

# Evidentiality ‘In’ and ‘As’ Context

## Corpus-Based Insights About the Mandarin V-*过* *guo* Construction

Vittorio Tantucci

Lancaster University, UK

Aiqing Wang

Lancaster University, UK

**Abstract** In this paper we argue that evidentiality can be a category of a linguistic system that emerges from the intersection between form, usage and ‘contextual situatedness’. We provide a multivariate corpus-based case study about the usage of the V-*过* *guo* construction in written Mandarin, and show how the text types in which the chunk appears significantly contribute to determine its pragmatic usage and its emergent meaning grounded in shared knowledge and collective recognition. This approach sheds new light on two critical issues. The first is that evidentiality is an important grammatical category of documentary, factual and academic prose in Mandarin Chinese. The second, much broader, claim of this paper is that generalisations about grammatical/semantic categories need to account for the usage of specific items in context. In this sense, ‘physical and sociocultural situatedness’ is as important a dimension as form and meaning in order to define categorial membership.

**Keywords** Chinese. Evidentiality. Context. Corpus-based. Multifactorial.

**Summary** 1 Introduction. – 2 The Mandarin V-*过* *guo* Construction. – 3 The Grammaticalisation of V-*过* *guo*. – 4 A Corpus-Based Account of V-*过* *guo* in Context. – 4.1 Data Retrieval and Annotation. – 4.2 Data Analysis. – 4.3 Evidential vs Experiential Categorisation in Context. – 5 Conclusions.



## 1 Introduction

It has been pointed out that the Mandarin<sup>1</sup> experiential marker 过 *guo*, originally expressing the past experience of a syntactic subject, has recently grammaticalised into an evidential construction (cf. Chappell 2001; Tantucci 2013, 2015a, 2015b, 2016c; Tantucci, Wang 2020b). In this study, we focus on the usage of V-过 *guo* in two comparable written corpora of Mandarin Chinese, namely the Lancaster Corpus of Mandarin Chinese (LCMC) (McEnery, Xiao 2004) and UCLA corpus of written Mandarin (Tao, Xiao 2012). The former includes texts from 1988 and 1992, whereas the latter includes texts from 2000 to 2005. Both corpora include one million words and are balanced with respect to the text types of which they are composed, so that they can be compared with one another. The aim of the present analysis is to shed light on the relationship between evidential reasoning and context and whether specific genres and textual environments favour the usage of evidential polysemies of V-过 *guo*. We are similarly interested in assessing whether the process of grammaticalisation of V-过 *guo* towards evidentiality is occurring at the expense of experiential usages of the same construct.

First of all, we can look at the formal and semantic differences between experiential and evidential usages of V-过 *guo*. Consider the two examples below:<sup>2</sup>

1. 她的鼻梁很细,我从来没有看过人有这么细的鼻梁,因而反把她年轻的、瘦削的脸衬得丰满起来。(LCMC / P: Romantic fiction)  
tā de bí-liáng hěn xì wǒ cónglái méiyǒu  
she SP nasal-bridge very narrow I never NEG  
kàn-**guo** rén yǒu zhème xìde bí-liáng yīn’ér fǎn  
see-EXP person have such narrow nasal-bridge thus turn  
bǎ tā niánqīng de shòuxuēde liǎn chèn de  
BA she young SP skinny face seem DEG  
fēngmǎn qǐlái  
chubby become  
‘The bridge of her nose is very thin, I have never seen anyone with such a thin one, and it makes her young, skinny face look chubby’.

<sup>1</sup> When Mandarin will be used in isolation, it will refer to present-day Mandarin (aka. currently spoken 普通话 *pǔtōnghuà* of Mainland China) throughout the present paper.

<sup>2</sup> The glosses follow the general guidelines of the Leipzig Glossing Rules. Additional glosses include: BA = ‘ba construction particle’; DEG = ‘complement of degree’; EMP = ‘emphatic marker’; EVD = ‘evidential’; EXP = ‘experiential’; ST = ‘structural particle’.

2. 本世纪以来, 长江发生过三次严重的洪灾, 其中1931年和1935年两次大洪水, 分别淹地 5090万亩和2264万亩, 死亡 14.5万人和14.2万人。(LCMC / J: Academic prose)

běn shìjì yǐlái Chángjiāng fāshēng-guo sān cì  
 this century since Yangtze.River happen-EVD three time  
 yánzhòng de hóngzāi qízhōng 1931 nián hé 1935  
 severe SP flood among 1931 year and 1935  
 nián liǎng cì dà hóngshuǐ fēnbié yān  
 year two time major flood respectively inundate  
 dì 5090 wàn mǔ hé 2264 wàn mǔ  
 land 5,090 ten.thousand mu and 2,264 ten.thousand mu  
 sǐwáng 14.5 wàn rén hé 14.2 wàn rén  
 die 14.5 ten.thousand person and 14.2 ten.thousand person  
 ‘Since the beginning of this century, there have been three severe floods  
 in the Yangtze River, including two major floods in 1931 and 1935, which  
 inundated 205,997 and 91,626 square meters of land and killed 145 thou-  
 sand and 142 thousand people respectively’.

In (1), the speaker is genuinely expressing some subjective/personal impression that directly underpins his/her own personal experience, namely *that s/he normally has never seen a nose as fine* as the one of the character that is being narrated. S/he is therefore establishing reference to his/her own subjective experience and personal impressions about a specific event or state of affairs. In Pragmatics, the notion of perlocutionary effects regards *what a speaker intends an utterance to achieve in an addressee* (cf. Austin 1962; Searle 1976). The perlocutionary effects of (1) are clearly not the ones of informing the reader of a piece of documented information, but most likely to share his/her emotional/sensorial experience and/or personal affects. Simply put, the usage of 过 *guo* in (1) cannot express a piece of collective knowledge (it cannot be marked by evidential functions such as *it is known that*, or *as it seems*), but only personal experience and related emotions resulting of the speaking subject as an individual.

