
'need to'	'be good at'	'only A can B'	'must'	'will/will not'> 'should (not)'	certainly can
Actors			Addressees		Actors
China >	Chinese people >	Socialism	Non-specific/1st person >	1st/3rd person >	CCP/ Chinese people of all ethnic groups

Table 10. Modal meaning per rhetoric unit.

In conclusion, the modal distribution per rhetoric unit is given in Table 10, which shows the consistency and logical progression of the speech and also the crucial role played by prescriptive argumentation, in the form of claim for actions aimed at a purpose (procedural) and, to a lesser extent, also in the form of claim addressing the listeners to persuade them to embrace a specific course of action (hortatory).

5. Conclusions

The results of this study can be summarized in the following six observations.

(I) The predominance of the procedural subdivision (in terms of length and modal density) over the hortatory one, hence the abundance of modals of practical necessity (anankastic) and the scarcity of modals resorting to moral necessity as a form of persuasion.

Out of four deontic expressions, two of them are injunctions to the West; moreover, a further negative injunction to foreign forces is conveyed through the negative evaluative adverb *jué bù* 'absolutely not'. The scarcity of deontic modals (and, notably, the absence of the prototypical deontic marker *yīnggāi* 应该 'should') does not imply the lack of a system of values. As highlighted in the numerous references to modern Chinese history, and in the repeated motto about learning from history (*yǐ shǐ wéi jiàn* 以史为鉴) the value system is the one already outlined, from the liberation of 1949 up to the present.

(II) The speech itself is structured as a roadmap to the attainment of the given goal. Thus, the prevalence of goal-oriented items, as *bìxū*, in the nine-musts (anankastic).

As shown in Section *4.1.2*, in impersonal constructions, *bìxū* gets a clear goal-oriented reading. In personal ones, it involves the speaker and the addressee (*wǒmen bìxū* 我们必须) and expresses a binding and unavoidable duty shared by the entire nation "on the new journey ahead" (*xīn de zhēngchéng shàng*), as suggested by the motto introducing the nine-musts in personal construction.

(III) The participant-external modalities are far more numerous than participant-internal ones, as shown also in the final slogan recalling the feasibility of goal achievement rather than the agent's success.

(IV) The role of multifunctional adverbs is more relevant compared to the evaluative ones.

The full significance of the procedural argumentation is better understood through the analysis of the evaluative adverbs occurring in the speech. While modal usage follows a specific design of the speech and is designed to attain a given rhetorical and persuasive goal, the occurrence of adverbs conveying the speaker's stance might reveal nuances and attitudes that are not consciously planned by the speakers. As reported above, we do not find in this speech many occurrences of evaluative adverbs. Among them, we have those typically used to express necessity and certainty but also multifunctional adverbs conveying degree (*zuì* 最, *gèng* 更, *gèngwéi* 更为) as well as scope (*dōu*), and linking adverbs (*jiù*). These multifunctional adverbs contribute to buttressing the core theme (in this speech, the 'only possibility' related to the role of the Party) and strengthening the sense of urgency in the implementation of the programme.

(V) The frequent use of adverbials emphasizing the meaning of continuity.

Other lexical items frequently used in adverbial function are those reinforcing the meaning of continuity as *bùduàn* 'continously' and *shǐzhōng* 始终 'all along' and those specifying the manner of performing actions as *quánmiàn* 'comprehensive' and *shēnkè* 深刻 'deep'. All these words, although not directly contributing to express the speakers' stance, are nonetheless strongly evocative of the commitments in "the journey ahead". Research results also show how epistemic adverbs (expressing certainty and inevitability, *yídìng* and *bìjiāng*, respectively) are more frequent than the others and that the distribution of the various adverbs across the genres also reveals variety. More specifically, the theme of inevitability, in the Chinese tradition *bìrán* 必然, is pervasive.

(VI) Inevitability is expressed in two ways: as 'only possibility', implying that the goal can only be accomplished under the guidance of the CCP; and as the historical inevitability of the realization of the Chinese dream (and inevitable failure of any attempt to stop it).

Further research (based on a significantly sampling of texts) is needed to test whether such features constitute a pattern in Chinese political discourse from a cross-linguistic perspective. This study contributes in shedding light on the linguistic devices, implicit intentions, and implications at work in this seminal speech.

References

Ai, H. (2012) "The Expression of Stance in Mandarin Chinese: A Corpus-Based Study of Stance Adverbs". *International Journal of Asian Language Processing*, 22(1): 1-14.

Alleton, V. (1984) *Les auxiliaires de mode en chinois contemporain*. Paris, Editions de la Maison des sciences de l'homme.

Biber, D. (2006) "Stance in Spoken and Written University Registers". *Journal of English for Academic Purposes*, 5: 97-116.

Cheng, M. (2019) "The Use of Modal Verbs in Political Debates", in C. Shei (Ed.) *The Routledge Handbook of Chinese discourse analysis*. Oxon, Routledge: 175-187.

Coates, J. (1983) *The Semantics of the Modal Auxiliaries*. London, Croom Helm.

Ernst, T. (2009) "Speaker-Oriented Adverbs". *Natural Language and Linguistic Theory*, 27(3), 497-544.

Hsie, C-L. (2005) "Modal Verbs and Modal Adverbs in Chinese: An Investigation into the Semantic Source". *USTWPL*, 1: 31-58.

--- (2008) "Evidentiality in Chinese Newspaper Reports: Subjectivity/Objectivity as a Factor". *Discourse studies*, 10(2): 205-229.

--- (2009) "Epistemic Stance Taking in Chinese Media Discourse", *Lilun Yuyanxue Yanjiu*, 12(3): 1-35.

Kimura, T. (2021) "China Still on the Revolutionary Journey". *China Today*, 2021/08/27. https://www.bjreview.com/Opinion/Voice/202108/t20210824_800256440.html (visited 2023/02/20).

Li, R. (2004) *Modality in English and Chinese. A Typological Perspective*. BocaRaton, Dissertation.com

Li, X. (2018) *The Reconstruction of Modality in Chinese-English Government Press Conference Interpreting a Corpus-Based Study.* Singapore, Springer.

Link, P. (2013) *An Anatomy of Chinese*. Cambridge, Harvard University press.

Longacre, R.E. (1974) "Narrative versus Other Discourse Genre. In R. Brend (Ed.) *Advances in Tagmemics*. Amsterdam, North Holland: 357-376.

Lü, S. 吕叔湘 ([1942] 2002). 中国文法要. Shenyan, Liaoning Jiaoyu Chubanshe.

Martin, J. (2022) "Rhetoric, Discourse and the Hermeneutics of Public Speech". *Politics*, 42(2):170-184. https://doi.org/10.1177/0263395720933779.

MOE - Ministry of Education (2021). 以史为鉴开创未来牢牢把握"九个必须". http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2021/2021_zt15/pl/202107/t20210706_542593.html (visited 2023/02/20).

Portner, P. (2009) *Modality* (Vol. 1). Oxford, Oxford University Press.

Riesigl, M. (2008) "Rhetoric of Political Speeches", in R. Wodak and V. Koller (Eds.) *Handbook of Communication in the Public Sphere*. Berlin/New York, De Gruyter: 243-269.

Sparvoli, C. (2012) *Deontico e anankastico. Proposta di ampliamento della tassonomia modale basata sull'analisi dei tratti distintivi dei modali cinesi inerenti dovere e necessità*. PhD Dissertation. Venice, Ca' Foscari University of Venice.

Sparvoli, C. (2017) "Competing markers: Acquisitive and core modals in Modern Mandarin Chinese". *Irish Journal of Asian Studies*, 3: 35-58.

Tian, Y.L. and R. Woo (2021) "Xi Warns against Foreign Bullying as China Marks Party Centenary". *Reuters* July 1st. https://www.reuters.com/world/china/beijing-set-celebrate-centenary-chinas-communist-party-2021-06-30/ (visited 2023/02/20).

van der Auwera, J. and A.V. Plungian (1998) "Modality's Semantic Map". *Linguistic Typology*, 2: 79-124.
Van Leeuwen, T. (2008) "News Genres". In R. Wodak and V. Koller (Eds.) *Handbook of Communication in the Public Sphere*. Berlin/ New York, De Gruyter: 343-362.
Xi, J. 习近平 (2021a) "在庆祝中国共产党成立100周年大会上的讲话". http://www.gov.cn/xinwen/2021-07/15/content_5625254.htm (visited 2023/02/20).
--- (2021b) "Speech at a Ceremony Marking the Centenary of the Communist Party of China". *Qiushi Journal*, 1 July http://en.qstheory.cn/2021-09/08/c_657713.htm (visited 2023/02/20).
Zhang, Y. 张谊生 (2014) 现代汉语副词研究. Beijing, Shangwu yinshuguan.
Zhongguo shehui kexue yuan yuyan yanjiu suo cidian bianji shi (2016) 现代汉语词典. Beijing: Shangwu yinshuguan.

THE SYNTAX OF *DE* (的)-OMISSION IN POST-NUMERAL POSITIONS

Sun Yangyu

University of Milano-Bicocca

1. Introduction[1]

In recent years, the syntax of DP has drawn a lot of attention. In Cartographic studies, many linguists (Cinque 2002; Giusti 2006; Poletto 2006, 2015; Laenzlinger 2015 among many others) demonstrate that the DP structure holds a parallel relation to the CP structure: DP is also composed by a "core area", which is the head noun (or the external head noun, in the sense of Cinque 2020), and a functional area with different types of modifiers hierarchically ordinated in a fixed sequence. This observation is proved to be universally valid, and from this point of view, cross-linguistic variations observed in the domain of DP become crucial to a further exploration of the syntactic essence of its universal structure.

---

[1] Abbreviation legend: CL = classifier; DE = nominal modifier marker *de*; NEG = negator; O-RC = object relative clause; RC = relative clause; S-RC = subject relative clause.

The DP structure in Mandarin Chinese also shows a clear affinity to the CP structure: nominal modifiers always precede their head noun, just as adverbs always precede the main verb; and the nominal modifier marker *de* (的) and the adverb marker *de* (地), except for being homophones, present a syntactic affinity in their surface use in DP and in CP (e.g., [AP-*de*-NP] and [AdvP-*de*-VP]). However, compared to verbal modifiers, nominal modifiers by nature exhibit a much more complicated system, and consequently the general nominal marker *de*, which attaches to almost all kinds of nominal modifiers, also becomes more intriguing to linguists, both syntactically and semantically.

This paper aims to provide a new perspective of considering the internal structure of DP, namely the phenomenon of *de*-omission, and present an even more articulated DP hierarchy than what has been proposed in the previous studies, exploiting the rigid pre-nominal order among the modifiers.

1.1 *de*-omission

In classical cases, the marker *de* can attach to every nominal modifier, whether it is a relative clause (RC), an adjective phrase (AP), a prepositional phrase (PP), or a noun (DP/NP). When there is more than one modifier in a single DP, it is natural to use multiple *de*s to mark their functional status.

(1)	[喜欢吃鱼**的**]	[好看**的**]	人
	Xǐhuān chī yú ***de***	*hǎokàn* ***de***	*rén*
	like eat fish DE	good-looking DE	person
	'a good-looking person who likes to eat fish'		

Alternatively, Sproat and Shih (1991: 593, footnote 9) mention a phenomenon which is called the "*de*-deletion": sometimes we can delete the first *de*(s) and only keep the final *de* immediately to the left of the head noun, without changing the grammatical judgement of the phrase. It can be applied to both adjectives and RCs:

Polysyllabic adjectives:

(2) [便宜] [好看]**的** 书
Piányí *hǎokàn-**de*** *shū*
Cheap interesting-DE book
'a interesting cheap book'

RCs:

(3) [喜欢吃鱼] [好看]**的** 那个人
Xǐhuān chī yú *hǎokàn-**de*** *nàgè rén*
like eat fish good-looking-DE that person
'that good-looking person who likes to eat fish'

In both (2) and (3), *de* after the first modifier is "deleted", whether it is an adjective or a RC, but the DP remain grammatical as their two-*de* variant. According to the authors, this is due to a general "stylistic" rule and does not imply that the higher modifier is a *de*-less direct modification.
However, if we investigate more deeply into this phenomenon, it turns out to be more than stylistic: the possibility of applying *de*-deletion is not at random but subject to syntactic restrictions. For instance, the judgement changes if we substitute the second adjective *hǎokàn* 好看 'interesting' in (2) with a disyllabic object RC (O-RC):

(4) *[便宜] [你买]**的** 书
* *Piányí* *nǐ mǎi-**de*** *shū*
Cheap you buy-DE book
Intended 'the cheap book you bought'

Moreover, in the two examples given by the authors, *hǎokàn* is respectively treated as an adjective ('interesting') and a RC ('good-looking'), totally depending on the context. This raises the question of the distinction between these two modifier categories: in the absence of an overt adjective marker, Mandarin adjectives can be sometimes confused with subject RCs (S-RCs).
Another curious point is the mechanism of this kind of "deletion": the non-existence of the first *de* can be logically explained either by a PF-deletion or by the absence of its external merge. The two methods

would have different implications for the formation of a [modifier-NP] phrase, but since this paper mainly focuses on the description of the phenomenon and does not investigate the mechanism which leads to it, in the following part I will call it "*de*-omission" instead of "*de*-deletion" to set away the question of whether *de* obligatorily merges with every non-direct modifier or not[2].

1.2 Main purpose and organization of the paper

This paper aims to investigate which syntactic restrictions can influence the possibility of using *de*-omission, and what is the implication to the universal syntactic structure of DP.

To achieve this, in Section 2.1 I briefly review some previous analyses about the nominal marker *de*, showing that even after a long history of study, its syntactic status is under discussion; in Section 2.2 I introduce some recent studies in Cartographic syntax about the DP structure, which provide a basic guideline to the detection of influencing factors of *de*-omission; in Section 2.3 I point out that the category of adjective in Mandarin is problematic, but I hold the idea that genuine adjectives exist in Mandarin, although not as pervasive as they are in other languages.

After these premises, I discuss the phenomenon of *de*-omission in Section 3: Section 3.1 tests various cases with controlled conditions on the influencing factors, and Section 3.2 summarizes the rules or the restrictions in using *de*-omission, which is followed by an analysis with which I argue that *de*-omission is essentially related to the functional hierarchy in the DP structure, reflecting the order of different modifier categories and suggesting a further refinement of the previous proposals; Section 4 concludes the discussion and addresses some unresolved issues.

2 An anonymous reviewer suggests that the PF-deletion option may fit here. As far as I am concerned, a possible drawback of the PF-deletion explanation is the fact that the "deleted" *de* triggered by the repetition is always the syntactically higher copy instead of the lower one, which is odd if we consider other cases of PF-deletion in the presence of a precedence.

2. Previous studies

2.1 Nominal marker *de* in Mandarin

In this subsection, I briefly review some previous studies about the nominal modifier marker *de* to show its complexity and the need of finding new perspectives. In Mandarin, *de* can be combined with every possible type of nominal modifier, including simple/complex adjectives, possessors, RCs, PPs and other non-predicative modifiers (Cheng and Sybesma 2009).

In the domain of APs, Liu (2016) shows that *de* is optional with non-predicative (in terms of Cinque 2010) adjectives but obligatory with predicative ones. However, the "optionality" is not arbitrary, the presence or the absence of *de* causes a change in the interpretation of the phrase. Paul (2005) argues that a modifier gets different readings according to its relative position to *de*: a modifier below *de* is interpreted as a "defining" property while one above *de* is interpreted as an "accessory" property.

In a broader sense, Sun (2015: 371-372) claims that there is a continuum of Chinese noun phrases from one Lexical pole (no use of *de*) to another Grammatical pole (necessary use of *de*). At the bottom of the continuum there are "nominalizations", where the nominal head is null but the obligatory *de* following the modifier clearly signals the noun category of the phrase. From this point of view, the use of *de* is associated with the uniqueness of the noun phrase: the most unique noun phrases, i.e., proper nouns, do not need *de*; while the most unspecified ones, composed only by a pure descriptive modifier, require the obligatory use of *de*.

However, these distinctions have no impact on cases of *de*-omission. In the next section we will see that the two or more modifiers may share the same property of accessory reading, in terms of Paul (2005), or the same specification function to the head noun, in terms of Sun (2015). In other words, we cannot draw parallels between the first *de*-less modifier in a *de*-omission context and a real *de*-less modifier immediately attached to the head noun.

Syntactically, the analyses are even more diversified. *De* has been argued to be a head-final C^0 (Cheng 1986), a non-root C head (Paul

2009, 2014), a D⁰ (Simpson 2002), a head of a more general ModifierP (Rubin 2002), a LINKER which hosts the inverted predicate at its specifier (den Dikken and Singhapreecha 2004, den Dikken 2006), or an underspecified classifier (Cheng and Sybesma 2009). Recently, Cinque (2020) also assumes that *de* in Mandarin is a C⁰ that participates in the derivation of the word order RC > *de* > head noun.

These proposals can be classified into two major groups: external to the modifier phrase ones and internal to the modifier phrase ones. External analyses (as in Figure 1) treat it as a functional head in the spine of DP, such as D⁰, LINKER, and underspecified classifier; while internal analyses (Figure 2) like C⁰ and Mod⁰ proposals deem it as a part of the modifier which do not directly select the head noun. Both groups of analyses are supported by good evidence, given that *de* always signals the nominal or predicative nature of its phrase, but at the same time it must be immediately preceded by a non-direct modifier or a RC.

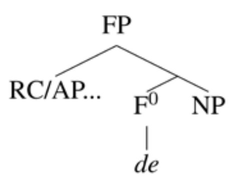


Figure 1. External analyses.

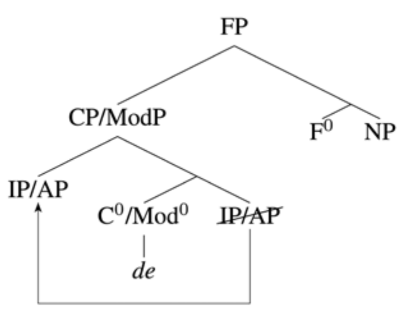


Figure 2. Internal analyses.

These two sets of analyses can both explain the fact that *de* always appear to the right of the modifier and to the left of the head noun, but they would also make different predictions regarding the constituency of a [modifier + *de*] sequence and the derivation of RCs. This paper focuses on the description of the phenomenon of *de*-omission, of which the syntactic mechanism is still unknown. Being aware of both options, I will leave the syntactic analysis of this marker for future research and concentrate on its implication to the functional hierarchy of DP in Mandarin.

2.2 Hierarchically organized nominal modifiers

As mentioned before, linguists have already found a parallelism between DP and CP, both of which have a functional area composed by extended projections organized in a fixed order. However, some language-specific particularities still remain to be explained.

Cinque (2018) observes that in other languages there is evidence which shows the following sequence in a DP structure (the symbol ">" stands for "hierarchically higher than")[3]:

(5) $RC_{finiteNon-Restr} > Dem > RC_{finiteRestr} > Num_{card} (Clf) > AP > RC_{reduced} > AP > N$

This hierarchy treats different types of nominal modifiers in different cartographic positions. It considers not only the basic AP vs. RC distinction, but also factors like finiteness and restrictiveness[4]. For instance, if a finite non-restrictive RC co-occurs with a finite restrictive RC, the non-restrictive one must precede the restrictive one. But as regard to the positions of RCs, it seems that in Mandarin Chinese the sequence does not hold well. The author shows three puzzles:

1. post-demonstrative RCs appear below cardinal numerals and classifiers rather than above them;
2. pre-demonstrative RCs are typically restrictive rather than nonrestrictive;

[3] Abbreviations: finiteNon-Restr. = finite non-restrictive; Dem = demonstrative; finiteRestr. = finite restrictive; $Num_{card.}$ = cardinal numeral; Clf = classifier.

[4] The notion of restrictiveness and finiteness are not relevant to the purpose of this paper. I refer to Cinque (2010) for the meaning of these terms.

3. the post-demonstrative, post-numeral, post-classifier position of the RC is apparently open to a nonrestrictive interpretation.

He proposes a possible solution to these puzzles that Chinese RCs can be analyzed as reduced nonfinite RCs. But the mystery remains: compared to reduced RCs in other languages (which are usually participle phrases), Chinese RCs can have an overt subject; in addition, they are unbounded and sensitive to islands. In this paper I will leave alone the discussion of finite vs. reduced RC, since it would go way beyond the main purpose. What I insist in this paper is that the problematic post-numeral area should be further refined based on evidence of *de*-omission.

To be added to the first puzzle is that some complex APs and more rarely, reduced RCs (which cannot take an explicit syntactic subject) are allowed (although not preferred) to appear between the demonstrative 'Dem' and the chunk of [cardinal numeral + classifier] 'Num_{card}(Clf)', as long as the cardinal number is not "one" or null, and the number is not a new information in the context. As shown in (6), this position is supposed to be a restrictive finite RC, but it is apparently nonfinite and can difficultly be restrictive.