The usage of 过 *guo* in (2) is rather different. In this case the syntactic subject of the sentence is inanimate, and the event that is reported has not been necessarily experienced by the speaker. A completely different speech act is performed in this case. The speaker is no more referring to his/her personal affects, or the ones of a syntactic subject. Rather, s/he is reporting or presenting (cf. Faller 2002; Tantucci 2016a, 2016b, 2016c) a piece of information that s/he has somehow acquired and which s/he could potentially provide evidence for. Interestingly, the text types in which these two usages occur also differ substantially. In the former case the narration occurs in a fictional context, and it is therefore more likely to be aimed to entertain or empathise with the reader. In the latter usage, the V-过 *guo* construction is used in academic prose and is functional to mark a

piece of information as a fact that can be considered as reliable and documented/documentable. Intersubjectively, we could say that usages such as (1) tend to be aimed at establishing empathy among interlocutors, whereas utterances of the kind of (2) aim to be persuasive and reliable. Finally, it is important to note that both contexts of usage in (1) and (2) do indeed require the post-verbal marker 过 *guo* and could not be uttered with an evidentially/experientially neutral perfective marker such as 了 *le*<sup>3</sup> (Tantucci 2013, 225).

§ 2 provides an overview of the V-过 *guo* construction and its different usages. It also provides the operational criteria to disentangle experiential versus evidential senses. § 3 is based on a diachronic discussion about the grammaticalisation of the V-过 *guo* construction and the semasiological formation of different polysemies. The main case-study in § 4 is then centred on the relationship between evidential vs experiential usages of 过 *guo* and the text types in which they tend to occur. In particular, we will be focusing on the following research questions:

- What is the distribution in different text types of evidential versus experiential usages of V-过 *guo*?
- Have there been any significant changes in the partition of usages of V-过 *guo* in the last thirty years?
- What is the relationship between the different senses of V-过 *guo* and the textual environment in which they occur?

## 2 The Mandarin V-过 *guo* Construction

In the literature, V-过 *guo* is commonly considered as a polysemous construction. It can express directionality (e.g. Li, Thompson 1981; Chen 2008), therefore emphasising the actional (i.e. underpinning *Aktionsart*, see Vendler 1967) movement in space of dynamic verbs, as in 拿过 *náguò* ‘to take/seize’, 走过 *zǒuguò* ‘to walk towards a certain direction’, 递过 *dìguò* ‘to hand over’, and others (Tantucci 2015a, 69). It can express completivity (cf. Bybee, Perkins, Pagliuca 1994, 51; see also Dahl 1985, 95 on conclusives) or traversativity (Tantucci 2015a), thus describing the phasal meaning of “do[ing] something thoroughly and to completion”, as conveyed by expressions such as *to shoot someone dead* or *to eat up*. The “lexical sources of completives [...] are all dynamic verbs or directionals, as they all suggest action or movement” (Bybee, Perkins, Pagliuca 1994, 59). They are actionally durative, as in 吃过 *chīguò* ‘to finish eating’ or 看过 *kànguò*

---

3 In the case of (1) this test would require a positive polarity.

‘to end up watching’.<sup>4</sup> In example (3) below, V-过 *guo* expresses that the *action of eating the noodles has been completed* or ‘traversed’ (Tantucci 2015a) so that a second action could be carried out or not.

3. 吃饭时没留意窗外, 吃过一碗刀削面走出小饭馆。(UCLA / G: Biography memoirs)
- chī-fàn shí méi liúyì chuāng wài chī-guò yì*  
eat-meal while NEG pay.attention window outside eat-COMPL one  
*wǎn dāoxiāomiàn zǒu-chū xiǎo fànguǎn*  
bowl noodles walk-out small restaurant  
‘I did not pay attention to the outside of the window while eating; I finished a bowl of noodles and walked out of the small restaurant’.

These particular usages of V-过 *guo* do not contribute to the illocutionary force of the utterance, as they merely intervene lexically on the *Aktionsart* (Vendler 1957) – elsewhere alternatively called lexical aspect (Olsen 1997), transformativity (Johanson 2000) or situation aspect (Smith 1997) – of a verbal compound [VV]. Simply put, it only marks the temporal constituency or the internal phase structure IPS (Johanson 2000) of a predicate, i.e. whether an action has been brought to completion or to some resultant state.

A third function of V-过 *guo* is the “experiential perfect” usage (Comrie 1976, 58; Li, Thompson 1981; Dahl 1985, 141; Carey 1994; Yeh 1996; Dai 1997; Smith 1997; Dahl, Hedin 2000; Xiao, McEneary 2004; Lin 2006, 2007; Chen 2008; Wu 2008), whereby the construction indicates the past experience of the syntactic subject, as in example (1) (§ 1) or in expressions such as 我去过北京 *wǒ qù guo Běijīng* ‘I have been to Beijing before’, see also (4) below:

4. 或许林徽因的心情也是这般, 从来没有固执地想过要什么, 也没有刻意去拒绝什么。(UCLA / G: Biography memoirs)
- huòxǔ Lín Huiyīn de xīnqíng yě shì zhèbān cónglái*  
perhaps Lin Huiyin SP mood also be like.this never  
*méiyǒu gùzhíde xiǎng-guò yào shénme yě méiyǒu*  
NEG stubbornly think-EXP want what also not  
*kèyì qù jùjué shénme*  
deliberately go refuse what  
‘Perhaps Lin Huiyin’s mood is also like this; she never stubbornly thought about what she wanted, nor did she deliberately refuse anything’.

<sup>4</sup> In Mandarin, both directional and completive usages of 过 *guò* retain the fourth tone, whereas more grammaticalised forms tend to be toneless. The *pinyin* notation of the rest of this paper will account for this distinction.