(6)	那	**脏兮兮**	的	三只	小猫
	Nà	***zāngxīxī***	*de*	*sān-zhī*	*xiǎo-māo*
	that	dirty	DE	three-CL	little-cat

'those three dirty kittens' (preferably nonrestrictive)

The rare use and the restrictions of this modifier position guide us to conclude that, even if this position may exist in Mandarin Chinese, the word order must be derived from a syntactic movement of the lower modifiers, namely APs and reduced RCs.

Along with Cinque (2010), this hierarchy naturally goes against the proposal of den Dikken and Singhapreecha (2004) and den Dikken (2006), in which all the pre-nominal modifiers result from an obligatory predicate inversion (as figured out in Sio 2006 and Paul 2012 among others). In the meanwhile, the universal direct modifier AP position suggests that there is no need to assume that all *de*-less modified DPs are compounds.

Except for factors like AP/RC distinction, finiteness, and restrictiveness, Larson and Takahashi (2007) also point out that the difference between stage-level and individual-level interpretation can affect the ordering of nominal modifiers.

In a more recent work, Cinque (2019) includes more subtle types of RCs (integrated vs. non-integrated; kind-defining; markedness, amount/maximalizing; infinitival; stage-level vs. individual-level) in world languages and tests their relative positions in the nominal functional area. He proposes that all these RC types have their own merge position by testing their relative orders, and the refined hierarchy is shown in Figure 3.

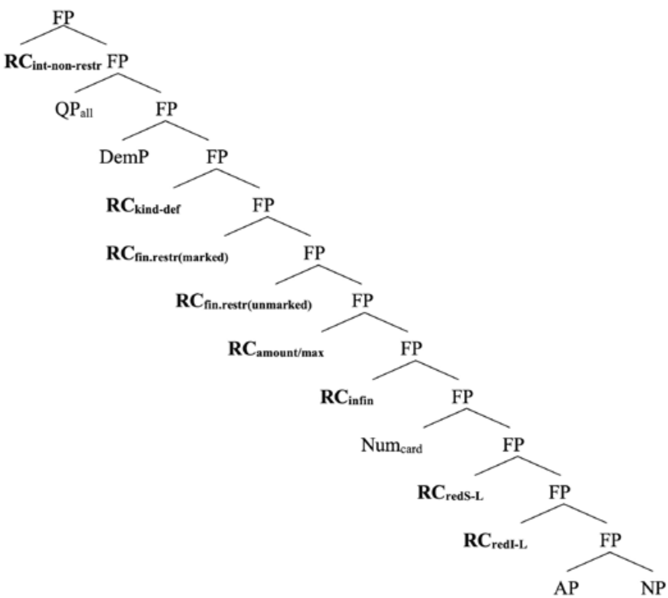


Figure 3. DP hierarchy in Cinque (2019).

In this paper I confine my discussion to post-demonstrative, post-classifier nominal modifiers (i.e., the ones lower than the FP holding "Num_{card}" in Figure 3), which according to this hierarchy only the AP/RC distinction and stage-level/individual-level interpretation should affect the test result.

2.3 A brief note on APs in Mandarin[5]

It is well-known to linguists that, in the absence of a real categorial morpheme that indicates the adjective state of a word, it is hard to recognize adjectives in Mandarin. Adjectives in Mandarin can be either marked by *de* or be *de*-less. Zhu Dexi (1999: 169) mentions that there are three types of nominal modifiers which can be attached to the noun without the use of *de*: Adjective + N (*xīn-fángzi* 新房子, 'new-house'), Noun + N (*mùtou fángzi* 木头房子, 'wood-house') and Quality + N (*jīn-fángzi* 金房子, 'gold-house = precious house'). Following this idea, if a *de*-less modifier is not a NP, we can infer that it must be an AP. The problem of identifying AP appears when modifiers are marked by *de*.

Some degree adverbs may test the adjectivehood of a word, but unlike "very" in English, according to Li and Thompson (1981: 339-340), the adverb *hěn* 'very' is compatible both to "adjectival verbs" and certain "experiential verbs":

(7)	他	很	高。	
	Tā	***hěn***	*gāo.*	
	3SG	very	tall	
	'He is very tall.'			
(8)	我	很	怕	狗。
	Wǒ	*hěn*	*pà*	*gǒu.*
	I	very	afraid	dog
	'I am very afraid of dogs.'			

From this point of view, it is difficult to trace the distinction between multisyllabic APs and non-dynamic verbs in Mandarin. Even a V-O structure cannot be a guaranteed criterion to the grammatical category, as some V-O constructions can be lexicalized as APs. For example, *yǒuqián* 有钱 'rich', literally means 'have.money' like a RC, but can also appear as a direct modifier without *de*: *yǒuqián-rén* 有钱人 'rich man', literally 'have.money-man'.

Huang *et al.* (2009: 22-24) propose a "*duì*-test" to distinguish an ad-

[5] This section is largely reduced compared to the unreviewed version, but I would like to thank the two anonymous reviewers for their precious suggestions about this topic.

jective from a verb, namely the preposition *duì* 对 'to, of' can take a semantic object and form a [*duì*+semantic-object+AP] sequence, while a verb cannot substitute the position of the adjective, just like in English the verb "love" can be semantically almost equivalent to the adjectival predicate "be fond of", where "of" is only compatible with the adjective, while the verb chooses an object without the use of "of".

(9)	这个	工作	**对**	你	很	合适。
	Zhè-gè	*gōngzuò*	***duì***	*nǐ*	*hěn*	*héshì.*
	this-CL	job	to	you	very	suitable

'This job is suitable for you.'
(Huang *et al.* 2009: 22)

But when an adjective is lack of semantic object, like *yǒuqián* 'rich', it cannot have a [*duì*+semantic-object+AP] form:

(10)	*张三	**对**	你	很	有钱。
	* *Zhāngsān*	***duì***	*nǐ*	*hěn*	*yǒuqián*
	Zhangsan	to	you	very	rich

Intended 'Zhangsan is very rich to you'

This test also fails in identifying all APs in Mandarin.
To sum up, in the presence of *de*, the grammatical status of APs in Mandarin is not evident. A proposal regarding this issue is beyond the main aim of this paper. Nevertheless, the existence of non-nominal direct modifiers without *de* (if they do not form a compound with the head noun), like in other languages with clear AP+NP/NP+AP constructions, suggests that we should better treat them as APs instead of RCs. Based on this conclusion, I admit that APs is a distinct grammatical category, but propose later that unlike simple APs, modified APs (e.g., "very rich", "not rich") syntactically behave like S-RCs in DP.
In the remaining part of this paper, I will ignore the direct *de*-less modifiers and concentrate on those with *de*, since only in the presence of it can the *de*-omission test make sense.

3. DP Hierarchy and *de*-omission

3.1 Factors that influence *de*-omission

In this section, I will test in which conditions *de*-omission is allowed and in which it is not, considering the above-mentioned two factors (AP vs. RC, stage-level vs. individual level), singularly or jointly, and proposing that the distinction between modified AP vs. simple AP / O-RC vs. S-RC also plays a role in licensing *de*-omission, which implies that this should also be added to the functional hierarchy of DP structure. To avoid a possible phonetic or prosodic effect to the word-order judgement, in the following examples the number of syllables in one modifier will be maintained the same as another modifier in the same phrase.

Before the tests, I would like to draw two general rules in Mandarin related to the use of *de*-omission:

(i) Two or more modifiers with all the same properties can always share one *de*:

(11) (simple) AP$_{\text{stage-level}}$>(simple) AP$_{\text{stage-level}}$>DE>N

这个	**悲伤**	**忧愁**	的	男生
Zhè-gè	***bēishāng***	***yōuchóu***	*de*	*nánshēng*
This-CL	sad	worried	DE	boy

‘this boy who is sad and worried’

(12) S-RC$_{\text{individual-level}}$>S-RC$_{\text{individual-level}}$>DE>N

这个	**喜欢吃鱼**	**不爱洗澡**
Zhè-gè	***xǐhuān chīyú***	***bù ài xǐzǎo***
This-CL	like eat-fish	NEG love take-shower
的	男生	
de	*nánshēng*	
DE	boy	

‘this boy who likes eating fish and doesn’t like taking shower’

(ii) If *de* appears after every modifier, the sentence can be (fully or marginally) accepted even when the hierarchy is violated:

(13) S-RC$_{\text{individual-level}}$>DE>S-RC$_{\text{stage-level}}$>DE>N:[6]

昨天	那个	学	语言学	的
Zuótiān	*nà-gè*	*xué*	*yǔyánxué*	*de*
Yesterday	that-CL	study	linguistics	DE
和你说话	的	男生		
hé-nǐ shuōhuà	*de*	*nánshēng*		
with-you talk	DE	boy		

'the boy who studies linguistics and who talked with you yesterday'

Rule (i) can be naturally considered as a simple case of symmetric coordination of the two modifiers since they are syntactically indistinct. I will use this rule to demonstrate that modified APs are interchangeable with, and thus syntactically treated as, S-RCs.

Rule (ii) implies that the presence of *de* licenses the interruption of the DP functional hierarchy. I will leave the exact syntactic mechanism of

---

6 This seems to go against the example made by Del Gobbo (2005: 298-299) who argues that an individual level RC cannot precede a stage level one at all:

(i) *[RC Wǒ zuótiān kànjiàn de] [RC xǐhuān qù yīnyuèhuì de] rén shì Zhāngsān.*
I yesterday meet DE like go concerts DE person is Zhangsan
'The person I met yesterday who likes to go to concerts is Zhangsan.'

(ii) **[RC Xǐhuān qù yīnyuèhuì de] [RC Wǒ zuótiān kànjiàn de] rén shì Lǐsì.*
like go concerts DE I yesterday meet DE person is Lisi
'The person who likes to go to concerts who I met yesterday is Lisi.'

(iii) *Nà-gè [RC zuótiān méiyǒu lái de] [RC hěn xǐhuān shàng kè de]*
that-CL yesterday not come DE very like go class DE
xuéshēng jiào Zhangsan
student call Zhangsan
'The student who didn't come yesterday who likes to come to class very much is called Zhangsan.'

(iv) **Nà-gè [RC hěn xǐhuān shàng kè de] [RC zuótiān méiyǒu lái de]*
that-CL very like go class DE yesterday not come DE
xuéshēng jiào Zhangsan
student call Zhangsan.
'The student who likes to come to class very much who didn't come yesterday is called Zhangsan.'

According to my informants, with a pause between the two RCs, (iv) is grammatical and (ii) is marginal, probably because the two RCs in (ii) are different both in the stage-/individual-level reading, and in terms of O-RC/S-RC.

it for future research. *De*-omission contexts, instead, are absolute in reflecting the allowed modifier orders and compatibilities.
After that we have excluded these cases, now we can test the possibility of using *de*-omission controlling the following factors:

1. Stage-level vs. Individual-level ("who talked to you" vs. "who likes linguistics");
2. Modified AP vs. simple AP ("very/not rich" vs. "rich");
3. AP vs. RC ("rich" vs. "who has a lot of money");
4. S-RC vs. O-RC ("who bought a yacht" vs. "which I gave to her").

First of all, *de*-omission seems to be banned when a stage-level multi-syllabic simple AP co-occurs with another individual-level one, while two stage-level ones or two individual-level ones can be freely coordinated before *de* according to Rule (i):

(14) *(simple) $_{\text{APstage-level}}$>(simple) AP$_{\text{individual-level}}$>DE>N

*这个	**愤怒**	**高大**	的	男生
**Zhè-gè*	***fènnù***	***gāodà***	*de*	*nánshēng*
This-CL	angry	tall.big	DE	boy

Intended 'this tall boy who is angry'

(15) *(simple) AP$_{\text{individual-level}}$>(simple) AP$_{\text{stage-level}}$>DE>N

*这个	**高大**	**愤怒**	的	男生
* *Zhè-gè*	***gāodà***	***fènnù***	*de*	*nánshēng*
This-CL	tall.big	angry	DE	boy

Intended 'this tall boy who is angry'

These phrases must be remedied by using two *de*s, no matter in which order, given Rule (ii), or by changing the individual-level simple AP to a direct *de*-less modifier:

(16) (simple) AP$_{\text{stage-level/individual-level}}$ >DE> (simple) AP$_{\text{stage-level/individual-level}}$> DE>N

这个	**愤怒/高大**	的	**高大/愤怒**	的	男生
Zhè-gè	***fènnù/gāodà***	*de*	***gāodà/fènnù***	*de*	*nánshēng*
This-CL	angry/tall.big	DE	tall.big/angry	DE	boy

'this tall boy who is angry'

(17) (simple) AP$_{\text{stage-level}}$>DE>(simple) AP$_{\text{individual-level}}$>N

这个	**愤怒**	的	**高大**	男生
Zhè-gè	***fènnù***	*de*	***gāodà***	*nánshēng*
This-CL	angry	DE	tall.big	boy

'this tall boy who is angry'

This implies that a simple AP$_{\text{stage-level}}$ is substantially kept different from a simple AP$_{\text{individual-level}}$ in the DP structure. In any case, *de* after the only stage-level simple AP must appear.

Then, the stage-level/individual-level distinction also affects S-RC, but in a different way: *de*-omission is fully allowed when S-RC$_{\text{stage-level}}$ precedes S-RC$_{\text{individual-level}}$, but the other way round is marginal:

(18) S-RC$_{\text{stage-level}}$>S-RC$_{\text{individual-level}}$>DE>N:

那个	**和你说话**	**喜欢数学**	的	男生
nà-gè	***hé nǐ shuōhuà***	***xǐhuān shùxué***	*de*	*nánshēng*
that-CL	with-you talk	like mathematics	DE	boy

'the boy who likes mathematics and who talked to you'

(19) ?S-RC$_{\text{individual-level}}$>S-RC$_{\text{stage-level}}$>DE>N:

?那个	**喜欢数学**	**和你说话**	的	男生
?*nà-gè*	***xǐhuān shùxué***	***hé nǐ shuōhuà***	*de*	*nánshēng*
that-CL	like mathematics	with-you talk	DE	boy

'the boy who likes mathematics and who talked to you'

As predicted by the hierarchy proposed in the previous works, within the same modifier category, a stage-level one is always higher than an individual-level one, and this is borne out by the markedness of (19). However, unlike the cases of AP in (14) and (15), these two kinds of S-RC are not categorially distinct, and thus *de*-omission is allowed with a preference.

When a S-RC co-occurs with an AP, simple APs and modified APs show different "compatibility" with S-RC in *de*-omission: a simple AP and a S-RC cannot share only one *de* even when they are of the same reading in (20) and (21), while a modified AP can appear with a S-RC with *de*-omission, as long as they have the same reading in the stage-/individual-level distinction (see (22) and (23)), or when the stage-level > individual-level order is maintained (see (24) and (25)):

(20) *(simple) $\text{AP}_{\text{stage-level}}$>$\text{S-RC}_{\text{stage-level}}$>DE>N:

*那个	**愤怒**	**背包**	的	男生
**Nà-gè*	***fènnù***	***bèibāo***	*de*	*nánshēng*
that-CL	angry	carry.bag	DE	boy

Intended 'the angry boy who is carrying a bag'

(21) *$\text{S-RC}_{\text{stage-level}}$>(simple) $\text{AP}_{\text{stage-level}}$>DE>N:

*那个	**背包**	**愤怒**	的	男生
**Nà-gè*	***bèibāo***	***fènnù***	*de*	*nánshēng*
that-CL	carry.bag	angry	DE	boy

Intended 'the angry boy who is carrying a bag'

(22) (modified) $\text{AP}_{\text{individual-level}}$>$\text{S-RC}_{\text{individual-level}}$>DE>N:

那个	**特别高大**	**喜欢数学**	的	男生
Nà-gè	***tèbié gāodà***	***xǐhuān shùxué***	*de*	*nánshēng*
that-CL	very tall	like mathematics	DE	boy

'the boy who is very tall and who likes mathematics'

(23) $\text{S-RC}_{\text{individual-level}}$> (modified) $\text{AP}_{\text{individual-level}}$ >DE>N:

那个	**喜欢数学**	**特别高大**	的	男生
Nà-gè	***xǐhuān shùxué***	***tèbié gāodà***	*de*	*nánshēng*
that-CL	like mathematics	very tall	DE	boy

'the boy who likes mathematics and who is very tall'

(24) (modified) $\text{AP}_{\text{stage-level}}$>$\text{S-RC}_{\text{individual-level}}$>DE>N:

那个	**特别愤怒**	**喜欢数学**	的	男生
Na-ge	***tèbié fènnù***	***xǐhuān shùxué***	*de*	*nánshēng*
that-CL	very angry	like mathematics	DE	boy

'the boy who is very angry and who likes mathematics'

(25) $\text{S-RC}_{\text{stage-level}}$> (modified) $\text{AP}_{\text{individual-level}}$ >DE>N:

那个	**和你说话**	**特别高大**	的	男生
Nà-gè	***hé nǐ shuōhuà***	***tèbié gāodà***	*de*	*nánshēng*
that-CL	with-you talk	very tall	DE	boy

'the boy who talked to you and who is very tall'

This contrast can be explained if modified APs in Mandarin are indistinctly analyzed as S-RCs, while simple APs, as I argued before, are kept apart as an individual modifier category, probably lack a predicate-like syntactic structure which distinguishes them from reduced or finite RCs, even if some APs are semantically predicative. This analysis fits the intuition that the head noun can be treated as the subject of a modified AP which functions as a predicate, just like the head-noun-subject to a S-RC.

Things should be different, then, with O-RCs: syntactically, modified APs cannot have a similar argument structure to O-RCs, and this is also represented by *de*-omission. Notice that I do not consider the individual-/stage-level distinction for O-RCs, since O-RCs always involve an external argument DP and thus rarely have an individual-level interpretation to the object head noun.

(26) *(modified) AP>O-RC>DE>N:

*这份	**十分贵重**	**他买给我**	的	礼物
**Zhè-fèn*	***shífēn guìzhòng***	***tā mǎi gěi wǒ***	*de*	*lǐwù*
This-CL	very-expensive	he buy to-me	DE	present

*O-RC> (modified) AP>DE>N:

(27)

*这份	**他买给我**	**十分贵重**	的	礼物
**Zhè-fèn*	***tā mǎi gěi wǒ***	***shífēn guìzhòng***	*de*	*lǐwù*
This-CL	he buy to-me	very-expensive	DE	present

Intended ‘this expensive present which he bought me’

The two phrases must be fixed by cancelling the use of *de*-omission, i.e., by adding an extra *de* after the first modifier, which fits Rule (ii):

(28) (modified) AP>DE>O-RC>DE>N:

这份	**十分贵重**	**的**	他买给我	的
Zhè-fèn	*shífēn-guìzhòng*	*de*	*tā mǎi gěi-wǒ*	*de*
This-CL	very-expensive	DE	he buy to-me	DE
礼物				
lǐwù				
present				

(29) O-RC>DE>(modified) AP>DE>N:

这份	**他买给我**	**的**	十分贵重	的
Zhè-fèn	*tā mǎi gěi wǒ*	*de*	*shífēn guìzhòng*	*de*
This-CL	he buy to-me	DE	very-expensive	DE
礼物				
lǐwù				
present				

‘this expensive present which he bought me’

The grammaticality judgement goes slightly better with the sequence S-RC > O-RC[7], while the inverted order is banned just as (27) with a modified AP:

(30) ?S-RC>O-RC>DE>N:

?那个	**和你说话**	**我不认识**	的	男生
?*Nà-gè*	***hé nǐ shuōhuà***	*wǒ bù rènshí*	*de*	*nánshēng*
that-CL	with-you talk	I NEG-know	DE	boy

'the boy who I don't know and who talked to you'

(31) *O-RC>S-RC>DE>N:

那个	**我不认识**	**和你说话**	的	男生
**Nà-gè*	*wǒ bù rènshí*	*hé nǐ shuōhuà*	*de*	*nánshēng*
that-CL	I NEG-know	with-you talk	DE	boy

Again, without *de*-omission, both the sequence of S-RC>DE>O-RC>DE>N and that of O-RC>DE>S-RC>DE>N are acceptable.

3.2 "Rules" of *de*-omission

In Section 3.1 I show that *de*-omission can be used as a tool to detect which kinds of nominal modifiers are identified as the same category. Adding other conditions which I did not mention before, the result of *de*-omission test is illustrated in Table 1. "Y" means that when the "first modifier" precedes the "second modifier", *de*-omission is allowed; "N" instead means that the sequence is not allowed, and the first modifier must be followed by a *de*; "?" stands for the marginality of the testing sequence.

[7] A possible explanation to the acceptability of (30) is the fact that other than S-RCs and O-RCs, complex RCs can also be modifiers in DP. If the stacked RCs are interpreted as "who talked to you (BUT) I don't know him" with a null-object in the "O-RC" part, it should be acceptable since the nominal modifier is syntactically a single RC, and no matter which method we adopt for the derivation of RC in Mandarin, it only happens once instead of twice, avoiding any conflicts in the hierarchy and in the interpretation of nominal elements.

First modifier / Second modifier	(simple) AP individual-level	(simple) AP stage-level	O-RC	(modified) AP individual-level	S-RC individual-level	(modified) AP stage-level	S-RC stage-level
(simple) AP individual-level	Y	N	N	N	N	N	N
(simple) AP stage-level	N	Y	N	N	N	N	N
O-RC	N	N	Y	N	?	N	?
(modified) AP individual-level	N	N	N	Y	Y	Y	Y
S-RC individual-level	N	N	N	Y	Y	Y	Y
(modified) AP stage-level	N	N	N	?	?	Y	Y
S-RC stage-level	N	N	N	?	?	Y	Y

Table 3. Result of the *de*-omission test.

Other than the fact that *de*-omission is always allowed with two modifiers of the exact same properties, as claimed in Rule (i) and shown by the diagonal Y-line in the table, from this table we can see that individual-level simple APs, stage-level simple APs and O-RC are totally independent from each other and from modified APs and S-RCs; while modified APs and S-RCs form a block with a preferred but not obligatory order of stage-level > individual-level. A S-RC sometimes can precede an O-RC without the aid of *de*, but as I explain in footnote 7, it might be due to a reanalysis of the modifiers as a single complex RC.

If the analysis is on the right track, this result suggests that:

1. Simple APs are radically different from modified APs;
2. Individual-level simple APs and stage-level simple APs are two distinct modifier categories;
3. Modified APs are analyzed as S-RCs;
4. O-RCs are different from S-RCs, although they are both RCs and usually analyzed as a unified modifier category;
5. Within the block of modified APs and S-RCs, the stage-level > individual-level order is still preferred.

Based on these observations extracted from data of *de*-omission, the hierarchy of the post-demonstrative, post-numeral functional area in Mandarin DP should be revised as follows:

(32) Dem > $\text{RC}_{\text{finiteRestr}}$ > Num_{card}(Clf) > (modified)AP/S-RC > O-RC > (simple)$\text{AP}_{\text{stage-level}}$>(simple)$\text{AP}_{\text{individual-level}}$ >N

4. Conclusion and discussion

In this paper I introduce the phenomenon of *de*-omission as a diagnostic tool to recognize nominal modifiers of the same syntactic type, and thus it is also a proof of a more articulated post-numeral DP functional area. Crucially, I point out that there is a clear cut between simple APs and modified APs: the former ones are distinguished into two sperate categories: stage-level simple APs and individual-level simple APs, while modified APs share the same syntactic status as S-RCs in the tested area. Moreover, O-RCs cannot be put together with S-RCs, as they cannot appear with S-RCs without the use of two or more *de*s. Therefore, the revised hierarchy in (32) based on previous studies and observations from *de*-omission cases also considers these new distinctions between modifiers.

The well-known complexity of the marker *de* in the DP domain still leaves us many mysteries. I believe that the phenomenon of *de*-omission can also have an impact on determining the syntactic nature of this marker.

Moreover, I have only talked about a very restrictive zone of DP, leaving all the pre-numeral area unexplained. Since *de*-omission also happens there, does it exhibit the same categorial sensitivity? Can it reveal the derivation of the famous pre-demonstrative RCs? I will leave these questions for future research.

References

Cheng, L.L.S. (1986) "*de* in Mandarin". *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 31(4): 313-326.

Cheng, L.L.S. and R. Sybesma (2009) "*De* as an Underspecified Classifier: First Explorations". *Yuyanxue Luncong*, 39: 123-156.

Cinque, G. (2002) *Mapping Functional Structure: A Project. Functional Structure in DP and IP*, vol. 1. Oxford/New York, Oxford University Press: 3-11.

--- (2010) *The Syntax of Adjectives: A Comparative Study*, vol. 57. Cambrige (MA), MIT press.

--- (2018) “A Note on Chinese Relative Clauses (Seen from a Distance)”. *Handout of conference talk at EACL-10* (2018).

--- (2019) “The Different Merge Positions of the Different Types of Relative Clauses”, in K. Ramshøj Christensen, H. Jørgensen and J.L. Wood (Eds.) *The Sign of the V – Papers in Honour of Sten Vikner*. Department of English, School of Communication & Culture: Aarhus University: 131-147. https://doi:10.7146/aul.348.93.

--- (2020) *The Syntax of Relative Clauses: A Unified Analysis*. Cambridge: Cambridge University Press.

Del Gobbo, F. (2005) “Chinese Relative Clauses: Restrictive, Descriptive or Appositive?”, in L. Brugè, G. Giusti, N. Munaro, W. Schweikert and G. Turano (Eds.) *Contributions to the XXX Incontro di Grammatica Generativa*. Venice, Libreria Editrice Cafoscarina: 287-305.

Den Dikken, M. (2006) *Relators and Linkers*. Cambridge (MA): MIT Press.

Den Dikken, M. and P. Singhapreecha (2004) “Complex Noun Phrases and Linkers”. *Syntax* 7(1): 1-54.

Giusti, G. (2006) “Parallels in Clausal and Nominal Periphery”, in M. Frascarelli (Ed.) *Phases of Interpretation*. Berlin/New York, De Gruyter: 163-84.

Huang, C.-T.J., Y-H.A. Li and Y. Li (2009) *The Syntax of Chinese*. Cambridge, Cambridge University Press.

Laenzlinger, C. (2015) “The CP/DP (non-) Parallelism Revisited”, in U. Shlonsky (Ed.) *Beyond Functional Sequence*: *The Cartography of Syntactic Structures*. Oxford, Oxford University Press: 128-150.

Larson, R. and N. Takahashi (2007) “Order and Interpretation in Prenominal Relative Clauses”. *MIT Working Papers in Linguistics*, 54: 101-120.

Li, C. and S. Thompson (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, University of California Press.

Liu, F.-H. (2016) “Syntactic Sources of Adjectives in Mandarin Chinese”. *Studies in Chinese Linguistics*, 37: 38-55.

Paul, W. (2005) “Adjectival Modification in Mandarin Chinese and Related Issues”. *Linguistics*, 43(4): 757-793. https://doi.org/10.1515/ling.2005.43.4.757.

--- (2009) “Consistent Disharmony: Sentence-Final Particles in Chinese”. *Cambridge Occasional Papers in Linguistics*, 5: 1-24.

--- (2012) “Why Chinese DE is not like French DE: A Critical Analysis of the Predicational Approach to Nominal Modification”. *Studies in Chinese Linguistics*, 33(3): 183-210.

--- (2014) “Why Particles are not Particular: Sentence-Final Particles in Chinese as Heads of a Split CP”. *Studia Linguistica*, 68(1): 77-115.

Poletto, C. (2006) “Parallel Phases: A Study on the High and Low Left Periphery of Old Italian”, in M. Frascarelli (Ed.) *Phases of Interpretation*. Berlin/New York, De Gruyter: 261-94.

--- (2015) “Word Orders in the Old Italian DP”. *The Cartography of Syntactic Structures*, 10: 109-127.

Rubin, E.J. (2002) *The Structure of Modifiers*. Draft version.

Simpson, A. (2002) “On the Status of ‘Modifying’ DE and the Structure of the Chinese DP”, in S-W. Tang and C-S. Liu (Eds.) *On the Formal Way to Chinese Languages*. Stanford (CA), CSLI: 74-101.

Sio, J.U-S.S. (2006) *Modification and Reference in the Chinese Nominal*. PhD Dissertation, Leiden University.

Sproat, R. and C. Shih (1991) “The Cross-Linguistic Distribution of Adjective Ordering Restrictions”, in C. Georgopoulos and R. Ishihara (Eds.) *Interdisciplinary Approaches to Language*. Dordrecht, Springer: 565-593.

Sun, C. (2015) “The Uses of DE as a Noun Phrase Marker”, in W.S-Y. Wang and C. Sun (Eds.) *The Oxford Handbook of Chinese Linguistics*. Oxford, Oxford University Press: 362-376.

Zhu, D. 朱德熙 (1999) 朱德熙文集第一卷 (语法讲义). Beijing, The Commercial Press.

THE ICONIC COGNITIVE PRINCIPLES OF MANDARIN CHINESE WORD ORDER: PEDAGOGICAL AND LEARNING PERSPECTIVES

Tommaso Tucci

Catholic University of the Sacred Heart (Milan)

Among the approaches dedicated to studying Mandarin Chinese (MC) word order, the cognitive-functional approach inaugurated by Tai (1985, 1989, and 1993) is particularly interesting because of its potential from the pedagogical point of view. Central to this approach, which was resumed in several studies on word order (e.g., Ho 1993; Hu 1995; Jiang 2009, 2017; Loar 2011; Morbiato 2017, 2019), is the exploration of recurrent patterns in MC syntax to formulate specific principles that motivate the mapping between form and function within the linguistic structure. Given its theoretical premises, this strand of analysis is well placed in that diverse landscape of approaches that can be labeled under the nomenclature of Cognitive Linguistics (CL), so much so that Tai's studies are considered "to mark the beginning of work on CL in China" (Shu *et al.* 2019: 17). In the present study, the focus is placed exclusively on the most cited cognitive iconic principles, leaving aside those of a functional nature for two different reasons, in addition to space constraints. First, because the potential for L2 instruction arising from the teaching of these principles is, to this day, almost unexplored, when on the other hand, the notions pertaining to functionalist approaches have already managed to integrate, albeit still too marginally, into the teaching curricula (e.g., topic-comment dichotomy, the notion of focus)[1]. Secondly, the concept of motivation evoked by the conceptualizations of these principles is consistent with the fundamental assumptions underlying CL, such as the non-autonomy of syntax and the embodied nature of language. Listed below are some of the most discussed cognitive principles in the literature on the Cognitive-functional approach to MC word order:

1. The Principle of Temporal Sequence (PTS).
2. The Principle of General-Preceding-Particular (GPP).
3. The Principle of Semantic Proximity (SP).
4. The Empathetic Principle (EP).

The choice to narrow the discussion to these principles is based on a critical analysis of the literature on the cognitive-functional framework. It aims to delimit a core set of fundamental iconic principles sound in motivating MC word order from a pedagogical perspective.
The present paper, therefore, aims to provide a review of the iconic principles related to the cognitive-functional framework, highlighting how the application of cognitive principles to the pedagogy of L2 MC is entirely in line with the key assumptions underlying the Applied Cognitive Linguistics (ACL) subfield. This will be suggested by first introducing the topic of linguistic motivation and iconicity as understood in CL and the cognitive-functional approach to MC word order (Section 1). Afterward, the four cognitive principles mentioned above will be briefly introduced and illustrated, providing an overview of their explanatory power according to the literature on the subject (Section 2). Finally, the explanatory power of these principles in L2

[1] See in this regard the provisions of *Il sillabo della lingua cinese* ('The Syllabus of the Chinese Language') (Ministero dell'Istruzione, dell'Università e della Ricerca 2016), a syllabus drafted by Italian Mandarin Chinese language teachers and scholars in order to standardize the linguistic contents to be presented in the curricula of MC classes held in Italian High Schools.

learning and pedagogy will be suggested by tying their treatment to the assumptions underlying the broader ACL framework (Section 3). It is hoped, therefore, that this paper will contribute to providing fresh perspectives within ACL approaches to L2 pedagogy.

1. Linguistic motivation and iconicity as a linguistic phenomenon

In the words of Dirven and Verspoor (2004: 13), motivation "refers to non-arbitrary links between a form and the meaning of linguistic expressions" (see also Radden and Panther 2004, 2011; Panther 2021). In the ACL literature, the term refers mainly to the pedagogical field. It is therefore used as an extension of the concept of 'semantic explanation', which refers primarily to grammatical structure in a vision that sees vocabulary and grammar unfolding in a single continuum. According to Taylor (2002, 2004), one of the strongest supporters of a pedagogical grammar based on the theoretical foundations of CL, the term 'motivation' is preferable to the latter because it would also include the connections within the linguistic system. In contrast, the former would motivate expressions only based on their conceptualization.

Among the various manifestations of linguistic motivation, iconicity is the one that pertains to the connections between form and meaning (Boers and Lindstromberg 2006; Radden and Panther 2004). There are numerous interpretations of iconicity in linguistics, starting with Charles S. Peirce's semiotic studies, according to which an icon is a sign that resembles its object (e.g., Peirce 1998 [1902]). A definition of iconicity often adopted in the literature on the subject is derived from the studies of Haiman (1980, 1985), where iconicity is understood as "a similarity between form and meaning as conceived by a human mind" (Tabakowska 2009: 133). Iconicity, and consequently motivation, can be intralinguistic or extralinguistic. The former refers to correspondences between units within the linguistic system, while the latter links linguistic signs to the phenomenological world, or rather to the conceptualization of reality as filtered by the cognition of the human being. The intralinguistic type is most studied in the generativist literature (e.g., Krivochen and Lackova 2020: 97), being committed

to sounding out the internal dynamics of a system considered autonomous. The assumptions underlying CL, on the other hand, see extralinguistic iconicity as proof of the inadequacy of the formalist view of the autonomy of syntax.
Iconicity is a crucial notion in the study of typological features of MC. Its manifestations are not limited to the writing system (*imagic* iconicity, cf. Tai 1993) but refer to the entire linguistic structure. The phenomena of syntactic iconicity are among the various manifestations of *diagrammatic* iconicity, a term used by Haiman to emphasize its relational nature. Scholars agree that the morphosyntactic structures of isolating languages as the MC are more iconic than morphologically richer languages, where the motivation behind different syntactic constructions has often lost its transparency (Haiman 1985). Tai (2002: 348), in this regard, states:

> It is important to point out that very few researchers of iconicity deny the existence of arbitrary grammatical principles. What they have attempted to show is that not all of the grammatical principles are arbitrary and autonomous. In my recent series of works, I have shown that Chinese grammar is more conceptually based and iconically motivated than English.

One of the basic assumptions of the cognitive-functional approach is the strong iconically motivated nature of MC grammar (Biq *et al.* 1996), in which the iconic principles have preeminence over the abstract ones in the symbolization process (Hsieh 1989). The MC word order regulating principles considered in this paper (cf. Section 2) are all extralinguistic iconic principles. The meaning of iconicity we refer to in the present study is, therefore, one in which content motivates a form, in which the external or the conceptual world is the model and the linguistic phenomenon is the icon (cf. Greenberg 1995). Consequently, the term 'motivation' refers to an 'external' one, i.e., based on the physical and socio-cultural reality in which the speaker is embodied. The concept of embodiment refers "to understanding the role of an agent's own body in its everyday, situated cognition" (Gibbs 2006: 1). It is a central assumption for theories about CL, which see language as not distinguished from general cognition. These considerations imply multiple repercussions for L2 instruction, which will be discussed in more detail in Section 3.

2. Mandarin Chinese word order iconic principles

2.1 The Principle of Temporal Sequence (PTS)

The PTS is a principle based on the linear order proposed by Tai (1985) as independently motivated. It states, "the relative word order between syntactic units is determined by the temporal order of the states which they represent in the conceptual world" (Tai 1985: 50). In other words, the sequence in the surface structure of linguistic units corresponds to the series of events in the phenomenological or conceptual world. It is a manifestation of what Langacker (2008: 80) calls 'temporal iconicity', in which "correspondence between conceiving of events in a given order and the order in which they are conceptualized and linguistically described by the speaker". Although it is a principle with universal characteristics, according to Tai (1985), this principle would be able to motivate the behavior of the majority of syntactic categories of MC, so much so that a violation of it would result in an agrammatical syntactic order.

In (1) (cf. Tai 1985: 54), it can be observed that if the order of the two events *zuò gōnggòngqìchē* 坐公共汽车 'to take the bus' and *dào zhèr* 到这儿 'to arrive here' is reversed this is also reflected in the meaning of the sentence. In (1a), the subject or topic *tā* 他 'he' first takes the bus and then arrives at the destination *zhèr*, whereas in (1b), the VP *dào zhèr* occurs temporally before the action of riding the bus.

(1) a. 他 坐 公共汽车 到 这儿
Tā zuò gōnggòngqìchē dào zhèr
he ride bus arrive here
'He came by bus'
'È arrivato qui prendendo l'autobus'

b. 他 到 这儿 坐 公共汽车
Tā dào zhèr zuò gōnggòngqìchē

he arrive here ride bus
'He came here to ride in a bus'
'È arrivato qui per prendere l'autobus'

The translations in (1a) point out that English and Italian are not required to respect this order, and the PTS can therefore be violated while still producing a grammatical sentence.
In addition to Tai's work, various studies have attempted to demonstrate the validity of this principle in word order analysis (Ho 1993; Hu 1995; Huang 2013; Jiang 2009, 2017; Loar 2011; Morbiato 2017; Zhang 1995, 2019). The literature generally agrees that the sentence internal constituents regulated by PTS can be verbal, prepositional (co-verbs), resultative, and, albeit limited to some types, adverbial (Loar 2011; Morbiato 2017). The PTS would also regulate the order between the main and subordinate clauses in the complex sentence: in MC, in the unmarked sentence, the subordinate clause typically precedes the principal clause (Loar 2011). This can be motivated in terms of PTS since the subordinate is processed conceptually earlier than the principal (Ho 1993; Jiang 2009) or otherwise delineates its frame of validity (Morbiato 2019).
Ho (1993) is the first scholar to explicitly extend the motivational range of PTS to the placement of preverbal temporal expressions. According to Ho, their order is "determined by their temporal relation to the verb" (1993: 153), and punctual time expressions thus occupy a preverbal position since they provide the temporal scope in which the action expressed in the predicate takes place. Ho does not elaborate further on his interpretation, which would seem to resemble the motivation given by Tai (1985) in the elaboration of another principle, the Principle of Temporal Scope (PTSC), according to which "if the conceptual state represented by a syntactic unit X falls within the temporal scope of the conceptual state represented by a syntactic unit Y, then the word order is YX". Jiang (2009) elaborates further on the remarks introduced by Ho by stating that the logic behind the preverbal positioning of time expressions is reflected in the fact that they represent "a temporal state that happens earlier than an action or marks the beginning of an action". Look, for example, at (2), taken from Tai (1985: 61):

(2) a. 他昨天走了
Tā zuótiān zǒu le
b. 昨天他走了
Zuótiān tā zǒu le
c. *他走了昨天
Tā zǒu le zuótiān
'He left yesterday.'
'Se ne è andato ieri.'

The ungrammaticality of (2c) according to PTSC would derive from the fact that the action expressed by the verb falls within the temporal scope of the time expression *zuótiān* 昨天 'yesterday' and would therefore go after it. Reasoning instead from the perspective of PTS, *zuótiān* is a temporal state whose inception precedes the action expressed by the verb. A visual representation of the action range of *zuótiān* at the temporal level, adapted from Huang (2013), is proposed in Figure 1.

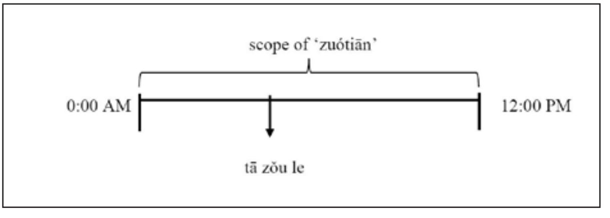


Figure 1. Temporal relations of '*tā zuótiān zǒu le*'.

In her cognitive-functional taxonomy of MC word order errors, Jiang (2009) thus does not explicitly include PTSC as a foundational principle and prefers to motivate the placement of temporal expressions relative to the verb according to PTS[2]. Instead, the relative order of time expressions is traced by Jiang to the more generic scope of conceptualization evoked by the Whole-Before-Part principle (cf. Section 2.2), and PTSC is considered as one of its sub-principles. Given these considerations, I agree with Jiang that the placement of time expres-

[2] More specifically, Jiang considers it as a sub-principle of the Whole-Before-Part (cf. General-Preceding-Particular, Section 2.2) in motivating the relative order of time expressions (she uses the term 'time expressions' in her taxonomy).

sions in relation to the verb can only be motivated for pedagogical purposes through PTS and that the relative order of time expressions is instead traceable to the principle of General-Preceding-Particular (cf. Section 2.2).

During instruction, the simplicity of the conceptual scheme underlying PTS can be represented by employing schematic visual diagrams consistent with the concept underlying dual coding strategies (cf. Section 3), as in the case of Figure 1 previously illustrated or in Figures 2 and 3 below, adapted from Huang (2013), where two different conceptualizations relevant to the adverbial or complement nature of the locative syntagm *zài chuán shang* 在船上 'on the boat' are represented, defining its preverbal or post-verbal placement in temporal frames (TF).