In (4) above, the function of V-过 *guo* is no more the one of expressing that a durative event has been completed, but rather to convey that the animate subject of the sentence, 林徽因 *Lín Huīyīn*, has never experienced a particular feeling, namely the one of *being obstinate in wanting something*. Table 2 below provides the diagnostics for identifying experiential usages of V-过 *guo*:

**Table 1** Diagnostics for identifying 过 *guo* as an experiential (adapted from Tantucci 2015a, 87)

过 <i>guo</i> as an experiential
Profiles the syntactic subject’s past experience.
Employed as a perfect in contexts where the syntactic subject has been through some experience before.
Frequently used with dynamic verbs.
Used generally in the first person, in negated statements or in second person questions (Dahl 1985; Dahl, Hedin 2000; Tantucci 2013).
It cannot collocate with the perfective post-verbal 了 <i>le</i> .*
It can collocate with the adverbials 曾经 <i>céngjīng</i> ‘once’ or 从来 <i>cónglái</i> ‘never’.
It cannot collocate with inanimate subjects.
It can collocate with absolute-state predicates (rare).
Not felicitous when collocating with IE adverbials such as 据了解 <i>jù liǎojiě</i> ‘it is understood that’, 好像 <i>hǎoxiàng</i> ‘apparently’, 众所周知 <i>zhòngsuǒzhōuzhī</i> ‘as everyone knows’.
* This is a diagnostic that helps distinguishing comparatively more grammaticalised usages of 过 <i>guo</i> (e.g. experiential and evidential) from cases where 过 <i>guò</i> is used as a completive or a directional complement, such as in 该联络的事宜都联络过了 <i>gāi liánluò de shìyí dōu liánluò guò le</i> ‘all the arrangements that required contacts where dealt with’ (LCMC / E14).

In Tantucci (2013; 2015a), it is also argued that 过 *guo* developed a more grammaticalised function underpinning knowledge ascription and evidentiality. At this stage of change of 过 *guo*, the notion current relevance for the here-and-now of the conversation underpins a presentative stance rather than an assertive one (Faller 2002). That is, while an assertive speech act has the sincerity condition that the speaker believes P and is unmarked with respect to its reliability, in the case of presentative utterances the speaker/writer merely ‘introduces’ a piece of knowledge s/he acquired somehow for the benefit of the addressee/reader. In this latter case, the speaker/writer marks the proposition as a piece of information that is somewhat ‘reliable’ and which can be potentially documented/confirmed. While experiential usages of 过 *guo* tend to occur in questions and in negative statements, evidential ones show a tendency to occur assertively, in the declarative mood (Tantucci 2013, 2015a; Tantucci, Wang 2020b). This functional and formal tendency is due to the presentative illocutionary force of evidential statements, and the fact that the perlo-

cutionary effects of P are distinctively the ones of informing a specific or generic addressee, rather than expressing subjective affective concern or empathy to the interlocutor. As a result, evidential usages of 过 *guo* tend to occur in the third person or in impersonal/subjectless constructions (Tantucci 2013, 2015a).

5. 在现实主义与古典主义之间, 出现过浪漫主义的“叛乱”。(LCMC / J: Academic prose)

*zài xiànshízhǔyì yǔ gǔdiǎnzhǔyì zhījiān chūxiàn-guo*  
at realism and classicism between exist-EVD  
*làngmànzhǔyì de pànlùn*  
romanticism SP rebellion

‘There used to exist a ‘rebellion’ of romanticism between realism and classicism’.

In the academic context of example (5) above, no experiential meaning is at issue. The author is not interested in sharing his/her own or someone else’s past experience with the reader. Rather, s/he purposely marks the proposition as a piece of knowledge that bears some sort of social recognition and which can be potentially confirmed and verified. In other words, a different ‘pragmeme’ is at play, viz. a different “situational prototype capable of being executed in the situation” (Mey 2001, 221). In this paper, we will argue that “contextual situatedness” (cf. Mey 2010; Haugh 2012) is a fundamental dimension that inherently informs meaning, and in particular contributes to determine the polysemic status of the V-过 *guo* construction. In Pragmatics, it is stressed that the physical and cultural environment plays a fundamental role in the encoding of the illocutionary force of an utterance. In other words, speech acts “in order to have an effect, must be situated” (Mey 2010, 2883; Capone 2005; Tantucci 2016c). The different intersection between contextual situatedness and illocutionary force that we find in (5) above determines a distinctive evidential reading of the utterance. In fact, in the same context, the merely perfective marker 了 *le* would not be idiomatic (to some degree not grammatical), as it would lack added evidential meaning that marks the proposition as a piece of ‘documented’ evidence, which bears collective recognition (\*出现了浪漫主义的“叛乱” *chūxiàn le làngmànzhǔyì de pànlùn*) (Tantucci 2013, 255). In table 2 below, we report the formal and functional diagnostics for identifying evidential usages of V-过 *guo*:

**Table 2** Diagnostics for identifying 过 *guo* as an interpersonal evidential (IE) (adapted from Tantucci 2015a, 88)

过 <i>guo</i> as an evidential
Profiles the speaking subject’s (Benveniste [1958] 1971; Traugott 2003; Langacker 2008) acquired information.
Employed in contexts characterised by an epistemic or presentative stance (Mushin 2001; Faller 2002), that is, the speaker/writer markedly ‘introduces’ a particular piece of knowledge s/he has acquired somehow.
Frequently in third person declaratives.
It cannot collocate with the perfective post-verbal 了 <i>le</i> .
It can collocate with the adverbials 曾经 <i>céngjīng</i> ‘once’ or 从来 <i>cónglái</i> ‘never’.*
It can collocate with inanimate subjects.**
It can collocate with absolute-state predicates (rare).
Felicitous when collocating with IE adverbials such as 据了解 <i>jù liǎojiě</i> ‘it is understood that’, 好像 <i>hǎoxiàng</i> ‘apparently’, 众所周知 <i>zhòngsuǒzhōuzhī</i> ‘as everyone knows’.
* This indicates that 过 <i>guo</i> reached a grammaticalisation stage where it can express aspectual discontinuity or anti-resultativity (e.g. Plungian, van der Awera 2006; Tantucci 2015a), which in turn is not possible for completive and directional usages of the same form.
** This is an important diagnostic as what is at issue in evidential usages is a piece of documented and/or socially recognised information, rather than the subjective experience of an individual. Impersonal usages (absent at earlier stages of the grammaticalisation of 过 <i>guo</i> ) are an important sign of this shift, as the absence of a syntactic subject is precisely due to the attempt to communicate <i>what has accordingly happened</i> , rather than <i>what has been once experienced by someone</i> , i.e. the syntactic subject of the sentence (Tantucci 2015a, 91).