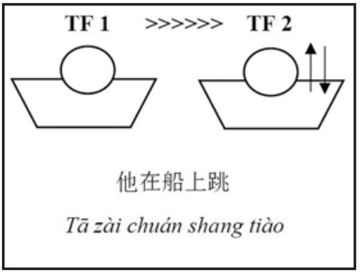


Figure 2. 'He is jumping on the boat'.

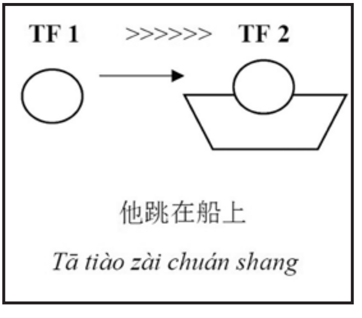


Figure 3. 'He jumped on the boat'.

2.2 The Principle of General-Preceding-Particular (GPP)

It is a principle that should be understood using the notion of logical scope (e.g., Morbiato 2020; Tai 1985), and therefore includes several labels with which the literature over the years has described principles of organization in frames and based on an image containment schema, such as the Principle of Whole-Before-Part (WBP, cf. Tai 1989), The Principle of Contained-Before-Contained (CBC, cf. Hu 1995) and The Principle of Temporal Scope (PTSC, cf. Tai 1985) previously discussed. These principles regulate word order patterns that can be traced back to the broader logical diagram expressed by the notion of 'general' and 'particular'. In the words of Ho (1993: 161), "constituents representing a global scope (general or whole) should precede those that represent a smaller scope (particular or specific)". I propose using the GPP label to define this conceptual scheme within a pedagogical grammar because it is the most logically inclusive and lends itself to including conceptualizations derived from principles such as WBP, CBC, and PTSC[3].

(3) a. 台湾，台北，罗斯福路三段，99号
Táiwān, Táiběi, Luósīfú lù Sān Duàn, Jiǔshíjiǔ Hào
'Taiwan, Taipei, Roosevelt Road, Section 3, No. 99'
b. 2334, North High Street, Columbus, Ohio, U.S.A.
c. Largo Agostino Gemelli 1, Milano, MI, Italia.

The most obvious manifestation of the iconic motivation of the GPP is shown in the relative order of spatial expressions as demonstrated by Tai (1989: 202) in (3a), where the most general entity, 'Taiwan', is placed first and the most particular one, the house number '99', is at the end of the address.

Differently from PTS, GPP is a language-specific principle. In the organization of spatial expressions, English presents an opposite pattern, 'particular-before-general' (or 'part-before-whole'), as demonstrated in (3b), and Italian resumes this arrangement, except for a single constituent ('1', the number) as shown in (3c). This diagrammatic difference is given by the nature of the iconic motivation of the principle, which is perceptual and experiential in nature and grounded in a culture-specific perception and con-

[3] Cf. the definitions by Hu (1995), Jiang (2009) and Morbiato (2020).

ceptualization of spatial relations (Morbiato 2017). It is a principle that reflects one of the theoretical tenets underlying CL, consistent with a weak position of linguistic relativism, namely that differences in word order patterns are partially due to language-specific conceptualizations resulting from different construal of the real world in different cultural experiences (Tai 2005: 563). The scope of iconic motivation of the GPP is not limited to the relative order of spatial expressions but also extends to various phrase-level constructions (e.g., percentages, fractions, partitive relations), clause-level constructions (e.g., preverbal spatial and locative constructions, double topic constructions, quantifier-float constructions) and to existential sentences that develop from a reference object or place (e.g. existential 'you' sentence). Structures such as the latter, exemplified by Hu (1995: 42) in (4), present particular difficulties for English and Italian-speaking learners whose L1 prefers a different spatial mapping.

(4)	桌子	上	放	了	一	本	书
	Zhuōzi	*shang*	*fàng*	*le*	*yī*	*běn*	*shū.*
	desk	top	put	PRT	one	CL	book

'There is a book on the desk'
'C'è un libro sul banco'

A further addition to the explanatory capabilities of the principle is given by Loar (2011), who traces the internal order of positional compound words (General Noun + Position Word) to the conceptualization of the containment schema of the Principle of Whole-Before-Part, which, as I have previously discussed, it is nothing more than a manifestation of the GPP. According to Loar (2011: 101), the spatial relationship between the two components forming the compound is grounded in the concepts of coincidence, interiority, superiority, inferiority, anteriority, posteriority, laterality, and opposition. In compound words such as (5), the superiority relation reflected by the containment scheme is rather concrete and transparent if one considers *shang* 上 'top' as a 'part' of the desk seen as 'whole' (i.e., its surface).

(5) 桌子　　上
Zhuōzi　　*shang*
desk　　top
'Sopra il banco'
'On the desk'

The schema takes on more abstract characteristics if we consider *shang* as the space projected from the table surface upward. In this case, the table NP is seen as the referent object that sets a reference frame for the position word (PW), consistent with the containment scheme evoked by the GPP. The conceptualization of the GPP is also extensible to the ordering of time expressions. The logic underlying the extension of a principle based on an image schema that has more properly to do with space is consistent with the view of the spatiotemporal-existential cline as a cognitive universal (Givón 1979). While in MC, as demonstrated in (6a), the order of temporal expressions conforms with the GPP, this is not the case in British English (6b) and Italian (6c), where constituents with a minor scope generally precede those with a major scope.

(6) a. 1980年，12月，22日，上午，10点钟
1980 nián, 12 yuè, 22 rì, shàngwǔ, 10 diǎnzhōng
b. 10 am, day 22, December, 1980
c. 10:00 di mattina, giorno 22, Dicembre, 1980

As with the PTS, the image schemas evoked by the GPP conceptualizations can be conveniently reproduced with visual diagrams that can be easily used in a pedagogical context.
The containment schema is represented in Figure 4 (own work), in which the scope of element X also includes those of entities Y and Z. In contrast, Y is superordinate to Z, which is the most 'particular' entity. However, representations employing two-dimensional spatial diagrams are also easily interpretable, such as the one in Figure 5 (own work), where the laterality relationship between the NP and the *yóubian* 右边 PW 'on the right' is represented by a selection of a single and 'particular' projection among the possible ones projected by the schematic surface of the 'general' reference object 'NP'.

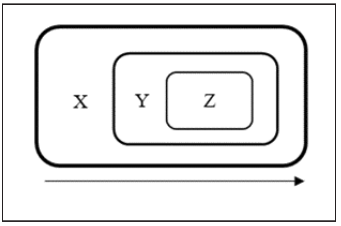


Figure 4. Visual diagram of containment.

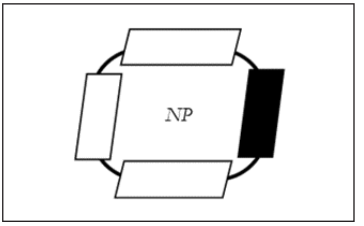


Figure 5. Visual diagram of 'NP + *yóubian*'.

2.3 The Principle of Semantic Proximity (SP)

The principle is based on one of the manifestations of diagrammatic iconicity that has been most investigated in the literature, that of 'distance' (*distance motivation*, e.g., Haiman 1985). This is an application to the MC word order of the *principle of proximity* enunciated by Givón (1991: 89): "entities that are closer together functionally, conceptually or cognitively will be placed together at the code level, i.e., temporally or spatially". One of the first applications of this principle to the study of MC can be found in the work of Tai (1993), who notes that the presence or absence of the structural particle *de* in linking modifier and head in nominal phrases reflects their semantic closeness: the lack of the particle would express a closer or more intimate link between head and modifier. The notion is taken up extensively by the studies by Lu (1998, 2009), who proposed the label of *Semantic Proximity* (SP). SP, although universal, is believed to have a higher motivational power in a language like MC, where the match-up between syntax and semantics is relatively more transparent (Tai 1993, 2005) and where word order plays a strong role in determining syntactic relations. Loar (2011) emphasizes several manifestations of the principle, including its pervasiveness in motivating the position of preverbal adverbials and the internal word order of the MC noun phrase, stating that the structural distance of a modifier from its head reflects in full its semantic proximity (proximity based on features such as inherence, relevance, intrinsicality, essentiality, and cruciality). (7) represents the general order of bare adjective modifiers proposed by Lu (1998: 65). The semantic proximity to the noun is motivated by their constancy. For example, in (8), the material *mù* 木 'wood' is considered to be a more constant property than the color *bái* 白 'white', the color is more constant than the size *xiǎo* 小 'small', and the size is more constant than the more transient modifier indicating 'age', *xīn* 新 'new'.

(7) age > size > color > origin > material > manner > function > noun

(8) 新小白木房
xīn xiǎo bái mù fáng
age > size > color > material > N
new small white wood house
'A new small white wood house'
'Una nuova e piccola casa bianca di legno'

The translations in (8) show that the relative order of modifiers in MC is also valid in English and Italian. Loar's study (2011) shows how the principle can motivate the order of different types of adverbials and complements (e.g., the order of action measure complement concerning verb-object compounds), which may be challenging to comprehend in the English or in the Italian language. More specifically, in the case of the iconic motivations of different types of adverbials that do not present proper temporal aspects (e.g., adverbials of modality, focus, emphasis, etc.), Loar states that it is necessary to resort to principles such as SP since the PTS cannot hold any explanatory power in these instances. A more recent study by Zhang (2019), in agreement with what has already been observed by Lu (1998), comes to the conclusion that SP plays sometimes a more significant role than PTS in motivating word order, also because it is capable of motivating particular constructions (e.g., the relative order of preverbal adverbials) in which the latter does not hold any explanatory power.

2.4 The Empathetic Principle (EP)

The designation of the principle is given by Ho (1993) and it is afterward adopted in Jiang's (2009) taxonomy, which includes the principle in a third domain apart from the cognitive and functional ones that she refers to as 'the sociocultural domain'. This is a macro-principle whose structuring stems from Kuno's (1987) studies on empathy and the notion of familiarity as understood by Siewierska (1988). Its motivational factor is grounded in the manifestation of two main iconic principles, the 'principle of sociocentrism' and the 'principle of egocentrism'. Ho (1993: 167) summarizes the logic behind the principle: "speakers are preferentially more interested in themselves than in others, more interested in what is socially, spatially and temporally nearer than what is socially, spatially and temporally farther away". Egodeictic iconicity is, for example, manifested by the fact that a sentence such as 'My sister is your daughter' is processed faster than a sentence such as 'Your daughter is my sister', because the speakers tend to mention themselves first or, on this specific instance, a person with whom they have a more intimate and personal relationship[4]. On the other hand, the order of the juxtaposed kinship terms in (9) is an example of sociocentric iconicity. Chinese society has strong Confucian and patriarchal roots, which is reflected in (9) by the precedence of the male kinship term *fùqin* 父亲 'father' over the female term *mǔqin* 母亲 'mother'.

(9) 父亲 母亲
Fùqin *mǔqin*
father mother
'Father and mother.'
'Padre e madre.'

While the egodeictic principle could be considered a universal principle, the sociocentric principle is specific because it reflects different social structures. For example, Panther (2021) points out that in English, male kinship terms are more frequently ordered in the first position than female ones (e.g., 'husband and wife', 'brother and sister', 'boy and girl'). This order would reflect the patriarchal nature of the Anglo-Saxon sociocultural model. In the Italian language, spoken in a country with evident patriarchal roots, the order of terms is sometimes interchangeable (as in English), e.g., *papà e mamma* / *mamma e papà* 'dad and mum/mum and dad', except for some juxtaposition with a more marked lexicalization (e.g., brother and sister, husband and wife)[5]. A reversal in the juxtaposition of the two terms still produces appropriate sentences. Note, how-

4 See Ho (1993) for other manifestations of this iconic principle.

5 Note that in some instances other types of motivations, often of prosodic nature, may concur in shaping the order of the juxtaposition (e.g., 'mum and dad').

ever, that the literature considers the asymmetry at the base of this principle to be more compulsory in MC, so much so that according to Jiang (2009), a reverse order of (9) would produce an inappropriate sentence. Given that this principle belongs to the broader extralinguistic area related to cognition (cf. Ho 1993), I do not consider it necessary to include the following principle in a separate domain, and it has therefore been included in the cognitive principles.

3. Learning and Pedagogical Perspectives

The centrality given by CL to linguistic motivation (and iconicity as one of its main manifestations) is one of the most promising features from the point of view of L2 learning and instruction, mainly because of its strong explanatory power, exploitable during classroom teaching and for the creation of pedagogical grammars aimed at promoting in teachers and learners a process of 'awareness-raising' of the target language structures (Taylor 1993).

Several scholars have tried to apply or suggest the use of CL's theories and linguistic motivation to second language learning and teaching (e.g., Achard 1997; Achard and Niemeier 2004; Boers and Lindstromberg 2006; De Knop and De Rycker 2008; Hijazo-Gascón and Llopis-García 2019; Holme 2009; Langacker 2001, 2008; Littlemore 2009; Pütz 2007; Pütz *et al.* 2001; Taylor 1993; Tyler 2008). Liu and Tsai (2021: 543) point out that a widespread belief in the literature is that the 'arbitrary' view of language given by formalistic theories is reflected in the production of equally arbitrary rules used in teaching strategies: this would lead the learner to an "overreliance on memorization, making language learning uninteresting and likely less effective". Boers and Lindstromberg (2006: 305) argue that "presenting segments of language as motivated is likely to be beneficial for learners as it may enhance comprehension, retention, cultural awareness and positive affect". Regarding comprehension, the assumption is based on the fact that learners are more likely to understand linguistic structures that reflect the phenomenological world they experience rather than an arbitrary explanation based on abstract mental rules. As Lakoff (1987: 346) states, "it is easier to learn something that is motivated rather than something that is arbitrary. It is also easier to remember and use motivated knowledge than arbitrary knowledge". Considering grammar as semantically motivated would thus add a step to the language teaching process: in the illustration of grammar rules in traditional instruction, given the arbitrary view of morphosyntax, the step of comprehension is often glossed over, moving directly from rule discovery to its application (Daloiso and Pascual 2017). Concerning long-term memorization (retention), Boers and Lindstromberg (2006) refer to the concept of *semantic* and *structural elaboration* as discussed in Barcroft (2002). These two mental operations are related respectively to the evaluation of meaning and form of linguistic items. According to Barcroft (2002: 324), semantic elaboration "can lead to greater recall and recognition of items than does structural elaboration". Mapping form and meaning through motivation leads to a semantic elaboration that should be capable of fostering deep processing, i.e., to stimulate the learner's mental operations of memorization at a deeper level, thus facilitating learning. The activation of this level of depth is studied, for example, in the theory of *dual coding* (cf. Clark and Paivio 1991), according to which a structuring of the input that stimulates a figurative representation of its semantic content is helpful in facilitating recall and therefore retention. In other words, inducing a reflection on the motivation of a connection between form and meaning could favor learning, bringing learners to process information at a 'deeper' level, thus increasing the possibility that the information remains imprinted in long-term memory and facilitating its acquisition. It is clear that the deep encoding that favors memorization pertains in fact to the semantic level more than to the more strictly syntactic one. Although dual coding theory and many other studies on the role of linguistic motivation have been concerned mainly with vocabulary and semantics, I believe that marked iconicity of MC word order would fit well to be applied to this strand of theory and approaches (cf. Section 2).

Further interesting implications from a pedagogical perspective lie in the nature of the motivations behind a language segment, which may reflect universal conceptualizations or culture-specific ones (cf. Tyler 2012). A deeper metalinguistic understanding of culture-specific motivations can help not only in increasing the learner's cultural awareness of the sociocultural context in which the target language is spoken but also in defining teaching strategies to cope with so-called conceptual errors resulting from the transposition (calque) of the L1 conceptual system to the L2 (cf. Danesi and Grieve 2010).

Finally, talking about the benefits of the "positive affect" previously mentioned by Boers and Lindstromberg (2006), it seems that linguistic motivation is able, indirectly, to positively influence also motivation (now considered in psychological terms), making the learning experience more stimulating and encouraging the learner to continue in the longer term the study of the target language. Learners' appreciation has been suggested by a series of satisfaction surveys administered to learners following linguistically motivated teaching (Boers and Lindstromberg 2006).

I conclude this section by recalling, however, a fundamental instance: while it is true that the strand of studies ascribed to CL agrees in seeing linguistic motivation as an explanatory concept (Radden and Panther 2011), it does not always hold an explanatory value that is easily understood by the learner for learning purposes. According to Boers and Lindstromberg (2006): "sometimes the motivation of meaning is pretty straightforward, but at other times it may be quite abstract and may even come across to learners [as] far-fetched". Meunier (2008) even underlines how, at times, it may even be the learners who expect a 'formalistic' (and therefore, arbitrary) illustration of the rules of grammar of the target language, because they are not used to a conceptually motivated teaching of L2. The motivation to which I appeal in this study, therefore, holds a useful explanatory power for structuring pedagogical interventions aimed at the learners, assuming first of all that, unlike the professional linguist, they hold limited metalinguistic knowledge, even in their L1, and they are not willing to make a considerable cognitive effort because it is seen as unnecessary or beyond their learning habits. For this reason, I believe it is useful to look for the help of linguistic motivation not only in the illustration of less salient grammatical rules or in those considered to be more difficult to acquire but, whenever possible, to start presenting grammar as motivated starting from the most elementary levels of competence, accompanying the learning process in parallel with the learner's acquisitional sequences. In this way, it is possible to accustom the learner to 'motivated' teaching from the beginning of his L2 learning process.

Within formal learning contexts, the introduction and illustration of language segments as motivated would occur predominantly in an explicit manner, during those interventions focused on the linguistic form aimed at stimulating and fostering the learner's language awareness. On some occasions, implicit teaching would not be sufficient to illustrate those linguistic structures based on different conceptual and cultural systems: instructional methodologies in which grammar is taught using an explicit approach, which motivates the conceptual steps that lead the L2 to be constructed differently from one's L1, would therefore have an explanatory power useful to facilitate learning (cf. Ellis and Robinson 2008; Yu and Wang 2018). However, unlike inductive or deductive approaches based on the explicit teaching of individual grammatical rules, in the ACL-inspired approaches what is explicitly illustrated by the language teacher are the conceptual and constructional schemas, each of which holds explanatory power with great grammatical generalizability. As for the iconic principles of MC word order illustrated in Section 3, it could be assumed that a recollection of the conceptual diagrams underlying the principles immediately preceding the production of the target language, during the planning phase, could help the learner to prevent possible word order errors, and therefore bring more tangible benefits from the point of view of appropriateness at the level of language production. In practice, this is facilitated by the fact that a few iconic diagrams dominate several syntactic patterns, and their evocation is therefore cognitively less expensive than the activation of a plurality of arbitrary rules. Their conceptualization could also be easily evoked during the post-production phases of monitoring and repair, particularly in written production. I have also shown that these principles' iconic motivations are universal (e.g., PTS) and language/culture-specific (e.g., GPP). If it is conceivable that the universal ones are easy to understand in the field of L2 learning because they are already part of the cognitive-conceptual system connected to the L1, the application of WO patterns governed by those language/culture-specific ones indicates the need for a repositioning of the conceptual apparatus towards that connected to the L2. If we start from the critical assumption for CL that the target language system is influenced by the cultural environment in which the native speaker is embodied, explicitly teaching these principles would foster the learner's metalinguistic and cultural awareness.

Several are the theoretical studies that have hypothesized a benefit in applying linguistic motivation in L2 teaching and learning; scarcer are the empirical ones, which have, however, generally found positive effects[6]. This type of study, applied to almost exclusively European target languages (mainly English, see Taylor and Huang 2018), has focused primarily on the conceptual teaching of vocabulary (idioms, phrasal verbs, prepositions) and morphemes (aspect and tense). Concerning MC, a pioneer series of works on the subject are collected in the 2011 volume *Rènzhī yǔfǎ yǔ duìwài Hànyǔ jiàoxué lùnjí* 认知语法与对外汉语教学论集 'Cognitive Grammar and the teaching of Chinese as L2' edited by Cui Xiliang. A review of acquisitional studies conducted on MC was more recently presented by Yu and Wang (2018). This study displays the expansion of theoretical studies about the potential of CL applied to MC pedagogy while highlighting the scarcity of empirical studies, primarily focused on categorization, construal, conceptualization, classifiers, figurative language, and specific constructions (see also Shu *et al.* 2019). On the other hand, studies on syntactic iconicity with clear implications for acquisition and learning are still scarce, although this is a line of research with great pedagogical potential. In addition to the work of Loar (2011) and the taxonomy of Jiang (2009), I recall the more recent study of Morbiato (2017), which is the first one to consider Italian-speaking learners and the different conceptualizations in the process of meaning construction and discourse organization between English, Italian, and MC.
To the best of our knowledge, empirical studies on the effects of an instruction motivated through word order iconic principles are still non-existent, if we exclude a quasi-experimental attempt by Tucci (2021), a cross-sectional investigation based on Italian language learners. In the study, an experimental group instructed on the role of PTS in the motivation of the preverbal and postverbal placement of the locative phrase '*zài* 在 + NP', outperformed in a post-treatment written test a control group instructed following the leading textbooks used in Italian Secondary Schools.

6 For a comprehensive and up-to-date review of studies inspired by CL, see Liu and Tsai (2021).

4. Conclusion

This paper sought to present the pedagogical potential of the cognitive-functional approach to MC word order by tracing the iconicity of some of its founding principles to the concept of linguistic motivation as understood in ACL. The review presented in Section 2 suggested that the motivations evoked by the principles are capable of holding a high explanatory power, so that their application in language pedagogy could replace the memorization of arbitrary rules, or at least could accompany their illustration. In the words of Loar (2011: xviii), "rules are easier to apply if one knows about the general principles upon which they are based".

The power of these principles in language pedagogy is thus primarily based on the fact that they are easy to understand and remember (Morbiato 2019) and hold a high explanatory value (Jiang 2009). This conceptual simplicity is essential to overcome one of the critical issues on which several scholars active in the field of ACL have drawn attention (cf. Broccias 2008; Tyler 2012), namely, to what extent it is effective to present motivations subordinated to conceptualizations of high complexity within formal education settings, where for curricular reasons the time to devote to linguistic awareness activities may be limited. In Section 2, it was also suggested how such simplicity could be well combined with visual instruction techniques (e.g., use of diagrams or images) during guided learning.

It is true that without adequate longitudinal studies it is difficult to speculate the benefit that the conceptualizations evoked by the principles can bring to the process of acquisition (understood in its unconscious and long-term aspect) of word order patterns. A preliminary illustration of the aforementioned principles suggested that they hold a different explanatory power based on the word order patterns they can motivate and on the conceptualizations evoked by their iconicity. In the case of conceptualizations where the interface between a syntactic construction and the semantics of the principles reaches such an abstract level that it excessively raises the cognitive load for the learner, an arbitrary and cognitively less demanding presentation would perhaps be more effective in illustrating the language. To answer this question, it is necessary to start conducting targeted experimental research on individual principles applied to specific syntactic constructions.

In conclusion, the present review suggests that the iconic principles pertaining to the Chinese language functional cognitive framework may prove to be a valuable tool in testing the assumptions inherent in ACL, given the commonality between their theoretical underpinnings and the similarity of their areas of application, first among them L2 learning and pedagogy. The strong iconicity of the Chinese language would also make it possible to broaden the scope of ACL analysis to word order, an area still almost entirely unexplored in this theoretical framework.

References

Achard, M. (1997) "Cognitive Grammar and SLA Investigation". *Journal of Intensive English Studies*, 11: 157-176.

Achard, M. and S. Niemeier (2004) *Cognitive Linguistics, Language Acquisition, and Foreign Language Teaching*. Berlin/New York, De Gruyter.

Barcroft, J. (2002) "Semantic and Structural Elaboration in L2 Lexical Acquisition". *Language Learning*, 52(2): 323-363.

Biq, Y.-O., J.H.-Y. Tai and S.A. Thompson (1996) "Recent Developments in Functional Approaches to Chinese", in C.-T.J. Huang and A.Y.-H. Li (Eds.) *New Horizons in Chinese Linguistics*. Dordrecht, Kluwer Academic Publishers: 97-140.

Boers, F. and S. Lindstromberg (2006) "Cognitive Linguistic Applications in Second or Foreign Language Instruction: Rationale, Proposals, and Evaluation", in G. Kristiansen, M. Achard, R. Dirven, and F.J. Ruiz de Mendoza Ibanez (Eds.) *Cognitive Linguistics: Current Applications and Future Perspectives*. Berlino/New York, De Gruyter: 305-355.

Broccias, C. (2006) "Cognitive Linguistics Theories of Grammar and Grammar Teaching", in S. De Knop and T. De Rycker (Eds.) *Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven*. Berlin/New York, De Gruyter: 67-90.

Clark, J. and A. Paivio (1991) "Dual Coding Theory and Education". *Educational Psychology Review*, 3: 233-262.

Cui, X. 崔希亮 (2011) 认知语法与对外汉语教学论集. Beijing, Beijing Language and Culture University Press.

Daloiso, M. and Pascual, G.J. (2017) “Bisogni Linguistici Specifici e apprendimento della grammatica: Il potenziale glottodidattico della Linguistica Cognitiva”. *EL.LE*, 6(3): 363-384.

Danesi, M. and C. Grieve (2010) “Conceptual Error Theory and the Teaching of Italian”. *Italica*, 87(1): 1-20.

De Knop, S., and T. De Rycker (Eds.) (2008) *Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven*. Berlin/New York, De Gruyter.

Dirven, R. and M. Verspoor (2004) *Cognitive Exploration of Language and Linguistics*. Amsterdam/Philadelphia, John Benjamins.

Ellis, N.C. and P. Robinson (2008) “An Introduction to Cognitive Linguistics, Second Language Acquisition, and Language Instruction”, in P. Robinson and N.C. Ellis (Eds.) *Handbook of Cognitive Lingusitics and Second Language Acquisition*. New York/London, Routledge: 3-24.

Gibbs, R.W. (2006) *Embodiement and Cognitive Science*. Cambridge, Cambridge University Press.

Givón, T. (1979) *On Understanting Grammar*. New York, Academic Press.

--- (1991) *Syntax: A Functional Typological Introduction: 002*. Amsterdam, John Benjamins.

Greenberg, J. (1995) “On Language Internal Iconicity”, in M. Landsberg (Ed.) *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*. Berlin/New York, De Gruyter: 57-64.

Haiman, J. (1980) “Isomorphism and Motivation”. *Language*, 56(3): 515-540.

--- (1985) *Natural Syntax: Iconicity and Erosion*. Cambridge, Cambridge University Press.

Hijazo-Gascón, A. and R. Llopis-García (2019) “Applied Cognitive Linguistics and Foreign Language Learning. Introduction to the Special Issue”. *International Review of Applied Linguistics in Language Teaching*, 57(1): 1-20.

Ho, Y. (1993) *Aspects of Discourse Structure in Mandarin Chinese*. Lewiston/Queenston/Lampeter, Mellen University Press.

Holme, R. (2009) *Cognitive Linguistics and Language Teaching*. London, Palgrave Macmillan.

Hsieh, H. (1989) “Time and Imagery in Chinese”, in J.H-Y. Tai and

F.F.S. Hsueh (Eds.) *Functionalism and Chinese Grammar. Monograph Series of the Journal of the Chinese Language Teachers Association* 1: 45-94.

Hu, W. (1995) *Functional Perspectives and Chinese Word Order*. PhD Dissertation, Ohio State University.

Huang, L. 黄良程 (2013) "时间顺序原则与华文语序教学". *Huawen Jiaoxue yu Yanjiu*, 50(2): 36-42.

Krivochen, D.G. and L. Lackova (2020) "Iconicity in Syntax and the Architecture of Linguistic Theory". *Studies in Language*, 44(1): 95-131.

Kuno, S. (1987) *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago, University of Chicago Press.

Jiang, W. (2009) *Acquisition of Word Order in Chinese as a Foreign Language*. Berlin/New York, De Gruyter.

--- (2017) "Acquisition of Word Order in Chinese as a Foreign Language: Replication and Extension", in I. Kecskes and C. Sun (Eds.) *Key Issues in Chinese as a Second Language Research*. New York/London, Routledge: 162-187.

Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago/London: University of Chicago Press.

Langacker, R.W. (2001) "Cognitive Linguistics, Language Pedagogy, and the English Present Tense", in M. Putz, S. Niemeier and R. Dirven (Eds.) *Applied Cognitive Linguistics I: Theory and Language Acquisition*. Berlin/Boston, De Gruyter: 3-39.

--- (2008) "Cognitive Grammar as a Basis for Language Instruction", in R.P. Roberson and N. Ellis (Eds.) *Handbook of Cognitive Linguistics and Second Language Acquisition*. London, Routledge: 66-88.

Littlemore, J. (2009) *Applying Cognitive Linguistics to Second Language Learning and Teaching*. New York, Palgrave.

Liu, D. and T-H. Tsai (2021) "Cognitive Linguistics and Language Pedagogy", in X. Wen and J.R. Taylor (Eds.) *The Routledge Handbook of Cognitive Linguistics*. New York, Routledge: 543-555.

Loar, J.K. (2011) *Chinese Syntactic Grammar: Functional and Conceptual Principles*. New York, Peter Lang.

Lu, B. (1998) *Left-Right Asymmetries of Word Order Variation: A Functional Explanation*. PhD Dissertation. University of Southern California.

--- (2009) "The Placement of Chinese Adverbials Revisited: What Differentiates Chinese Word Order from Other SVO Languages", in J. Xing (Ed.) *Studies of Chinese Linguistics: Functional Approaches*. Hong Kong, Hong Kong University Press, 23-44.

Meunier, F. (2008) "Corpora, Cognition and Pedagogical Grammars: An Account of Convergences and Divergences", in S. De Knop and T. De Rycker (Eds.) *Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven*. Berlin/New York, De Gruyter: 91-120.

Ministero dell'Istruzione, dell'Università e della Ricerca (2016) Direzione generale per gli ordinamenti scolastici e per la valutazione del sistema nazionale di istruzione, Ufficio VIII - Internazionalizzazione del sistema educativo di istruzione e formazione, *Il sillabo della lingua cinese. Quadro di riferimento unitario per l'insegnamento della lingua cinese nella scuola secondaria di secondo grado.* Roma, MIUR.

Morbiato, A. (2017) "Information Encoding, Mandarin Chinese Word Order and CSLA: A Cognitive-Functional Account", in I. Kecskes (Ed.) *Explorations into Chinese as a Second Language*. Cham, Springer: 83-107.

--- (2019) "Conceptual Word Order Principles and Mandarin Chinese Grammar", in C. Shei (Ed.) *Routledge Handbook of Chinese Discourse Analysis*. Abingdon/New York, Routledge: 85-101.

--- (2020) "Cognitive-Functional Principles and Chinese Linear Order: The Containment Schema". *Cognitive Linguistic Studies*, 72: 307-333.

Panther, K. (2021) "Motivation", in X. Wen and J.R. Taylor (Eds.) *The Routledge Handbook of Cognitive Linguistics*. New York, Routledge: 268-296.

Peirce, C.S. (1998 [1902]) *Logic, Considered as Semeiotic*, in J. Ransdell (Ed.) *An Overview of Charles Peirce's Philosophical Logic, Constructed from MS L75, Version 1*, *Analytical Reconstruction*. Arisbe, The Peirce Gateway.

Pütz, M. (2007) "Cognitive Linguistics and Applied Linguistics", in D. Geeraerts and H. Cuyckens (Eds.) *The Oxford Handbook of Cognitive Linguistics*. Oxford, Oxford University Press: 1139-1159.

Pütz, M., S. Niemeier and R. Dirven (Eds.) (2001) *Applied Cognitive Linguistics II: Language Pedagogy*. Berlin/Boston, De Gruyter.

Radden, G. and K. Panther (2004) "Introduction: Reflections on Motivation", in G. Radden and K-U. Panther (Eds.) *Studies in Linguistic Motivation*. Berlin, De Gruyter: 1-46.

--- (2011) "Introduction: Reflections on Motivation Revisited", in K-U. Panther and G. Radden (Eds.) *Motivation in Grammar and the Lexicon*. Amsterdam, John Benjamins: 1-28.

Shu, D., H. Zhang and T. Li (2019) "General Introduction: 30 Years of Cognitive Linguistics in China", in D. Shu, H. Zhang and L. Zhang (Eds.) *Cognitive Linguistics and the Study of Chinese*. Amsterdam, John Benjamins: 7-44.

Siewierska, A. (1988) *Word Order Rules*. London, Croom Helm.

Tabakowska, E. (2009) "Iconicity", in F. Brisard, J. Ostman and J. Verscheuren (Eds.) *Grammar, Meaning and Pragmatics*. Amsterdam, John Benjamins: 129-145.

Tai, J.H-Y. (1985) "Temporal Sequence and Word Order in Chinese", in J. Haiman (Ed.) *Iconicity in Syntax*. Amsterdam, John Benjamins: 49-72.

--- (1989) "Towards a Cognition-Based Functional Grammar of Chinese", in J.H-Y. Tai and F.F.S. Hsueh (Eds.) *Functionalism and Chinese Grammar. Monograph Series of the Journal of the Chinese Language Teachers Association*. 1: 187-226.

--- (1993) "Iconicity: Motivations in Chinese Grammar", in M. Eid and G. Iverson (Eds.) *Principles and Prediction: The Analysis of Natural Language.* Amsterdam, John Benjamins: 153-174.

--- (2002) "Temporal Sequence in Chinese: A Rejoinder", in L. Lien and K. Chui (Eds.) *Form and Function: Linguistic Studies in Honor of Shuanfan Huang*. Taipei, Crane Publishing: 331-351.

--- (2005) "Conceptual Structure and Conceptualization in Chinese". *Language and Linguistics*, 6(4): 539-574.

Taylor, J.R. (1993) "Some Pedagogical Implications of Cognitive Linguistics", in R.A. Geiger and B. Rudzka-Ostyn (Eds.) *Conceptualizations and Mental Processing in Language*. Berlin/Boston, De Gruyter: 201-226.

--- (2002) *Cognitive Grammar*. Oxford, Oxford University Press.

--- (2004) "The Ecology of Constructions", in G. Radden and K-U. Panther (Eds.) *Studies in Linguistic Motivation*. Berlin, De Gruyter: 49-73.

Tucci, T. (2021) "Il Principle of Temporal Sequence nell'apprendimento del sintagma locativo ‘zài + luogo’: un’indagine preliminare su discenti italofoni”, in C. Romagnoli and S. Conti (Eds.) *La lingua cinese in Italia. Studi su didattica e acquisizione*. Roma, Roma TrE Press: 115-139.

Tyler, A. (2008) “Cognitive Linguistics and L2 Instruction”, in N. C. Ellis and P. Robinson (Eds.) *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York/London, Routledge: 456-488.

--- (2012) *Cognitive Linguistics and Second Language Learning: Theoretical Basics and Experimental Evidence*. London, Routledge.

Tyler, A., L. Huang and H. Jan (2018) *What is Applied Cognitive Linguistics? Answers from Current SLA Research*. Berlin/Boston, De Gruyter.

Yu, N. and B.P.Y. Wang (2018) “Cognitive Linguistics Approaches to Chinese Second Language Acquisition”, in C. Ke (Ed.) *The Routledge Handbook of Chinese Second Language Acquisition*. Abingdon/New York, Routledge: 31-47.

Zhang, M. (1995) “Iconicity and Word Order Change in Chinese”, in J. Camacho and L. Choueiri (Eds.) *The Proceedings of the Sixth North American Conference on Chinese Linguistics*, vol. II. GSIL, University of Southern California: 249-263.

--- (2019) “时间顺序原则与像似性的’所指困境’”. *Shijie Hanyu Jiaoxue*, 33(2): 166-188.

THE IDENTIFICATION AND COMMUNICATION OF EXPRESSIONS OF ANGER IN ITALIAN AND CHINESE USING EMOTIONAL SCRIPT*

Valeria Varriano

"L'Orientale" University of Naples

Serena Zuccheri

University of Bologna (Forlì Campus)

1. Introduction

The socio-pragmatic aspects regulating emotional communication are a neglected aspect in textbooks on Chinese language and culture targeting Italian-speaking learners. Correspondingly, we might anticipate a similar lack in textbooks on Italian language and culture targeting Chinese-speaking students. Yet, the correct use of expressions relating to emotional experiences is extremely important to avoid miscommunication. To address this gap, we have designed an experimental didactic activity focusing on anger as a primary emotion and inspired, in part, by Mrowa-Hopkins and Strambi (2005). The aim is to help Chinese and Italian language learners identify expressions of anger. We specifically target situations where existing emotional-lexicon rep-

---

* Although this paper was collaboratively produced, Sub-sections 2.2 and *3.2.2* were written by Valeria Varriano only and Sub-sections 2.1 and *3.2.1* by Serena Zuccheri only. Sections and Sub-sections 1, 2.3, 3.1 and 4 were written by both.

ertoires define an emotion (Varriano and Zuccheri 2022), but do not help teach how to communicate an emotional state.

This paper first reviews the experimental didactic activity we designed and simultaneously conducted among Italian and Chinese students from four universities ("L'Orientale" University of Naples, the University of Bologna, the Beijing University of International Business and Economics - UIBE, and the Zhejiang International Studies University - ZISU). Secondly, we focus on the identification, learning and communication of terms and expressions relating to anger through the compilation of emotional scripts. These scripts are based on the analysis of the verbal, non-verbal and paralinguistic elements observed in two television series, *L'amica geniale* (My Brilliant Friend) (2018-to present) produced in Italy, and *Yǐ jiārén zhī míng* 以家人之名 - *Go Ahead* (In the Name of the Family) (2020) in China.

The body of the paper is divided into three main sections. The first (Section 2) addresses the theoretical background underlying our experimental didactic activity. We specifically consider the reasons why we chose to study anger as an emotion from a psychological and linguistic point of view (Section 2.1). Secondly, we examine the concept of a 'script' as used in the fields of Psychology and Artificial Intelligence (AI) as well as linguistic research. We then present the script concept as used in our research activity (Section 2.2). Finally, we consider the Collectivism versus Individualism dimension within a culture. We see this dimension as forming the basis for a learner's recognition of the main cultural differences which may exist between China and Italy. We also consider the ongoing validity of the hypothesis that cultures differ in their display rules regarding expressions of anger (Section 2.3) as raised by Matsumoto *et al.* (2010) and as based on Hofstede *et al.*'s cultural dimensions model (2010). The second section (Section 3) considers the students involved in the didactic activity, the selected television series and the script framework which the students completed. We also discuss the questionnaire we provided to help guide the students in analysing their selected scenes. The provisional findings based on the scripts collected from the students are then reviewed and discussed.

2. Theoretical and methodological background

2.1 Anger: the most frequently externalised emotion to be kept under control

Interest in emotions as a component of language has led philosophers, anthropologists, neuroscientists, linguists, psychologists and many others to speculate since ancient times, given emotions "seem to come from the most untamed, animal part of ourselves" (Whatt Smith 2015: 23). This has also led to thinking about the abstract concepts emotions refer to and trying to answer the naturally elusive question, "If different people have different ways of conceptualising emotions, might they feel differently about them as well?" (*Ibid.*: 25). One of the main reasons we chose to analyse anger as an emotion in this context is that it conceptually brings together a diverse range of knowledge. Anger is omnipresent in our Western mythological, literary and religious traditions. It is similarly a major factor in various aspects of Chinese culture, philosophy, literature and traditional medicine.
As a primary emotion, there are widely recognised, established and universal bodily expressions of and physical reactions to anger (e.g., gnashing of teeth, faster heartbeat, etc.). However, it cannot be taken for granted that different cultures or languages, like Chinese and Italian, have the same linguistic-metaphorical conceptualisations of anger[1]. Since the 1980s, metaphor has, in fact, been viewed as a cognitive

---

1 Santangelo's research on emotions (2003) represents a pioneering effort in Chinese studies and involved in-depth, interdisciplinary, textual analysis of the perception of emotions in Chinese literary culture during the Ming (1368-1644) and Qing (1644-1911) dynasties. Bond's studies (1993) in the field of Psychology are also of particular relevance. These investigated how events giving rise to emotions are perceived in Chinese culture. Bond speculates that there are differences in perception between cultures, particularly in China and America, according to the mapping done by Hofstede *et al.* (2010). That is, perceptions of emotion in America are determined more by individualistic traits whereas Chinese culture is characteristically and contrastingly more hierarchical and collectivist in its interpersonal orientations. Thus, it is not useful to assume there is an established level of universality in reactions to a given emotion. Rather, upon considering the main aspects of Chinese culture (philosophy, literature, traditional medicine, etc.), it emerges that "what differs in Chinese culture is the lower frequency, intensity

as well as linguistic tool for organising abstract concepts around the tangible physiological reactions within human experience (Lakoff and Johnson 1980; Lakoff and Kövecses 1987; Johnson 1987). A conceptual metaphor for anger, such as the idea that 'ANGER IS THE HEAT OF A FLUID IN A CONTAINER' (Gevaert 2005; Kövecses 2010), exists in a number of languages. For example, in English there is the expression 'You make my blood boil' and its equivalent in Italian is the phrase '*ribollire di rabbia*' (be boiling with anger). In both cases, the idea is that the body is like a container of emotions and experiences an increase in temperature and pressure as the result of anger (defining a sensory-body experience).