Evidentiality has been defined as “the existence of a source of evidence for some information” (Aikhenvald 2004, 1), the “encoding of the speaker’s (type of) grounds for making a speech act” (Faller 2002, 2), or the communication of a piece of “acquired knowledge” (Tantucci 2013, 214). Evidentials relativise or measure the information status of the sentence (Rooryck 2001a, 125; 2001b), yet in many languages, such as English, do not constitute a grammatical category and are generally communicated through adverbials or discourse markers such as *apparently* and *allegedly* (see Mushin 2001, 54; Narrog 2009, 10), predicates conveying an evidential meaning such as *it seems that*, *it appears that*, and *I saw that*, pragmatic strategies (see Aikhenvald 2004), or overtly expressed contextual elements providing some type of information. In our view, in languages where evidentiality does not correspond to a distinctive inflectional category, it is precisely the intersection between form, usage, and context that define an evidential reading. Similarly, it could also be argued that, even in languages where evidential systems are highly complex and grammaticalised (e.g. mostly spread through Northern, Central America, Eastern Europe, central and Southeast Asia; Aikhenvald 2004, 303), there is still a crucial intersection between contextually situatedness and usage of

those forms (see for instance the hybrid case of Gitksan evidentials, which are entirely optional and not paradigmatically organised; Peterson 2010). The inherent relationship between contextual situatedness and formal usage of some evidentials is an argument that has been put forward by Squartini (2012) in the discussion of the subcategory of circumstantial evidentiality, but also by Capone (2005; 2010) and Tantucci (2016c) concerning the crucial role of physical and sociocultural context for the encoding of so-called ‘evidential pragmemes’.

A crucial dimension that is missing from the classification in table (2) above is therefore the one of ‘contextual situatedness’ of the V-过 *guo* construction. That is to say, the diagnostics that are reported in each table take into account formal and functional elements of usage, yet they overlook the textual and sociocultural environment of each polysemy. In this sense, a multivariate corpus-based analysis can shed important light on the holistic relationship between form, illocutionary force and context. Significant intersections of the variables subsumed by formal, pragmatic and contextual dimensions are referred to as **illocutional concurrences (IC)** (Tantucci, Wang 2018, 2020a, 2020b; Formato, Tantucci 2020). Namely, ICs encompass converging factors at different levels of verbal experience that contribute, both locally (i.e. at the morphosyntactic level) and peripherally (i.e. at the illocutionary level), to the encoding of contextually and culturally situated speech acts. The final discussion of this paper will be devoted to the inherent relationship between contextual situatedness and schematic categorisation of form and meaning. A specific focus will be placed on the interdependence of conventional association of linguistic functions and the situation type in which they are used as an important factor of semantic and grammatical change.

### 3 The Grammaticalisation of V-过 *guo*

In this brief section, we discuss the importance of context in the diachronic reanalysis of V-过 *guo* as an evidential construction. This claim will be further discussed in § 4, where we will provide a detailed multivariate analysis of the synchronic usage of V-过 *guo* in the LCMC and the UCLA corpora of Mandarin Chinese.

During the 唐 Tang dynasty (618-907 AD), 过 *guò* starts to occur in the second slot of [VV] constructions with a specific completive/traversative meaning (Cao 1995, 38), therefore expressing lexically the phase where an action has been completed/traversed. Different from early directional usages, this new function collocates with durative verbs that do not necessarily express physical movement:

6. 每至义理深微常不能解处, 闻醉僧诵过经, 心自开解。(纪闻 *Jìwén*, 太平广记 *Tàipíng guǎngjì*, 异人异僧释证卷 *Yìrén yìsēng shìzhèng juàn*, 第 *dì* 81-101 卷 *juǎn*, Cao 1995, 38)

*měi zhì yǐlǐ shēnwēi cháng bù néng*  
 every arrive argumentation mysterious often NEG can  
*jiě chù wén zuì sēng sòng-guò jīng*  
 comprehend place hear drunk monk recite-COMPL scripture  
*xīn zì kāi jiě*  
 heart self open understand

‘Every time the argumentation would become too difficult and mysterious, all the parts that s/he could not comprehend would then become clear after s/he listened to that drunk monk reading through them’.

From (6) above, we can see that 过 *guò* now starts to convey completeness/traversativity, as it marks the phasal meaning of completing/traversing an action, rather than marking a syntactic subject’s past experience. Cao notes that during the Tang dynasty the phasal meaning of 过 *guò* merely

indicates the action itself, and never stresses the subsequent results of the event [...] this is evident from the missed co-occurrence with resultative verbs such as 关 *guān* ‘to close’, 锁 *suǒ* ‘to lock’, 盛 *chéng* ‘to fill’ or absolute states such as 老 *lǎo* ‘be/grow old’, 冷 *lěng* ‘to be cold’, 红 *hóng* ‘to be red’, 白 *bái* ‘to be white’ and others. (1995, 40)<sup>5</sup>

A possible operational model that can inform the stages of semantic and grammatical change of V-过 *guo* is the Invited inferencing theory of semantic change (IITSC) (Traugott 1999; Traugott, Dasher 2002, 5; see also Dahl 1985, 11). IITSC states that inferences pragmatically induced from the speaker/writer to the addressee/reader tend to become conventionalised and determine new semantic polysemies within a construction. In a subsequent stage of reanalysis, due to its semantic element of discontinuity to the present, the V-过 *guo* construction starts to be encoded as a perfect with a conventionalised meaning expressing past-experience of an animate subject. Earliest evidence of this is found between the Tang and the Song (960-1279 AD) dynasties whereby 过 *guo* starts to collocate with mental verbs or verbs referring to the syntactic subject’s past experience, as in the case of 尝 *cháng* ‘to taste’, 验 *yàn* ‘to experience’, 问 *wèn* ‘to ask’ (Lin 2004, 45), albeit it is not frequently used before the Yuan dynasty (1271-1368 AC) (Cao 1995, 43; Lin 2004, 42):

<sup>5</sup> Translated and readapted from Chinese. Unless otherwise indicated all translations are by the Authors.

7. 看文字须仔细, 虽是旧曾看过, 重温亦须仔细。(朱子语类 *Zhūzǐ yǔlèi*, 卷一〇 *jǔan yīlíng*, Cao 1995, 41)
- |                 |              |           |             |               |            |             |                |
|-----------------|--------------|-----------|-------------|---------------|------------|-------------|----------------|
| <i>kàn</i>      | <i>wénzì</i> | <i>xū</i> | <i>zǐxì</i> | <i>suíshì</i> | <i>jiù</i> | <i>céng</i> | <i>kàn-guo</i> |
| see             | character    | must      | careful     | although      | old        | once        | see-EXP        |
| <i>chóngwēn</i> | <i>yì</i>    | <i>xū</i> | <i>zǐxì</i> |               |            |             |                |
| review          | also         | must      | careful     |               |            |             |                |
- ‘When you look at a character you must be attentive, even if it is one that you saw before, you still have to be attentive’.