This is a complex kind of metaphor, as Kövecses (1995) reminds us, which actually contains three sub-metaphors: (1) the body is like a container for emotions, (2) anger is like heat, and (3) emotions are like fluids. It is, therefore, really a set of conceptual metaphors that refers to the humoral theory of emotion and medicine, whose roots lie with Hippocrates (c. 460-375 BC). According to him, the body is a container for anger and the liquid inside the body, be it blood or bile, is affected by anger. Thus, the expression 'You make my blood boil' exists in English. Commensurately, in Italian there are the expressions '*farsi il sangue amaro*' (make your blood bitter) and '*rodersi il fegato dalla rabbia*' (let your liver be eaten away by anger). The general idea is that pressure in the body builds as the result of anger, causing its fluids to overheat, and so on. Some studies (King 1989; Yu 1995, 2002) show that metaphorical expressions of anger in Chinese are similar to the conceptual metaphors proposed by Lakoff and Kövecses (1987). However, in Chinese, the substance within the 'container' is construed in a relevantly and culturally distinct way. In Chinese, 'Anger is the hot gas in a container', echoing the concepts of *yīn* 阴 and *yáng* 阳 and of the five elements of traditional Chinese medicine (wood, fire, earth, metal, and water)[2]. Anger is also metaphorically conceptualised and apparent in metonymic connections made within Chinese culture and language (Kövecses 2000), e.g., the notion of having an excess of *qì*

and duration with which emotions are typically experienced" (Bond 1993: 245).

[2] In traditional Chinese medicine anger is related to the liver and wood element (Yu 1995).

气[3]. Kornacki (2001) analyses anger in terms of the Natural Semantic Metalanguage theoretical model proposed and extended by Wierzbicka and others. This analysis focuses on identifying five Chinese terms (*nù* 怒, *shēngqì* 生气, *nǎohuǒ* 恼火, *fèn* 愤, and *tǎoyàn* 讨厌) and their potential equivalence with the English words 'angry' and 'anger'. It describes cognitive scenarios in which such concepts either differ or are similar in both languages. Alternatively, Xu and Tao (2003) see anger as belonging to a sub-category of psychological emotions (*jīyú xīnli gǎnshòu de qínggǎn* 基于心里感受的情感) identified by 12 terms: *fènnù* 愤怒; *fènhèn* 愤恨; *jīfèn* 激愤; *shēngqì* 生气; *fènmèn* 愤懑; *fènkǎi* 愤慨; *fèn nù* 忿怒; *bēifèn* 悲愤; *wōhuǒ* 窝火; *bàonù* 暴怒; *bùpíng* 不平; *huǒ* 火. More recently, a study by Lee (2019) has developed a more highly articulated classification of primary emotions using previous taxonomies proposed by Plutchik (1980) and Turner (2000). This more recent classification considers the various intensities (high-medium-low) with which a primary emotion can be expressed and how it combines with other primary emotions. In the case of anger, this classification turns out to be linked to 34 nouns, verbs and adjectives.

From a psychological perspective, anger "is a central and prototypical emotion because in it – perhaps more than in other emotional states – it is possible to identify a clear functional origin, characteristic antecedents, constant expressive and physiological displays, and predictable tendencies to action" (D'Urso and Trentin 1998: 309). Anger is seen as a primitive emotion because it is detectable in humans from an early age. As an emotion, it is visceral, short-lived, animal-like and pathological in origin and triggered by numerous physical and psychological factors (*Ibid.*: 309-328). In cognitive psychology, psychophysiology and neuroscience, recent studies have confirmed that emotional states arise when a goal of fundamental importance to individual survival or well-being is either not achieved or compromised (Poggi 2008: 17). Similarly, Ortony *et al.* (1988), focusing on the cognitive aspect of emotions, divides them into three main groups: (1)

[3] In this regard, Yu (1995) asserts that anger is conceptualised as a destructive force capable of harming the angry subject and those in their vicinity in both Chinese and English.

emotions arising from reactions to events experienced in a positive or negative way, (2) emotions arising from approved or unapproved agents and (3) emotions arising from objects that may be pleasant or unpleasant to a person. Scherer *et al.* (1986) have conducted a related cross-cultural study of emotional experience and reactions in seven European countries and Israel. This research shows that in situations eliciting a primary emotion like anger, responses to it can be traced to factors including: (1) the failure (of friends, relatives and strangers) to conform to social norms and rules, (2) inconvenience or inappropriate rewards, and (3) failure to reach one's goals. A study by D'Urso (2001: 38-39) identifies four main common situations that generated anger: (1) perceiving an unpleasant situation of a physical or material nature (e.g., threats to personal integrity and material goods); (2) encountering disturbances or obstacles in one's activities (e.g., hindrances to the achievement of one's aims); (3) experiencing psychological frustration (e.g., unpleasant experiences, the interruption of or deprivation from pleasant experiences or suffering harm to one's image, self-esteem, dignity or that of loved ones); (4) injustices (e.g., either suffered or envisaged against oneself or others).

By defining anger as one of the most socially based of all emotions, Matsumoto *et al.* (1989) recognise that it is the most frequently externalised of all primary emotions. Furthermore, expressive reactions to anger, whether verbal or non-verbal, are also greater than the related physiological ones. Displays of anger as a negative emotion in any social context depends on the degree to which it is controlled by the established social norms and rules existing in every culture (Matsumoto *et al.* 2010). Cultures may thus differ in their display rules (Ekman and Friesen 1971) in terms of expressing anger. As a result, different cultural groups or languages involved in a communicative exchange may express and control anger in varying ways at diverse levels.

2.2 The script-based approach

The script concept (Schank and Abelson 1977; Fehr and Russell 1984; Shaver *et al.* 1987) has been developed in the fields of Psychology and Artificial Intelligence (AI) in order to define the shared, culturally specific knowledge that lets people think about, express and regulate emotions. According to Galati, the script model has been "particularly useful in explaining the internal structure of emotion [...] as a causative sequence of events rather than a list of components, capturing the dynamic and processual characteristic that makes it an event rather than a state" (1993: 171). The script model 'recounts' a cultural view of the distinctive phases of an emotional experience. That is, it reflects a procedural knowledge whereby we can distinguish three distinct stages comprised of (1) emotional 'antecedents' (the circumstances which elicit a certain type of emotion), (2) 'responses' (the representations of reactions related to the antecedents) and (3) 'controls' (the actions that moderate the emotional reaction).

From a linguistic point of view, the script-based approach (Kövecses 1986, 1990, 2000; Lakoff 1987; Wierzbicka 1994) has made it possible to conceptually evaluate emotions by adopting, on the one hand, a universalistic view of them, and, on the other, an awareness of the experiential nature of emotional concepts. This involves embracing the idea that emotional conceptualisation, rather than physiological reactions, may vary between different languages and cultures. Script-based semantics stems from Raskin (1985) who considered jokes as texts made up of two overlapping and opposing scripts. According to Raskin, a semantic script is a collection of concepts related to a particular task or event which are thus, causally linked. Scripts are then very useful in the semantic analysis of emotional expressions. Presuming there is a specific knowledge about emotions within a culture, it is essential to understand the emotional-knowledge structures these cultures may transmit. Such knowledge structures are currently referred to as emotional scripts because they conceive of emotional episodes as unfolding over time between two or more actants (Shaver *et al.* 1987; Fehr and Baldwin 1996; Fitness 1996, 2000). Furthermore, these scripts can be explained using the learned patterns which form part of the reality in which one lives. Understanding people's emotional scripts is important since the script model plays a decisive role by shaping and influencing perceptions, expectations, judgments and memories of emotional episodes as social cognitive research has shown (Smith 1995; Fitness 1996). Inspired by Corno and Pozzo (1991), in the context of our didactic activity, we refer to a 'script' as a framework created to describe a specific, ordered and predictable sequence of events which are part of a stereotypical emotional scenario. This scenario, in turn, is perceived as natural by second language learners. It is embedded within their memory and evoked in their thinking whenever a similar linguistic situation presents itself.

2.3 Cultural differences in displaying and expressing anger and the cultural dimensions model

Anger has been described as one of the most 'active' of primary emotions because it is expressed using a relatively wide and varied range of verbal expressions (threats, insults, swear words, violent expressions, interjective forms, short and imperative sentences and indirect forms of communication), non-verbal expressions (facial expressions, gestures and body orientation) and paralinguistic signals (voice volume, tone, etc.). Since the 1990s, many of Matsumoto's studies[4] have focused on investigating cultural similarities and differences in facial expressions of anger. The assumption is that if "cultures can produce universality in emotional responses (...), [they] can also differ in their display rules regarding angry expressions" (Matsumoto *et al.* 2010: 130-131). In exploring "how the expression and perception of anger is at once universal, biologically-based, yet also influenced by culture" (*Ibid.*: 134), Matsumoto *et al.* consider the scale of cultural values found in the cultural dimensions model developed by Hofstede *et al.* (2010)[5].

The aspect of this cultural model we focus on in our research is the

[4] The studies conducted by Matsumoto alone, or with other scholars, are numerous and we mention only a few here. Please refer to Matsumoto's website for a complete bibliography of his work (https://www.davidmatsumoto.com/, visited 2023/02/20).

[5] Hofstede became interested in national cultural differences, in particular value differences, in 1974, after gaining access to an extensive body of survey data on IBM branch-office employees from over 50 countries. He considered values as at the core part of a culture and understood them as "broad tendencies to prefer certain states of affairs over others", such as 'evil versus good', 'decent versus indecent' or 'forbidden versus permitted' (Hofstede *et al.* 2010: 9). The empirical findings based on the IBM data revealed comparable and measurable aspects of national cultures. Hofstede *et al.* (2010) refers to these aspects as 'cultural dimensions', defining them specifically as: Power Distance (from small to large); Collectivism versus Individualism; Femininity versus Masculinity; Uncertainty Avoidance (from weak to strong); Long-Term Orientation versus Short-Term Orientation; and Indulgence versus Restraint.

dimension of Collectivism versus Individualism, for which Italy and China have significantly different scores (Italy scores 76 and China 20 with respect to this dimension). From the perspective of our experimental didactic activity, the Collectivism versus Individualism dimension may first represent a starting point for understanding learner recognition of some main cultural differences between the two countries. Secondly, it may serve as the basis for the identification, learning and communication of anger in all its expressions.

Focusing on Collectivism versus Individualism involves considering that in individualist societies, where “everyone is expected to look after him or herself and his or her immediate family”, individual interests prevail over group interests; in contrast, in collectivist societies, group interests prevail over personal interest and “people from birth onward are integrated into strong, cohesive in-groups, which throughout people’s lifetime continue to protect them in exchange for unquestioning loyalty” (Hofstede *et al.* 2010: 92).

Some of the key differences between collectivist and individualist societies identified by Hofstede *et al.* in relation to general and family norms are of interest to our research. One important difference relates to being in harmony with one’s social environment. In most collectivist societies, maintaining this harmony is a fundamental virtue that extends to other spheres beyond the family. For this reason, open and direct discussions with other people are avoided as well as use of the word ‘no’, since saying no could mean starting a conflict. On the contrary, in individualist societies to speak one’s mind is a virtue and handling conflicts is part of everyday life both within and beyond the family.

The concepts of shame and guilt are strictly related to harmony. Indeed, in a collectivist culture, if a group member breaks a social rule, the remaining group members feel shame because the collective sense of duty has been betrayed. In contrast, in individualist societies, guilt prevails in a similar situation. The difference between the two concepts is that:

> Shame is social in nature, whereas guilt is individual; whether shame is felt depends on whether the infringement has become known by others. This becoming

> known is more of a source of shame than the infringement itself. Such is not the case for guilt, which is felt whether or not the misdeed is known by others (Hofstede *et al.* 2010: 110)[6].

Based on this theoretical model, Matsumoto *et al.* (2010) assume that members of collectivist cultures should show more positive emotions towards members of in-groups and more negative emotions towards members of out-groups with regard to emotional displays. Vice versa, members of individualistic cultures should show more negative emotions towards members of in-groups and more positive emotions towards those of out-groups.

3. The experimental didactic activity

3.1 The television series, recipients and script framework

We aim to answer several research questions based on the theoretical and methodological issues discussed above. Firstly, given that cultures can differ in their display rules regarding anger, we speculate on which didactic activities may help learners articulate the linguistic expressions (in terms of syntactic and lexical structures) that convey anger. A related issue is how to teach subjects to identify expressions of anger when lexical repertoires only define the emotion itself (Varriano and Zuccheri 2022), which alone is insufficient for teaching how to communicate an emotional state. Secondly, we question if Matsumoto's hypotheses are still valid in light of our comparisons of verbal, non-verbal and paralinguistic displays of anger in Chinese and Italian.

The term 'script' is a word borrowed from the world of cinema and television. As previously mentioned, a script describes what happens

[6] When speaking of shame, we must mention the concept of losing or saving face. These are deeply rooted in many collectivist groups and highly representative of Chinese culture. These concepts are linked, respectively, to either expressing individual honour and prestige (honouring someone/saving face, *gěi miànzi* 给面子) or the feeling of being humiliated (losing face, *diū miànzi* 丢面子).

in a given scene. In the context of our research, then, what is the aim of compiling scripts to identify expressions of anger? Will this help students understand the different ways of expressing anger in Italian versus Chinese culture? To put it another way, will compiling these scripts help address the current gap on the socio-pragmatic-cultural factors regulating the intercultural communication of emotion as evident in Chinese and Italian language textbooks?

Using a film or a television series to determine emotions is something Ekman *et al.* (1972) has already considered given that the interplay between characters creates easy-to-analyse interactions. Moreover, the emotions displayed and expressed in certain scenes are more intense than in real life (*Ibid.*). From our perspective, another fundamental reason to consider a television series for didactic purposes is related to the recent pandemic period. This event has forcibly distanced us from our affections and interests as well as freely travelling to other countries. As a result, intercultural exchanges have clearly decreased. At the same time, television and its fictional stories, despite their artificiality, stand as some kind of concrete and reliable example of intercultural and interpersonal interaction (Washburn 2001; Quaglio 2009: 148-149).

Therefore, the script-based approach, alongside considering the cultural dimension of Collectivism versus Individualism, constitute the theoretical and methodological framework on which the proposed didactic activity in a classroom setting is based.

The subjects involved in the activity were fifty-three Italian and Chinese MA students majoring in Chinese or Italian, respectively, as a second language at four universities: "L'Orientale" University of Naples (Department of Asian, African and Mediterranean Studies), the University of Bologna (Department of Interpreting and Translation), the Beijing University of International Business and Economics - UIBE (Department of Italian Language), and the Zhejiang International Studies University - ZISU (Department of Italian language). We selected two television series for consideration which were both well known in their home country and which recount realistic stories. Programmes based on fictional situations or distant cultural-historical ones far removed from the present, e.g., costume dramas, were excluded from consideration.

The Chinese series chosen for the Italian students was *Yī jiārén zhī míng - Go Ahead* (In the Name of the Family), a 46-episode drama series aired on Hunan TV in 2020. Watched by nearly 30,000 viewers, it came in 24th on the Top-Ranked Programme List on MyDramalist.com, a website for fans of Asian TV shows, dramas and films. It also had a 9/10 rating among 13,399 users. This Chinese series recounts the story of three children growing up in an atypical family. Their family is atypical because the three children, Ling Xiao, He Ziqiu, and Li Jianjian (the youngest) are not siblings. However, they end up in the same household after a series of unfortunate events. Ling Xiao, the eldest of the children, moves in with his parents, Ling Heping and Chen Ting, who reside in the same building where Li Jianjian and his father, Li Haichao, a widower, live. A resourceful neighbour introduces Li Haichao to the third child in the story, He Ziqiu, and his mother. He Ziqiu's mother is eventually forced to go to prison and thus, leaves her son in the care of Li Haichao. Meanwhile, Chen Ting, unable to get over her daughter's death, decides to abandon her son and husband (Ling Xiao and Ling Heping). As a result, the three children (Ling Xiao, He Ziqiu and Li Jianjian) find themselves living under the same roof with Li Haichao and Ling Heping, who afterwards become their main parental figures.

It was more difficult to choose an Italian series for the Chinese students. This is because Italian programmes purchased by Chinese broadcasters for Chinese television are rare if non-existent[7]. One exception is the first season of the series *L'amica geniale* (My Brilliant Friend). This series is based on the novel by the same name by Elena Ferrante and is one of the very few Italian television programmes to have achieved extraordinary international success. It is also highly appreciated in China, where it won the Magnolia Award for best foreign series at the 26th Shanghai Television Awards

[7] The National Radio and Television Administration (NRTA) is the mainland China agency in charge of controlling and administrating exchange activities between Chinese and foreign television companies and controls their content by regulation. As a result, there are restrictions on the number of successful foreign programmes which are accessible in China.

in 2020. The series tells the story of the close and controversial friendship between Elena 'Lenù' and Raffaella 'Lila'. The story is set in a poor neighbourhood on the outskirts of Naples and depicts the history of Italy over the second half of the 20th century, as its characters experience the historical, political, and social dynamics of the country at the time.

In both of the series, the characters establish their interpersonal relationships in social and family environments where numerous verbal and non-verbal conflicts occur. These programmes thus, constitute excellent source material for identifying anger and expressions of it.

The key focus of our analysis was to unravel the wide array of words and expressions used to convey anger. That is, we intended to catalogue these expressions as well as thoroughly understand the shades of meaning a second language learner attributes to them. By 'comprehending a meaning' we refer to the cognitive way mental images are formed. This idea spurred our decision to analyse texts within situations and contexts. Based on these, we infer the representative patterns that most strictly and rigorously correspond to how people construct meaning (Rumelhart 1980). To that end, the analysis carried out by our students relied on a syntagmatic framework based on a script (Schank and Abelson 1977). More specifically, the students defined all the factors involved in the identified scenes of anger. Here we relied on them exploring the terminology used to define the intimate relationships between actors and the antecedents to becoming angry. We asked our students to form four groups, watch the first four episodes of each TV series, and then identify different scenes that they could perceive as anger scenes (see Table 3). The aim was to come up with a description of all of the elements linked to anger and outline the emotional expressions of anger present in the selected scenes. These elements ranged from verbal and non-verbal signals (e.g., facial expressions, gestures, etc.) to paralinguistics (voice, intonation, etc.). The script framework used is described in Table 1 below.

Episode number
Initial timecode of the scene
Final timecode of the scene
Relationship of intimacy between the actors
Type of antecedent: according to D'Urso (2001) or Scherer *et al*. (1986)
Emotional expression through body orientation (towards the interlocutor; distance from the interlocutor; remaining in place, etc.) [Screenshots are highly suggested]
Emotional expression through non-verbal signals (facial expressions or gestures) [Screenshots are highly suggested]
Emotional expressions through paralinguistic signals (voice volume, elocution speed, pauses, hesitations, etc.)
Emotional expressions through verbal signals (interjections, swear words, insults, violent expressions, short and imperative sentences, forms of indirect communication, etc.)

Table 1. Script framework designed for the experimental didactic activity.

As shown in Table 1, we suggested students take screenshots to document facial expressions, gestures, and, where possible, the body orientation of subjects when expressing anger. When selecting the scenes, we also suggested ensuring that the expressive elements referred only to those aspects of speech that convey information to the receiver about the speaker's (or angry subject's) state during the verbal communication of anger. Finally, as shown in Table 2 below, we provided the students with a short, guiding questionnaire to help them analyse their selected scenes.

In the scenes you selected, anger seems:
(a) not expressed
(b) over-expressed
(c) expressed more verbally than non-verbally
(d) expressed less verbally (screenshots are recommended in the case of non-verbal expression through facial expressions, gestures, and body orientation)
If you answered (c), anger seems to be expressed through:
- terms/words
- metaphorical or metonymic expressions/phrases
- indirect, oblique or unspoken forms
- syntactic structures
- other (specify).

Table 2. Guiding questionnaire for the analysis of the scenes selected.

3.2 Preliminary analysis of data collected

3.2.1 Yǐ jiārén zhī míng - Go Ahead: *scripts and data*

The television series *Yǐ jiārén zhī míng - Go Ahead* is characterised by many tense situations that are well suited to analysing expressions of anger. The first important step in the analysis was to identify the scenes of anger. The results obtained reveal the interesting fact that the two Italian groups identified a different number of scenes (Table 3) in each episode.

Episode number	**Number of scenes selected (University of Bologna)**	**Number of scenes selected ("L'Orientale" University of Naples)**
One	8	12
Two	6	3
Three	7	4
Four	9	4

Table 3. Number of scenes selected from *Yǐ jiārén zhī míng - Go Ahead*.

This difference is partly due to the fact that the students in Naples identified all the scenes in which there is a verbal reproach, such as that stemming from a concerned parental figure in the story. An example of this can be found in the first episode when Li Haichao scolds his daughter because she has not finished eating the food on her plate (Ep. 01, 4:24-4:28). Based on Li and Meng's study (2007), the students stated that the selection of this, and several other similar scenes, makes sense since the Chinese written communicative mode, which is by nature reserved, implicit and indirect (*hánxù* 含蓄), may conceal an expression of anger or even a mild reproach. Not directly and explicitly expressing oneself can be traced back to ancient Chinese philosophy (Fang and Faure 2010). In any case, recognition of this led the students to indicate that a display of anger could also be concealed in the expression *dà rètiān de* 大热天的 (Ep. 1, 03:39) ("it's boiling today"). This is evident by their proposing translations for this phrase such as "Will you hurry up?", "Can't you see we're busy?" or "Will you stop wasting time?". In the case of identifying reprimands, the student observations may be limited, however. This is because a television series cannot be considered just as written discourse. Rather, it must also be seen as the result of "prefabricated orality" (Baños-Piñero and Chaume 2009). Chen Ting's declaring her annoyance over the hot day in addressing another character, however, is rightly identified by the students as a manifestation of anger expressed at a low level intensity (Plutchik 1980).

Figure 1. A facial expression usually not associable with anger.

The students' willingness to 'unmask' the anger behind some presumed indirect expressions contrasted with their frequent disregard for the character's facial expressions in the same scenes. According to the students, Chen Ting's response to little Li Jianjian (*Tā bù chī xuěgāo, xièxiè* 他不吃雪糕, 谢谢, He does not eat popsicles, thank you), when the little girl tells her that she bought a popsicle for her son, should also be understood as a display of anger despite the fact that her facial expression is not one of anger (Figure 1). The students, however, indicated there was a hidden meaning behind Chen Ting's words which should be understood as meaning "Go away. I don't want you to meet him or approach him".

These observations arose after students, acting on their own initiative, reviewed several related essays addressing intercultural communication and Hall's anthropological theory of high-context and low-context cultures (1976). However, these comments on hidden meaning are at odds with the communicative reality. As the image shows (Figure 1), the students miss the facial expressions accompanying what the characters say. This necessarily leads to some gaps in detecting expressions of anger through facial expressions. These facial expressions were considered by the students at the University of Bologna, but often discounted by the other student group. This group instead preferred to consider specific contexts, focusing only on subtitles, which led to underestimating the relevance of non-verbal expression. Focusing more on the subtitles than on the other components characterising scenes of anger, however, led the Italian students from Naples to make another observation. They noted that a discourse marker such as *wǒ gàosù nǐ* 我告诉你 'I tell you', which is widely used in numerous contexts, when found at the beginning of a sentence in the series generally signalled the start of an expression of anger (Ep. 2, 9:19).

In some cases, especially when the three children were fighting, their expressions of anger were not identified because cheerful music was playing in the background. It is as if the programme creators wanted to soften the anger between the characters with the music. This anger, however, is confirmed by other verbal and non-verbal expressive elements, especially the character's facial expressions and gestures (Ep. 3, 40:18-40:50). Still, the choice of accompanying music misled some students.

Another interesting aspect emerged in the identification of the scenes. Most students in the Naples group considered the frustration felt by a paternal figure (Ep. 4, 36:09) as another way to display anger, for example, when a parent recognised the child had grown up or was facing a series of problems.

While compiling the scripts, we also asked the students to evaluate the relationship of intimacy that existed between the actors in the selected scenes, which in most cases involved characters who were family or knew each other. As a result, the anger was, therefore, expressed mainly towards members of the in-groups. As previously mentioned, this contradicts Matsumoto *et al.*'s assertion that members of a collectivist culture should show more positive emotions towards members of in-groups and more negative emotions towards those of out-groups[8].

[8] In the four episodes examined, the students from Bologna University, for instance, identified 20 scenes in which anger is directed towards in-group members and 10 in which it is directed towards out-group members.

As asserted elsewhere (Varriano and Zuccheri 2022), very few occurrences refer directly to the emotive lexicon of anger (Xu and Tao 2003; Lee 2019). The words *shēngqì* 生气 and *qì* 气 ('get angry' and 'anger', respectively) are only used to describe one's state (not necessarily at the exact moment when anger is displayed) or another person's state (Ep. 2, 6:46). At the sentence level, students in both groups noted that interrogative, imperative, prohibitive[9] and declarative sentences can be vehicles for anger, with examples of these given in Table 4 below[10].

Sentence type	**Source text examples**
Interrogative sentences (57%)	(1) 你到家有半个小时吗 *Nǐ dào jiā yǒu bàn gè xiǎoshí ma* Have you stayed at home over half an hour? (2) 你是不是想太多了 *Nǐ shì bú shì xiǎng tài duō le* That is just your speculation
Imperative and prohibitive sentences (22%)	(1) 今天哪都别去 我告诉你 *Jīntiān nǎ dōu bié qù wǒ gàosù nǐ* Tell you what, you cannot go anywhere today (2) 闭嘴 *Bì zuǐ* Shut up
Declarative sentences (21%)	(1) 真是觉得搬到哪都一样 *Zhēnshi juédé bān dào nǎ dōu yīyàng* No matter where we move, it is all the same (2) 那是我身上掉下来的肉 *Nà shì wǒ shēnshang diào xiàlái de ròu* It is my daughter! I gave her birth!

Table 4. Sentence type examples in *Yǐ jiārén zhī míng - Go Ahead*.

[9] Our groups considered imperative and prohibitive sentences even if most of them appeared to be without the final modal particles *le* 了 and *ba* 吧 (Paternicò *et al.* 2021: 495-498).

[10] The official English subtitles accompany the examples given in the source text columns of the following tables in order for the non-Chinese reader to understand the meaning of the original lines.

In addition to nouns and phrases indicating offense, insults and idiomatic expression (Table 6), the use of the final modal particles *ya* 呀 and *a* 啊, and especially *ne* 呢, *ma* 吗 and *ba* 吧, in association with the interrogative substitutes, were also found to be vehicles of anger (Table 5).

Parts of speech	Source text examples
Nouns (6%)	(1) 坏人 *Huàirén* bad guy (2) 猪八戒 *Zhū Bājiè* Zhu Bajie[11] (3) 神经病 *Shénjīngbìng* You psycho (4) 小兔崽子 *xiǎo tù zǎizi* Little bastard
Adjectives/verbs (1%)	好气哦 *hǎo qì ò* It is so annoying!
Final modal particles (93%)	(1) 陈婷你过分了啊 *Chén Tíng nǐ guòfèn le* Chen Ting you have crossed the line (2) 是 我疯了 *Shì wǒ fēng le* Yes! I am crazy! (3) 大姐 你有病吧 *Dàjiě nǐ yǒu bìng ba* Hey, what is wrong with you?

Table 5. Parts of speech in *Yī jiārén zhī míng - Go Ahead*.

Finally, among the syntactic structures conveying anger, the students identified, albeit relatively infrequently, the constructions 'shì 是... *de* 的' and '*lián* 连... *dōu* 都/*bù* 不' and the resultative complement (Table 6).

11 Zhu Bajie is one of the main characters from the Wu Cheng'en (1510-1580) novel *Xīyóu jì* 西游记 (*A Journey to the West*). He is part human and part pig and extremely fat and ugly.

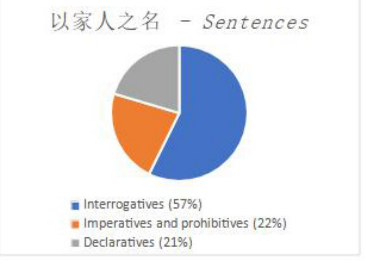
Syntactic and lexical structures	Source text examples
是…的 structure (37%)	(1) 都是他 是他害死的 *Dōu shì tā shì tā hài sǐ de* It is him! He killed his sister! (2) 都是你惯的 *Dōu shì nǐ guàn de* It is your fault (3) 李爸不是为了钱 才照顾我的 *Lǐ bà bù shì wèile qián cái zhàogù wǒ de* Dad Li does not take care of me for money
连…都/不 structure (13%)	(1) 平常 不给他外婆 打电话也就算了现在你看 连电 话都不接了 *Píngcháng bù gěi tā wàipó dǎ diànhuà yě jiù suànle xiànzài nǐ kàn lián diànhuà dōu bù jiē le* He does not call his grandma himself, but now he even refuses to answer the phone (2) 这么多年不管不顾的 连电话都不打一个 *Zhème duō nián bù guǎn bù gù de lián diànhuà dōu bù dǎ yī ge* She has been gone for years without even calling him
Idiomatic expressions (25%)[12]	(1) 你觉得我**莫名其妙**是吗 *Nǐ juédé wǒ mò míng qí miào shì ma* A mistery to you? (2) 你这**皮糙肉厚 没心没肺** *Nǐ zhè pí cāo ròu hòu méi xīn méi fèi* You are totally a heartless man! (3) 我真的不知道你 为什么要回来还要对我的家人**指手画脚** *Wǒ zhēn de bù zhīdào nǐ wèishéme yào huílái hái yào duì wǒ de jiārén zhǐ shǒu huà jiǎo* I do not know why you came back and judge my family

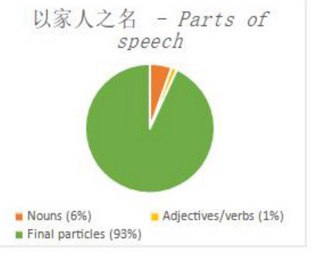
[12] Given the elusive nature of Chinese *chéngyǔ* 成语 (four-character idiomatic expressions), at this stage of the research, we decided to include them in this table, mainly considering the syntactic or lexical function they may perform within a sentence.

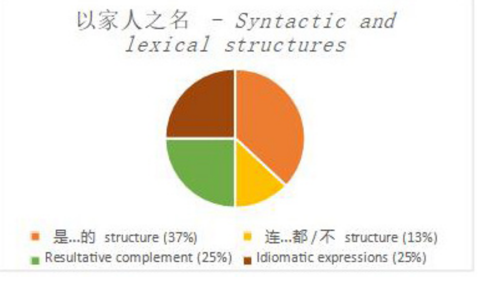
Resultative complement (25%)	(1) 走开 *Zǒukāi* Get out! (2) 再上我揍死你 *Zài shàng wǒ zòusǐ nǐ* I will beat you if you do it again! (3) 饿死你 *Èsǐ nǐ* Starve to death!

Table 6. Examples of syntactical and lexical structures in *Yǐ jiārén zhī míng - Go Ahead.*

We provide several graphs below that indicate the overall articulation of occurrences at the sentence level as well as in parts of speech and syntactic and lexical structures. These show the frequency percentages for the specific forms and expressions identified with anger in the source text by the two student groups. We note these graphs and percentages aim to provide some indicative rather than exhaustive insight.







3.2.2 L'amica geniale (My Brilliant Friend)*: scripts and data*

The first season of *L'amica geniale* intentionally shows a series of scenes of anger. As a matter of fact, Elena's narrative voice explains:

> As a child, I imagined tiny animals that came out of ponds, the abandoned train cars, the stones, the dust, and made their way into the water, the food, the air, making our mothers as angry as starving dogs. The men were always getting furious, but then they calmed down, whereas the women flew into a rage that had no limit and no end (S. 01, Ep. 1, 25:37-26:34)[13].

The incidents of anger recounted in the series deal with relationships within and beyond the family environment. They come from daily life as well as battles for supremacy in physical if not realistic spaces.
Like the case of *Yī jiārén zhī míng*, the analysis was carried out simultaneously by the groups at the two Chinese universities. We chose then to carry out this parallel analysis with Chinese students considering Italian texts in order to understand whether they reflected a different sensitivity towards expressions of anger. We then considered the results from the Chinese students alongside those of the Italian ones. Thus, these results do not comprehensively describe how anger is expressed in Italian (or the Neapolitan dialect). The results, however, do again reveal important differences in the number of scenes identified by the groups (Table 7).

Episode	Number of scenes (UIBE)	Number of scenes (ZISU)
One	6	9
Two	7	6
Three	6	4
Four	6	5

Table 7. Number of scenes selected in *My Brilliant Friend*.

The difference in the number of scenes identified can be traced back to the way the two groups identified anger. Some Chinese students considered a reproach as a manifestation of anger. In particular, one

[13] The minute count shown is taken from the HBO episodes subtitled in Chinese.

student group from ZISU interestingly indicated scenes involving the teachers (Ep. 1). In contrast, one student group from UIBE chose those involving parents or fights among siblings (Ep. 3). The teacher-student relationship is perceived as a reason for teachers to get angry since a pupil's negative marks and bad performance are considered a "hindrance to achieving one's goals" (an item shown in the script model). The reproachful tone and its paralinguistic markers (voice intonation, tone, etc.), the insults ("You're a donkey", Ep. 1), the slanders ("Look at the dunce!", Ep. 1), and other idioms ("and you can lead a donkey to water, but you can't make him drink", "Giving a donkey a bath is only a waste of soap and water", Ep. 1) are viewed as displays of anger rather than antecedents of anger. Imperative sentences uttered by teachers and parents have likewise been considered as expressions of anger, such as, "I won't hear of it. You have to go on to high school" (Ep. 3, 41:20). Teachers-parent tensions also represent another area where the groups slightly differed in identifying the same scenes. The following is an example of a scene where the teacher argues with Lila's mother, who says, "You think I'll speak to my husband, and he'll change his mind? […] Do you think my daughter, who's a girl, will go to middle school like a lady while her brothers will go to work? Maybe things are like that in your house. But not in ours". This teacher-parent confrontation belongs to relationships that extend beyond the family environment where, in this case, a more controlled form of anger is noticed. This type of 'reined-in anger', almost gentle if you will, was perhaps unexpected by students and may explain why these expressions of anger are not always identified.

Paying great attention to gestures is a common trait among all the student groups. Gestures, in fact, were examined in detail in every scene, even when verbal expressions did not bear the same meaning, as in the final scene of Episode 4, where we witness a real war fought to the sound of fireworks. Each gesture was understood in terms of its meaning. As impressive as the explicit violence is in almost all its manifestations, we note that the link between the gestures, territorial disputes and efforts to assert domination was not noted. This is the case despite the narrator making this link explicit by saying, "The last night of the year was like a night of battle, in the neighbourhood and all of Naples. The smoke from the gunpowder made everything hazy. If there's a civil war, I thought, […] they would have the same faces as these, the same expressions, the same postures". Conversely, intimidating elements were always identified as an expression of anger in social situations and family relationships. In evaluating the relationships between the actors in the scenes of anger selected by the Chinese students, what is remarkable is that they have difficulty in clearly identifying relationships of intimacy in the Italian series. As previously mentioned, the relationship between a teacher and their pupils is perceived as sometimes intimate and sometimes detached. In some scenes, anger is expressed simultaneously by several characters. Here, it was notably difficult to establish whether an emotion like anger is addressed to in-groups or out-groups. This is because the social and cultural setting depicted in the television series represents a district of Naples that in many cases appears as a sort of extended family.

Regarding the overall results, it is worth pointing out that some expressions were assigned a fixed place in terms of anger. This is the case, for instance, with scatological words like *cazzo* (fuck) and *pezzo di merda* (piece of shit) or the interjection *basta* (enough). Such terms were seen as signalling annoyance and the idea of pushing to have that annoyance or a prolonged suffering end. In the case of the Chinese series, these kinds of terms were also linked to anger. Another element worth considering is that the Chinese students define the use of frequent repetitions of terms or phrases throughout the Italian series as a peculiar linguistic manifestation of anger. Emphatic repeated expressions such as "Go back home! I said get the heck back home!" (Ep. 1), for example, which are expectedly not subtitled in English, thus, become part of the emotional sphere.

Regarding the identification of syntactic structures, a predominance of interrogative forms is confirmed, even though imperative and prohibitive sentences are also frequently noted. Compared to the analysis of the Chinese series, the difference in the Italian one is mainly in the much larger number of insults, offensive remarks and obscenities, as well as emphatic structures. The Chinese students did not detect the presence of such expressions as a culturally specific, foreign element; instead, they deemed them appropriate in the various situations involved.

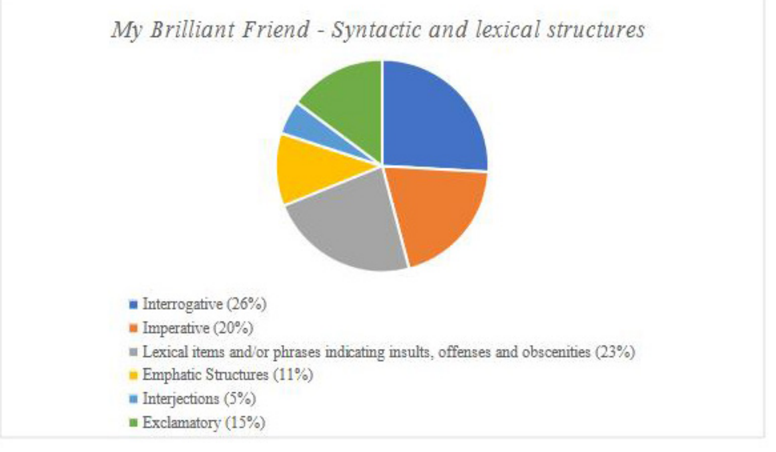
The use of dialect, however, also aroused interest, and the frequent switching into dialectal expressions in angry situations was noticed.

Confirming this is the scene when Elena's mother says, "Do I have to say it in proper Italian?" (Ep. 2), when she is angry about her daughter's wish to continue her studies.
Table 8 and the graph below show some of the possible expressions of anger in Italian by category. The English translation given generally comes from the subtitles provided they were adequate for analysis.

Syntactic and lexical structuress	**Source text examples**
Interrogative sentences (26%)	(1) 'E capito levame 'e mmane 'a cuollo? *Get your hands off me!* (Ep. 1) (2) Ma chi te crire d'essere, famme capì? *Who do you think you are?* (Ep. 2) (3) 'O dialetto nun t'o ricuorde chhiù? *You've forgotten the dialect?* (Ep. 2) (4) Te spezzo 'e coscie si te faie truà nata vota faccia a faccia cu chillo, 'e capito? *I'll break your legs if I see you with him!* (Ep. 3)
Imperative sentences (20%)	(1) Ti devi stare zitta e devi parlare solo quando sei interrogata. *Keep quiet and speak only when you're questioned, is that clear?* (Ep. 2) (2) C'amma ribbellà a chilli strunzi. *We gotta do something, stand up to those assholes.* (Ep. 4)
Lexical items and/or phrases indicating insults, offenses, and obscenities (23%)	(1) 'A signora d' 'o cazzo! *Lady my ass!* (Ep. 1) (2) Si' 'na cessa! *Ugly bitch!* (Ep. 1)
Exclamatory sentences (15%)	(1) T'aggi"a accìdere! *I'll kill you!* (Ep. 1) (2) E quante è vero Dio e levo da faccia da terra! *I swear to God, I'll wipe them off the face of the earth!* (Ep. 4)

Emphatic Structures (11%)	(1) Gnè gnè gnè... *[English translation missing in the subtitles]* (Ep. 1) (2) Si' 'na cessa, si' 'na granda cessa! *Ugly bitch! Big ugly bitch!* (Ep. 1) (3) T'abbasta accussì? T'abbasta accussì? *Is that enough for you? Is that enough for you?* (Ep. 2) (4) Sorema è na zoccola? Sorema è na zoccola? Sorema è na zoccola? *My sister's a slut? - My sister's a slut? - My sister's a slut?* (Ep. 3)
Interjections (5%)	(1) E basta! *Enough!* (Ep. 2) (2) Maronna e ' sta Oliviero e comme me fa 'ncazza, ue'! *God! That Oliviero pisses me off!* (Ep. 3)

Table 8. Examples of syntactic and lexical structures in *My Brilliant Friend*.



4. Conclusions

The development of a specific script model to identify expressions of anger in Chinese and Italian when interpreted as a second language has made it possible to identify two distinct learning steps, at least in this first phase of research. One is strictly connected to the syntactic as well as lexical and linguistic expressions observed by the second-language learners. This makes it possible to highlight that in Chinese, anger is expressed more at the sentence level, which is particularly the case with interrogative and imperative forms. Furthermore, final modal particles perform an important function as well in sentences expressing anger. Conversely, the Chinese students' analysis of the Italian texts highlights the importance of certain sentence types (interrogative and imperative), the use of swear words and the emphatic repetition of terms, entire phrases or sentences in expressing anger. These findings lead us to emphasise the importance of a didactic approach that, on the one hand, clarifies the value of expressions of anger at the sentence level. On the other, there is a need to go beyond just an obsessive search for the translation equivalents of Italian and Chinese phrases or terms. This is because anger is still expressed in a language or culture even if certain linguistic forms for it may not exist.