In the case of (7), 过 *guo* no longer simply intervenes on the *Aktions-art* of the predicate on a lexical level. It has now developed a new grammaticalised function of experiential perfect (e.g. Comrie 1976). It therefore expresses current relevance of a previous experience occurring in a vague, discontinuous past. The bulk of the literature focusing on the aspectual features of 过 *guo* is distinctively focused on this particular usage. The main aspectual features of the experiential V-过 *guo* that emerge from the literature are the following:

- a. It marks an eventuality having at least one occurrence in the past.
- b. It has a ‘class’ meaning, viz. refers to an event type, rather than a specific instantiation.
- c. It expresses aspectual discontinuity to the present/or a reference time.
- d. It encodes only repeatable eventualities.
- e. It marks an event that is temporally independent from others in the discourse.
- f. It expresses a past eventuality often in the form of someone’s past experience (e.g. Chao 1968; Li, Thompson 1981; Yeh 1996; Dai 1997; Smith 1997; Lin 2006, 2007; Wu 2008; Li 2011).

In experiential usages of V-过 *guo*, the original actional meaning of ‘having been through an action’ that was originally encoded on a lexical level, has now turned into a more speaker-based meaning whereby some animate subject’s past experience becomes at-issue for the here-and-now of the speech event.

While both completive or resultant states are attested to be common lexical sources of perfects (i.e. resultative, hot-news, existential, experiential meanings; see McCawley 1971; Portner 2003; Dahl, Hedin 2000), in the case of 过 *guo*, aspectual discontinuity and ‘absence’ of results are themselves the trigger of specifically experiential and subsequent evidential reanalyses of the chunk: i.e. 我年轻过 *wǒ niánqīng guo* ‘I have been young (albeit I am not anymore)’ (see Comrie 1976; Carey 1994; Dahl 1985; Dahl, Hedin 2000; Chappell 2001; Li 2011; Tantucci 2013 for specific discussions about the typological features of experiential perfects).

It is acknowledged that experiential and existential perfects express relevance to the present without expressing a resultative continuation of the past event up to the moment of speech. This is the case of a well-known example:

8. The Earth has been hit by giant asteroids before. (Portner 2003, 464)

Usages involving a discontinuous past such as (8) show that relevance needs to be intended as having a primarily discursive nature, rather than having to do with the actionality or some temporal/physical contiguity/continuity of the event to the utterance time. Most crucially, Portner notes that the experiential and existential perfects of the kind of (8) “provide evidence for something, not that it indicates any results” (2003, 464; cf. Rubovitz 1999 about the semantic-pragmatic correspondence between existential/experiential perfects and evidential reasoning).

The notion of discontinuity to the present becomes an important element of further semantic and grammatical reanalysis of V-过 *guo*. At this point in time, invited inferences being conveyed by the speaker/writer can be semantically and pragmatically associated with some reliability behind the proposition, whereby the truthfulness of P becomes markedly “at-issue” (Faller 2002; Tantucci 2016a, 2016b). In fact, due to the inherent anti-resultativity of the construction, an event marked with 过 *guo* is necessarily communicated either in the form of personal experience or as a piece of interpersonally shared knowledge (Tantucci 2015a). Crucially, earliest usages of V-过 *guo* as an experiential perfect seem to be limited to collocations with animate subjects, mental verbs or verbs profiling the syntactic subject’s personal experience in the past (Cao 1995; Lin 2004; Liu 2009, 231). However, Tantucci (2013, 224-5; 2015a) notes that during the Qing dynasty (1644-1912 AD) V-过 *guo* undergoes a new stage of semantic and grammatical reanalysis. This is a stage where V-过 *guo* collocates with subjectless or impersonal constructions with a new interpersonal evidential (IE) meaning. At this stage, V-过 *guo* is no longer used to mark an event in the form of an animate subject’s passed experience, but rather as a piece of knowledge shared by the speaker/writer together with a generic third party in society. Tantucci (2013, 2015a) notes that this trend is confirmed by the rise of the subjectless construction 发生过 *fāshēng-guo* ‘it happened before that’, as the valency of 发生 *fāshēng* in Mandarin normally does not include an experiencer. Earliest collocations of this verb with 过 *guo* are a clear sign of new evidential reanalysis of the chunk. Something similar is at stake for the verb 有 *yǒu* ‘to exist, to be there’, expressing an existential meaning rather than a possessive one. Early evidential usages of 有过 *yǒu-guo* ‘there has been before’ in the PKU-CCL-COR-

PUS<sup>6</sup> also date back to the Qing dynasty:

9. 这一天城里的街道,居然也打扫干净了,只怕从有上海城以来,也不曾有过这个干净的劲儿。(CCL / 清 Qīng / 二十年目睹之怪现状 èrshínián mùdǔ zhī guài xiànzhuàng)
- zhè yī tiān chéng lǐ de jiēdào jūrán yě  
this one day city in SP street unexpectedly also  
dǎsǎo-gānjìng le zhǐ pà cóng yǒu Shànghǎi  
clean-up.clean PFV only be.afraid since exist Shanghai  
chéng yǐlái yě bùcéng yǒu-guo zhè ge gānjìng de  
city since also never exist-EVD this CLF clean SP  
jìn’er  
degree/energy  
‘On this day, streets in the city had unexpectedly been cleaned thoroughly; I am afraid since the existence of Shanghai, the city has never been this clean’.

In example (9), there is not an animate syntactic subject to which some past experience is ascribed. The speaker/writer is similarly not referring to his personal life, as s/he cannot have experienced the full history of the city of Shanghai. S/he is rather referring to a piece of information that could be confirmed by other members of his/her own community of practice, thus expressing a proposition bearing collective recognition (cf. Searle 2010). Usages such as the one above are defined as interpersonal evidentials (IE) since,<sup>7</sup> as while no specific source of evidence is encoded by the construction, a piece of information is marked as shared knowledge within a community of practice, ideally paraphrasable as *it is known that*.

After the 民国 Mínguó period (1912-1949), the PKU-CCL-CORPUS includes a fairly balanced collection of texts, which is no longer limited to fictional registers, but also includes factual prose from press, academic journals and biographies. In Tantucci (2015a), it is shown that it is precisely in these textual environments that evidential usages of V-过 *guo* become increasingly frequent. From (9) above, we can observe that it is precisely the anti-resultativity of V-过 *guo* that

<sup>6</sup> The PKU-CCL-CORPUS is one of the largest corpora of Mandarin Chinese available and includes both a balanced synchronic and a diachronic section of written language. The total size of corpus data is approximately 200 million Chinese characters. Texts written in traditional Chinese in PKU-CCL-CORPUS contain approximately 101 million Chinese characters (486 documents, 54 folders, 202,305,825 bytes), and the texts written in modern Chinese contain 115 million Chinese characters (157 documents, 23 folders, 229,700,435 bytes).

<sup>7</sup> E.g. Tantucci 2013, 2015a, 2015b, 2016a, 2016c, 2017a, 2017b, 2020; Tantucci, Wang 2020a; Arslan et al. 2014; Jarque, Pascual 2015; Brugman, Macaulay 2015; Guardamagna 2017; Van Olmen 2019.

prompts further speculations concerning the evidence behind the proposition. In this sense, all the evidence that is provided subsequently is pragmatically aimed at filling a ‘temporal gap’ between the event and the reference time.

The diagram below summarises the present data about the grammaticalisation pathway of the V-过 *guo* construction:

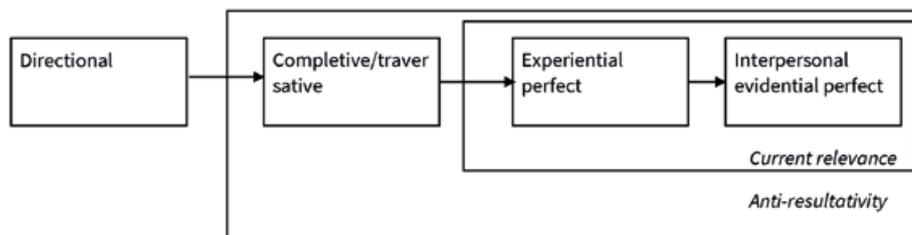


Figure 1 The pathway of change of the V-过 *guo* construction

As we can see from figure 1, a first step towards the grammaticalisation of V-过 *guo* is the transition from meaning expressing directional-ity of actions in space to a new aspectual meaning (completive/traversative) expressing that some action has been completed or ‘traversed’ by an animate subject [fig. 1]. This is an important stage of change of the construction, as the event is never conceptualised as entailing a resultative state. This element of anti-resultativity becomes crucial for further stages of change, as it persists (cf. Hopper 1982) in later usages conveying past experience of an animate subject. Anti-resultativity, in connection with discursive current relevance, contributes to express that an event has been experienced by the subject in a vague past, without specific reference to when this happened. Usages of the construction in the third person or in impersonal contexts contribute to a new evidential reading of the events that are referred to. In Fludernik, a distinction is made between “natural narrative proper” and “retelling of other people’s stories” (2006, 14). The crucial grammatical distinction between the two consists in first-person versus third-person narration (Norrick 2013a). Norrick notes that differences in first-person versus third-person narration underpin idiosyncratic features of the two types of narratives in relation to their form and function. They reflect differences in terms of teller perspective, story introduction, epistemic authority, and function (Norrick 2013a, 2013b). Frequent third-person-shift and impersonal usages are here considered as a very important factor contributing to the rise of novel interpersonal evidential polysemies of V-过 *guo*. Formal features as such, intersecting with specific text types and ‘contextual situatedness’, holistically affected the last stage of grammaticalisation of the construction from experientiality to interpersonal evidentiality.

## 4 A Corpus-Based Account of V-过 *guo* in Context

In this section, we provide the results of a corpus-based study from two synchronic corpora of Mandarin Chinese:

- the synchronic Lancaster Corpus of Mandarin Chinese (LCMC) (McEnery, Xiao 2004), a one-million-word balanced corpus designed as a Chinese match of the Freiburg-LOB Corpus of British English (FLOB), including texts from the years 1988 and 1992.
- The UCLA corpus of written Mandarin (Tao, Xiao 2012), also a one-million-word balanced corpus, designed as a match of the LCMC, including texts from 2000 to 2005.

The partition of texts of the LCMC is reported in the table below:

**Table 3** Text types of the LCMC

Code	Text category	Samples	Proportion
A	Press reportage	44	8.8%
B	Press editorials	27	5.4%
C	Press reviews	17	3.4%
D	Religion	17	3.4%
E	Skills/trades/hobbies	38	7.6%
F	Popular lore	44	8.8%
G	Biographies/essays	77	15.4%
H	Miscellaneous	30	6%
J	Science	80	16%
K	General fiction	29	5.8%
L	Mystery detective fiction	24	4.8%
M	Science fiction	6	1.2%
N	Martial arts fiction	29	5.8%
P	Romantic fiction	29	5.8%
R	Humour	9	1.8%
<b>Total</b>		<b>500</b>	<b>100%</b>

With this survey we aimed at answering three research questions:

- What is the distribution in the two corpora of evidential versus experiential usages of V-过 *guo*?
- Have there been any significant changes between the 1990s and the beginning of the 21st century in the partition of usages of V-过 *guo*?
- Is there a relationship between the formal and functional categories of V-过 *guo* and the textual environment in which it occurs?