The second learning step, however, involves the cultural sphere. As previously mentioned, several studies consider anger as an area in which it is easier to observe how cultures achieve the goal of maintaining social order. They do this by calibrating the biologically-based emotional systems of their members. This occurs because emotions are a primary source of motivation for behaviour, and behaviour needs to be managed to maintain social order. This concern further stems from the idea that anger is a potentially destructive emotion in any culture. It may disrupt social bonds, cause physical and psychological harm and invite retaliation. This research has aimed to identify elements that would allow foreign language learners to be more aware of expressions of anger. To that end, we assert that Matsumoto *et al.*'s (2010) hypotheses prove to be somewhat outdated in terms of the direct analysis phase. Indeed, the students involved in this experimental activity were unanimously in agreement in establishing the antecedents eliciting anger and revealing that the control phase for a negative emotion such as anger is rare in both cultures. The question that remains unanswered is if we can claim that an emotion such as anger tends to be expressed and displayed towards out-groups in a culturally recognised collectivist country such as China. Concurrently, in Italy, which is characterised by more individualism, is this emotion addressed more towards in-group members? This is not the place to draw a definitive conclusion in this regard, nor claim that Matsumoto *et al*. are entirely incorrect. However, it raises the issue of having a learning phase in the future that more extensively considers socio-cultural and pragmatics aspects. Thus, we entertain the possibility of collecting additional data at some point to investigate this aspect in more detail. It may involve combining this research with an analysis of the universality versus non-universality of facial expressions of anger. Moreover, another aspect we might further explore relates to non-direct verbal confrontations in which conflict takes on indirect and oblique forms. This is something which the students in the activity discussed here were unable to grasp, despite the fact it may sometimes be a fundamental part of the linguistic manifestation of anger.

References

Baños-Piñero, R. and F. Chaume (2009) "Prefabricated Orality. A Challenge in Audiovisual Translation". *InTRAlinea Online Translation Journal*, «The Translation of Dialects in Multimedia», special issue. www.intralinea.org/specials/article/Prefabricated_Orality (visited 2023/02/20).

Bond, M.H. (1993) "Emotions and Their Expression in Chinese Culture". *Journal of Nonverbal Behaviour*, 17(4): 245-262.

Corno, D. and G. Pozzo (1991) *Mente, linguaggio, apprendimento. L'apporto delle scienze cognitive all'educazione*. Firenze, La Nuova Italia.

D'Urso, V. (2001) *Arrabbiarsi*. Bologna, Il Mulino.

D'Urso, V. and R. Trentin (1998) *Introduzione alla psicologia delle emozioni*. Bari/Roma, Editori Laterza.

Ekman, P. and W.V. Friesen (1971) "Constants across Cultures in the Face and Emotion". *Journal of Personality and Social Psychology*, 17: 124-129.

Ekman, P., W.V. Friesen and P. Ellsworth (Eds.) (1972) *Emotion in the Human Face. Guidelines for Research and an Integration of Findings*. New York, Pergamon Press Inc.

Fang, T., and O.G. Faure (2010) "Decoding Chinese Communication Style: A Yin Yang Perspective". Paper presented at the China Europe International Business School, EAMSA, India. http://www.eamsa.org/wp/wp-content/uploads/2015/10/hdhtyle.pdf (visited 2022/04/04).

Gevaert, C. (2005) "The ANGER IS HEAT Question: Detecting Cultural Influence on the Conceptualization of ANGER Through Diachronic Corpus Analysis" in N. Delbecque, J. van der Auwera and D. Geeraerts (Eds.) *Perspectives on Variation: Sociolinguistic, Historical, Comparative*. Berlin/New York, De Gruyter: 195-208.

Hall, E.T. (1976) *Beyond Culture*. New York, Anchorbooks/Doubleday.

Harkins, J. and A. Wierzbicka (Eds.) (2001) *Emotions in Crosslinguistic Perspective*. Berlin/New York, De Gruyter.

Hofstede, G., G.J. Hofstede and M. Minkov (2010) *Cultures and Organizations: Software of the Mind*. London, McGraw Hill.

Johnson, M. (1987) *The Body in the Mind*. Chicago, Chicago University Press.

King, B. (1989) *The Conceptual Structure of Emotional Experience in Chinese*. Unpublished PhD Dissertation, Ohio State University.

Kornacki, P. (2001) "Concepts of Anger in Chinese", in J. Harkins and A. Wierzbicka (Eds.) *Emotions in Crosslinguistic Perspective*. Berlin/New York, De Gruyter: 255-289.

Kövecses, Z. (1995) "Anger: Its Language, Conceptualization, and Physiology in the Light of Cross-Cultural Evidence", in J.R. Taylor and R.E. McLaury (Eds.) *Language and the Cognitive Construal of the World*. Berlin/New York, De Gruyter: 181-196.

Kövecses, Z. (2000) *Metaphor and Emotion. Language, Culture, and Body in Human Feeling*. Cambridge/New York, Cambridge University Press.

Kövecses, Z. (2010) "Cross-Cultural Experience of Anger: A Psycholinguistic Analysis", in M. Potegal, G. Stemmler and C. Spielberger (Eds.) *International Handbook of Anger*. New York, Springer: 157-174.

Lakoff, G. and M. Johnson (1980) *Metaphors We Live By*. Chicago, University of Chicago Press.

Lakoff, G. and Z. Kövecses (1987) "The Cognitive Model of Anger Inherent in American English", in D. Holland and N. Quinn (Eds.) *Cultural Models in Language and Thought*. Cambridge, Cambridge University Press: 195-221.

Lee, S.Y. (2019) *Emotion and Cause. Linguistic Theory and Computational Implementation*. Berlin, Springer.

Li, X. and Meng, M. (2007) "A Cross-Cultural Characterization of Chinese and English Written Discourse". *Intercultural Communication Studies*, 16(3): 90-98.

Matsumoto, D., H.G. Wallbot and K.R. Scherer (1989) "Emotions in Intercultural Communication", in M.K. Asante, W.B. Gudykunst and E. Newmark (Eds.) *Handbook of International and Intercultural Communication*. Newbury Park, Sage: 225-246.

Matsumoto, D., S.H. Yoo and J. Chung (2010) "Expressions of Anger Across Cultures", in M. Potegal, G. Stemmler and C. Spielberg (Eds.) *International Handbook of Anger*. New York, Springer: 125-137.

Matsumoto, D., S.H. Yoo, J. Fontaine and A.M. Anguas-Wong (2008) "Mapping Expressive Differences Around the World: The Relationship between Emotional Display Rules and Individualism Versus Collectivism". *Journal of Cross-Cultural Psychology*, 29(1): 55-74.

Mrowa-Hopkins, C. and A. Strambi (2005) "How Angry Can You Be in French and Italian? Integrating Research and Teaching for the Development of Pragmatic Competence in L2 Classrooms". *FULGOR*, 2 (2): 48-62.

Ortony, A., G.L. Clore and A. Collins (1988) *The Cognitive Structure of Emotions*. Cambridge, Cambridge University Press.

Paternicò, L.M., V. Varriano and H. Tian (2021) *Grammatica della lingua cinese*. Novara, UTET.

Plutchik, R. (1980) *Emotions: A Psychoevolutionary Synthesis*. New York, Harper-Collins.

Poggi, I. (2008) (Ed.) *La mente del cuore. Le emozioni nel lavoro nella scuola nella vita*. Roma, Armando.

Quaglio, P. (2009) *Television Dialogue. The Sitcom* Friends *vs. Natural Conversation*. Amsterdan/Philadelphia, John Benjamins.

Raskin, V. (1985) *Semantic Mechanism of Humour*. Dordrecht, Kluwer Academic Publishing Group.

Rumelhart, D.E. (1980 [2018]) "Schemata: The Building Blocks of Cognition", in R.J. Spiro, C.B. Bertram and W.F. Brewer (Eds.) *Theoretical Issues in Reading Comprehension*. London, Routledge: 33-58.

Santangelo, P. (2003) *Sentimental Education in Chinese History. An Interdisciplinary Textual Research on Ming and Qing Sources*. Leiden/Boston, Brill.

Scherer, K.R., H.G. Wallbott and A.B. Summerfield (1986 [2010]) (Eds.) *Experiencing Emotion. A Cross-Cultural Study*. Cambridge, Cambridge University Press.

Shank, R.C. and R. Abelson (1977) *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale/New Jersey, Lawrence Erlbaum Associates Publishers.

Turner, J.H. (2000) *On the Origins of Human Emotions: A Sociological Inquiry into the Evolution of Human Affect*. Stanford (CA), Stanford University Press.

Varriano, V. and S. Zuccheri (2022) "Rabbia tradita o tradotta? Il caso della traduzione cinese della serie *Gomorra* e del film *Perfetti sconosciuti*", in G. Zanoni and S. Zuccheri (Eds.) *Emozioni: sentirle, parlarne, tradurle*. *MediAzioni*, 33: D29-D65. https://doi.org/10.6092/issn.1974-4382/15264.

Whatt Smith, T. (2015) *The Book of Human Emotions*. London, Profile Books.

Washburn, G.N. (2001) "Using Situation Comedies for Pragmatic Language Teaching and Learning". *TESOL Journal*, 10(2): 21-26.

Xu, X. 许小颖 and J. Tao 陶建华 (2003) "汉语情感系统中情感划分的研究", in *Proceedings of the 1st Chinese Conference on Affective Computing and Intelligent Interaction*. Beijing, China, December 8-9: 1-7.

Yu, N. (1995) "Metaphorical Expressions of Anger and Happiness in English and Chinese". *Metaphor and Symbolic Activity*, 10(2): 59-82.

Yu, N. (2002) "Body and Emotion: Body Parts in Chinese Expression of Emotion". *Pragmatics and Cognition*, 10(1/2): 341-367.

CONTRIBUTORS

Giorgio Francesco Arcodia is Associate Professor of Chinese Language and Linguistics at Ca' Foscari University of Venice. His work on Sinitic languages has appeared in multiple journals and in edited volumes, and he is the author of *Lexical Derivation in Mandarin Chinese* (Crane Publishing, 2012) and, with Bianca Basciano, of *Chinese Linguistics: An Introduction* (Oxford University Press, 2021). His research focuses on word formation, Chinese dialectology, and the typology of Sinitic languages.

Linda Badan is Associate Professor in Linguistics at the Department of Translation, Interpreting and Communication at Ghent University. Her main research interests include theoretical linguistics at the interface between syntax and prosody and syntax and pragmatics, comparative syntax between Romance and Sinitic languages and the acquisition of Italian as second language.

Bianca Basciano is Associate Professor of Chinese at Ca' Foscari University of Venice. Her research focuses on Chinese morphology and the syntax-semantics interface, and her work has appeared in multiple journals and edited volumes. She is the author of several entries of the *Encyclopedia of Chinese Language and Linguistics* (Brill) and co-author of the entry on Morphology in Sino-Tibetan languages in the *Oxford Research Encyclopedia of Linguistics.* She is the co-author of the book *Chinese Linguistics: An Introduction* (Oxford University Press, 2021).

Chiara Bertulessi is a Postdoctoral Researcher in Chinese Studies at the University of Milan. She received her PhD from the University of Milan, obtained with a dissertation in critical Chinese lexicography. Her research interests revolve around the study of Chinese lexicography and the analysis of political discourse in contemporary China. She is author of the book *L'ideologia nel discorso lessicografico cinese. Analisi critica dello Xiandai Hanyu Cidian* 现代汉语词典 (LED, 2022).

Marco Casentini is a PhD in "Civilizations of Asia and Africa" at the Italian Institute of Oriental Studies of the Sapienza University of Rome. He is currently a research fellow at the Ca' Foscari University of Venice. His major research interests include corpus analysis, discourse markers, discourse grammar, the syntax-prosody interface analysis of Topics in Chinese, the interpretation of covert subject pronouns and L2 Chinese acquisition.

Sergio Conti is Assistant Professor at Roma Tre University (Italy). His research interests primarily lie in the field of Chinese applied linguistics, including the acquisition and teaching of Chinese as a second language. He is also a member of the Lab on Language-based Area Studies for Sustainability Transformations (LabLAST, Ca' Foscari University of Venice), working on an Ecological Discourse Analysis framework for Chinese environmental texts.

Gloria Gabbianelli is Assistant Professor in the Department of Communication Sciences, Humanities and International Studies, at the University of Urbino Carlo Bo. She teaches Chinese language in MA and BA courses. Her research interests include Chinese as a second and as a foreign language, Chinese language acquisition and education, teaching methodology and teaching materials. Her research publications mainly focus on the acquisition and teaching of Chinese as a foreign language.

Alessia Iurato is a PhD Candidate at Ca' Foscari University of Venice and University of Bremen (Universität Bremen). She is writing her PhD dissertation titled "The Acquisition of Chinese 是*shì*…的 *de* clefts by L1 Italian learners: Triangulating corpus and experimental data". She is the author of a few papers on Chinese learner corpora and received the "2022 Hawai'i International Conference on Chinese Studies Graduate Student Award" for the best research paper (University of Hawai'i, Mānoa). Her research interests include L2/LS Chinese acquisition, learner corpus linguistics, Chinese syntax and pragmatics.

Carmen Lepadat obtained her PhD with honors from Sapienza University of Rome, discussing a thesis on the pragmatics of the utterance right periphery in Mandarin. Currently, she is a Postdoctoral Research Fellow at Roma Tre University, working on the acquisition of Chinese directional complexes, and she teaches Chinese language and culture at the University of Siena. Her main research interests include pragmatics, information structure, intersubjectivity and modality.

Feng Lisi is a PhD Candidate at the University of Milan. She completed a BA in Italian Studies in China (2017) and a MA in Linguistics (2019) (Cum Laude) at the University of Padua. Her research interests focus on cultural and linguistic exchanges between Italy and China. She has published articles in various journals. She is an OIM project (Crusca) member and an academic assistant at Accademia Ambrosiana.

Luisa M. Paternicò, PhD in Civilizations, cultures e societies of Asia and Africa at Sapienza University of Rome (2011), is currently Associate Professor of Chinese Language at "L'Orientale" University of Naples, where she teaches Chinese Language at both Bachelor and Master degree. Her research interests include history of Chinese linguistics and of Chinese language teaching, together with modern didactics of Chinese language (both Mandarin and Cantonese). Main monographs: *When the Europeans Began to study Chinese* (Leuven, 2013); with E. Raini, *Il cinese tecnico-scientifico* (Milan 2014); with J. Lam, *Corso di lingua cantonese* (Milan 2017); with V. Varriano and H. Tian, *Grammatica della lingua cinese* (Turin 2021).

Chiara Romagnoli received her PhD from Sapienza University of Rome and is Associate Professor of Chinese language since 2014 at Roma Tre University. Her research interests include Chinese vocabulary acquisition, discourse markers and lexicography. She is the author of several works, including research papers, a monograph on Chinese functional words (*Grammatica cinese. Le parole vuote del cinese moderno*, Hoepli, 2012) and editor of a volume on Chinese pedagogy in Italy (*La lingua cinese in Italia. Studi su didattica e acquisizione*, Roma TrE Press, 2021, with S. Conti).

Carlotta Sparvoli is Associate Professor at Ca' Foscari University of Venice. Her primary research fields are the semantics of modal expressions and the acquisition and teaching of L2 Chinese. Her most recent publication is "The Expression of Modality in Classical Chinese: Notions, Taxonomy and Distinctive Features", *Oxford Research Encyclopedia*. She is an Executive Board member of the International Association of Chinese Linguistics (IACL) and serves as President of the Italian Association of Chinese Linguistics.

Sun Yangyu is a post-doc research fellow at the University of Milano-Bicocca. She obtained her PhD in Linguistics from the University of Padua in 2022. Her research interests include theoretical comparative syntax and Chinese syntax, with a special focus on Right Dislocation, modal verbs, adverbs, the DP structure and the *ba*-construction.

Tommaso Tucci is a PhD candidate in linguistic and literary sciences at the Catholic University of the Sacred Heart. He holds a master's degree in Asian and Mediterranean African Languages and Civilizations from Ca' Foscari University of Venice. He is currently a tenured teacher of Chinese language and culture in an Italian upper secondary school and a lecturer at the Confucius Institute of the Catholic

University of the Sacred Heart in Milan. His main research interests include syntax, pragmatics, and learning and teaching of Chinese language as L2.

Yuan Huahung teaches French at the Department of Applied Foreign Languages at National Taipei University of Business, Taiwan. She obtained her PhD in Linguistics at Paris-Cité University. Her research interests include phenomena at syntactic-semantic interface in Chinese languages.

Valeria Varriano is Associate Professor of Chinese Language and literature at "L'Orientale" University of Naples. She is engaged in a study of the representation of history in film, fiction, and TV entertainment in the post-Tiananmen era (1989-present). She also conducts studies following a multidisciplinary approach, which merges linguistic and sociolinguistic analysis with cultural and media studies. Her interest in education has spurred her to find the didactic implications of the studies she has conducted. Her latest publications are: with S. Zuccheri "Rabbia tradita o tradotta? Il caso della traduzione cinese della serie Gomorra e del film Perfetti sconosciuti", with Velardi and Cigliano *Shangke le*, *Corso di lingua cinese per le scuole superiori* (2022) and with Paternicò and Tian *Grammatica della lingua cinese* (2022).

Serena Zuccheri is Full Professor of Chinese Language and Translation at the University of Bologna (Forlì Campus). Her recent research interests are non-professional translation (fansubbing), linguistic impoliteness and emotional communication. She wrote several articles and monographs, and her most recent publications are: *Zheng Xiaoqong. Versi di umana resistenza* (2022); with V. Varriano "Rabbia tradita o tradotta? Il caso della traduzione cinese della serie Gomorra e del film Perfetti sconosciuti" (2022).

Ultimi volumi pubblicati:

Chiara Elefante, *Traduzione e paratesto.*

Francesca Biagini, *L'espressione della finalità in russo. Uno studio contrastivo con l'italiano.*

Francesca Biagini e Svetlana Slavkova (a cura di), *Contributi italiani allo studio della morfosintassi delle lingue slave*.

Elio Ballardini, *Traduire devant la justice pénale. L'interprète traducteur dans les codes de procédure pénale italiens aux XIX^e^ et XX^e^ siècles.*

Ahmad Al-Addous e Sara Nanni (a cura di), *Drammi. Due opere teatrali di Sa'd Allah Wannus.*

Rachele Antonini (a cura di), *La mediazione linguistica e culturale non professionale in Italia.*

Félix San Vicente, Ana Lourdes de Hériz, María Enriqueta Pérez Vázquez (eds), *Perfiles para la historia y crítica de la gramática del español en Italia: siglos XIX y XX. Confluencia y cruces de tradiciones gramaticográficas.*

Amalia Amato e Gabriele Mack, *Comunicare tramite interprete nelle indagini di polizia. Implicazioni didattiche di un'analisi linguistica.*

Gloria Bazzocchi e Raffaella Tonin (a cura di), *Mi* traduci *una storia? Riflessioni sulla traduzione per l'infanzia e per ragazzi.*

Francesca Gatta (a cura di), *Parlare insieme. Studi per Daniela Zorzi.*

Adele D'Arcangelo (a cura di), *Promuovere la competenza interculturale nella didattica della traduzione. L'esperienza della Scuola Interpreti e Traduttori di Forlì.*

Roberta Pederzoli, Licia Reggiani, Laura Santone (dir.), *Médias et bien-être. Discours et représentations.*

Elide Casali, *Il bambino e la lumaca. Rileggere Piero Camporesi (1926-1997).*

Derek Boothman (a cura di), *La traduzione come luogo di incontro e di scontro*.

Félix San Vicente, Gloria Bazzocchi, Pilar Capanaga (eds), *Oraliter. Formas de la comunicación presencial y a distancia*.

Motoko Ueyama, Irena Srdanović (eds), *Digital resources for learning Japanese*.

Raffaella Baccolini, Roberta Pederzoli, Beatrice Spallaccia (eds), *Literature, Gender and Education for Children and Young Adults / Littérature, genre, éducation pour l'enfance et la jeunesse*.

Adele D'Arcangelo, Chiara Elefante, Valeria Illuminati (eds), *Translating for Children Beyond Stereotypes / Traduire pour la jeunesse au-delà des stéréotypes*.

Mariachiara Russo, Emilia Iglesias Fernández, María Jesús González Rodríguez (eds), *Telephone interpreting*.

Anabela Cristina Costa da Silva Ferreira, *A língua do Direito de Família*.

Greta Zanoni, *Lo sviluppo della competenza pragmatica in italiano: dalla rete alla classe*.

Patrick Leech, *Cosmopolitanism, dissent, and translation. Translating radicals in eighteenth-century Britain and France.*

Valeria Illuminati, *La traduzione dei classici per l'infanzia in una prospettiva di genere*.

Gloria Bazzocchi *et al.* (coords), *Nosotros somos nos y somos otros. Estudios dedicados a Félix San Vicente*.

Finito di stampare nel mese di aprile 2023
per i tipi di Bologna University Press