## 4.1 Data Retrieval and Annotation

To answer each question it was necessary to design a solid annotation scheme that could grant a high inter-rater reliability (85%). We took into account a number of formal, functional and contextual dimensions, so that we could gather a holistic understanding of the behavioural profiles (cf. Gries 2010) of the construction. We therefore focused on: whether the polarity of the sentence was negative or positive; the corpus in which the chunk appeared; the verb (both as a token and as a type) collocating with 过 *guo*; the text-type where the V-过 *guo* was used; whether sentence final particles were present in the utterance; the type of the location of the force of each usage; the person of the verb (e.g. first singular, 3rd plural, and so on); and whether the function was evidential rather than experiential. The function of the construction was also the dependent variable of our analysis, and was based on the assessed set of criteria given in tables 1 and 2, in § 2. In table 4 below is given an example of one string of annotation:

**Table 4** Example of an annotated string of the usage of V-过 *guo* in the LCMC and UCLA

Function	Person	Verb	Verb_type	Ill_force	Sent_final	Text_type	Corpus	Polarity
Evid	3s	说	say	asser	N	B	LCMC	P

The utterance in table 4 has been annotated as an evidential usage, in the third person singular, collocating with the verb 说 *shuō*, which is a verb of saying (annotated as ‘say’). The illocutionary force of the utterance is assertive, it does not include sentence final particles, the text type corresponds to Press-editorials [tab. 3], the corpus in which it occurs is the LCMC and the polarity of the sentence is positive.

We retrieved all the usages including verbs with the highest MI<sup>3</sup> score from both the LCMC and the UCLA. Mutual Information (MI) expresses the extent to which observed frequency of co-occurrence differs from expected frequencies. It measures the strength of association among specific words or word types (in our case the strength of association of 过 *guo* with a preceding verb). The MI<sup>3</sup> score is used to rebalance MI score so as to give more weight to frequent words and less to infrequent words, by ‘cubing’ observed frequencies (cf. Oakes 1998, 171-2).

## 4.2 Data Analysis

After the retrieval of the top 15 verbs with the highest MI<sup>3</sup> score from both corpora, we first sought to answer our first two research ques-

tions, underpinning respectively the distribution in the two corpora of evidential versus experiential usages of V-过 *guo* and whether any changes between the 1990s and the beginning of the 21st century have occurred in the partition of usages of V-过 *guo*. We thus looked at the general distribution of experiential and evidential usages in the two corpora. We then performed a test of independence to assess whether there were significant mismatches based on chi-square and 'Pearson residuals'.

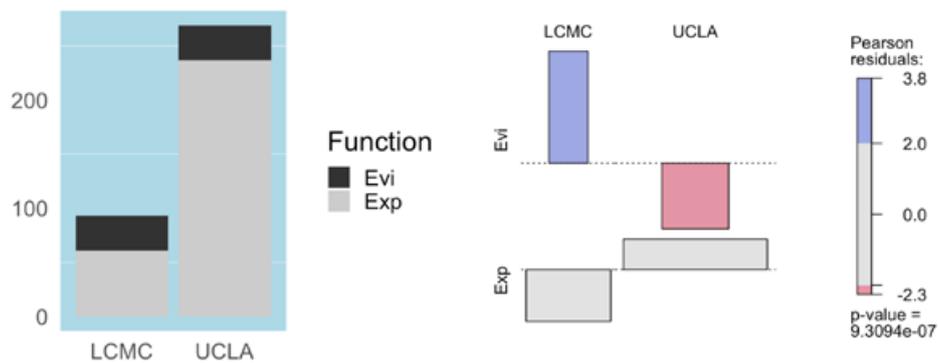


Figure 2 Distribution and test of independence of evidential vs experiential usages in LCMC and UCLA

The bar plot on the left hand side of figure 2 indicates a much more frequent usage of the V-过 *guo* construction [fig. 2]. It also shows a remarkably higher frequency of experiential usages (light grey) in contrast with the evidential ones (black) in the UCLA in comparison with the LCMC. This mismatch is statistically significant as indicated by the p-value ( $< 0.0005$ ) from the chi-square test, given at the bottom right hand side of figure 2. To explain, the plot on the right-hand side above is called assocplot (R package: *vcd*, cf. Hornik, Zeileis, Meyer 2006) and allows the analyst to visualise significant mismatches between observed and predicted frequencies deriving from a chi-square test. These mismatches are commonly called 'Pearson residuals'. If the observed frequency is greater than expected, the residual is positive. If the observed frequency is smaller than expected, it is then negative (Levshina 2015, 218). A blue colour (if any) indicates a significantly positive mismatch, whilst a red colour (if any) indicates a negative one, while the width of the bars is based on frequency.

From figure 2 we can clearly conclude that the frequency of evidential usages of V-过 *guo* is significantly higher in the LCMC corpus in comparison with the UCLA. This first result is not an obvious one. If we consider the relatively recent development of evidential functions of V-过 *guo*, one may expect it to progressively increase throughout the decade in between the LCMC and the UCLA. Quite

the opposite emerges from figure 2, as it is the experiential function the one that increases dramatically. This tendency supports the idea that constructional change and grammaticalisation are not necessarily incremental (e.g. Tantucci, Culpeper, Di Cristofaro 2018; Tantucci, Di Cristofaro 2019). Once a division of labour among functions of one construction is established, the frequency of comparatively more recent usages (such as the case of V-过 *guo* used as an evidential) is not necessarily going to further increase at the expense of comparatively older ones (e.g. V-过 *guo* used as an experiential).

It is now time to bring to the fore the role of context and text types in the encoding of evidential rather than experiential functions of V-过 *guo*. To begin with, we plotted a multiple correspondence analysis (MCA) (e.g. Nenadic, Greenacre 2007) on a two-dimensional plane. In this model, associations among variables are measured by calculating the chi-square distance between different categories of the variables and between observations. These associations are then represented graphically as a map, which eases the interpretation of the structures in the data: the closer the distance between variables, the stronger the statistical correspondence (Levshina 2015).

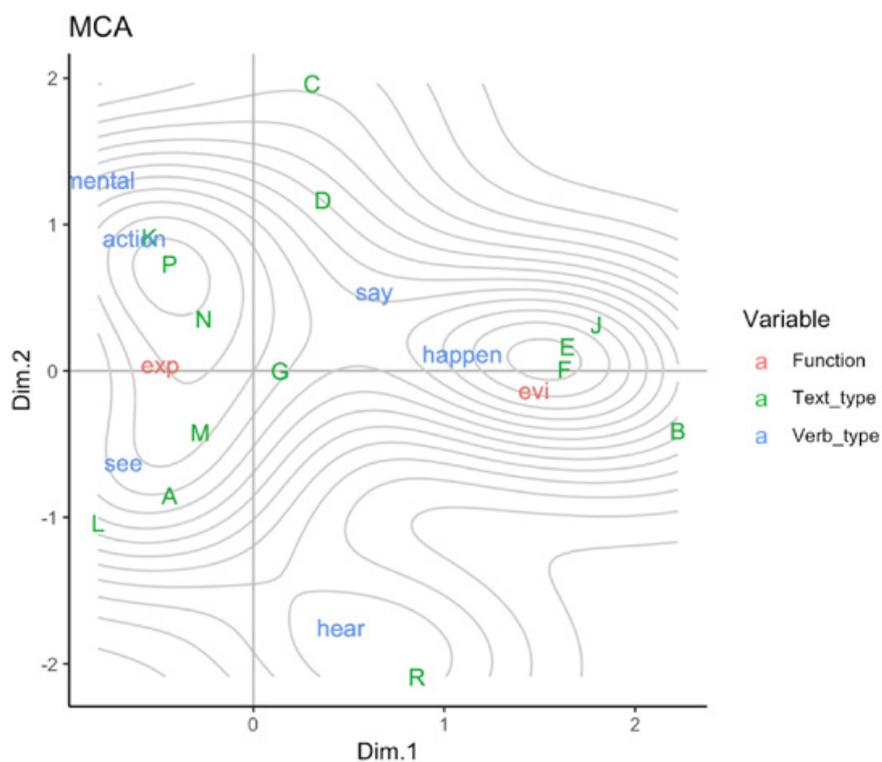


Figure 3 MCA of the relationship between function of 过 *guo*, text types and verb types

In the plot above the two dimensions represent 84.1% of variation among the three variables, which is a good approximation for MCA visualisation (Levshina 2015, 382). What counts for the interpretation of the data is the degree to which Function (i.e. Experiential vs Evidential), Text Type and Verb Type cluster together, therefore indicating a large-scale convergence in the way people use the V-过 *guo* construction, pragmatically, semantically and in contextually situated text types.

We can first note a clear division between the left and the right hand side of the plot, with two distinct clusters including text types (green) and verb types (blue) around respectively experiential and experiential usages (red). More specifically, at the left-hand side of the map there are experiential functions of V-过 *guo*, in turn attracting a different set of text types and verb types. More specifically, experientials are strongly attracted to verbs of action, or physical perception, such as 见 *jiàn* 'to see' or 看 *kàn* 'to watch', and mental verbs such as 想 *xiǎng* 'to think/plan'. These tend to form a cluster with text types K (General fiction), P (Romantic fiction), N (Martial arts fiction), M (Science fiction), A (Press reportage), L (Mystery detective fiction), and G (Biographies/essays). Most of these textual environments are fictional, whereby emotions and distinctive features of characters are often expressed through reference to their past experiences. The only exception regards A (Press reportage), which is undoubtedly a factual genre, yet also strongly based on a narrative stance of past events, which are very often experienced by the reporter or by other people who are being interviewed. Consider the extract below from the LCMC:

10. 这个问题是个很大的问题，因为我是从车间里滚出来的，发电机我也开过，上海二六轰炸的时候，我开发电机，我当厂长，我都是干过活的。(LCMC / A: Press reportage)

zhè ge wèntí shì ge hěn dà de wèntí  
 this CLF question be CLF very big SP question  
 yīnwèi wǒ shì cóng chējiān lǐ gǔn-chūlái de fādiànjī  
 because I be from workshop in work-out SP generator  
 wǒ yě kāi-**guo** Shànghǎi èrliùhōngzhà de  
 I also operate-**EXP** Shanghai February.Sixth.Incident SP  
 shíhòu wǒ kāi fādiànjī wǒ dāng chǎng zhǎng wǒ  
 time I operate generator I be factory director I  
 dōu shì gàn-**guo** huó de  
 all be do-**EXP** work EMP

'This is a very big question, because I used to work in a workshop and operated generators before; at the time of the February Sixth Incident in Shanghai, I operated generators and was a factory director—I was indeed engaged in my work'.

In the case above, the narrator is being interviewed about his previous experience working in a factory in Shanghai. This is a very interesting contextual environment. In fact, the usage of the construction is clearly experiential, yet, in this and similar contextual environments, someone’s personal experience is not shared merely to establish empathy among interlocutors, but more specifically to count as evidence about some broader factual information that has been reported by the interviewer. Nonetheless, interpersonal evidential pragmatic markers, such as 据了解 *jù liǎojiě* ‘it is understood that’, 好像 *hǎoxiàng* ‘apparently’, 众所周知 *zhòngsuǒzhōuzhī* ‘as everyone knows’, would not be compatible with this usage, which indicates that V-过 *guo* in (10) can still be considered as prominently experiential.

Back to the map, we can see that evidential polysemies are rather attracted to verbs of saying (e.g. 说 *shuō* or 讲 *jiǎng*) or verbs inherently expressing the occurrence of some event, such as 出现 *chūxiàn* ‘to appear’, 发生 *fāshēng* ‘to happen’, 有 *yǒu* ‘to exist, to occur’, and so on. The convergence of these verb types and evidential usages of 过 *guo* is at stake in texts such as E (Skills/trades/hobbies), F (Popular lore), J (Science) and B (Press editorials). The latter all tend to be geared to registers whereby information needs to be reported as a piece of evidence, rather than some past event that contribute to shape the personality or the personal history of a specific persona/character. In this case, events are presented to the reader as facts that can be potentially verified. The per-locutionary effects of these usages are not the ones of getting to know someone better, but rather to inform the reader of a piece of socially shared knowledge.

11. 自上个世纪七十年代开始, 有过四次较大的 METI 项目, 但每一次都有人站出来反对。(UCLA / J: Science)

zì	shàng	ge	shìjì	qīshíniándài	kāishǐ	yǒu- <b>guo</b>	sì
from	last	CLF	century	70s	begin	have-EVD	four
cì	jiào	dà	de	METI	xiàngmù	dàn	měi yí
time	relatively	big	SP	METI	project	but	every one
cì	dōu	yǒurén	zhàn-chūlái	fǎnduì			
time	always	someone	stand-out	object			

‘Starting from 1970s, there have been four relatively big METI projects, but every time there was always someone standing out to object’.

In (11) above, the stance of the speaker/writer is not centred on the identity of a specific persona, rather s/he uses 过 *guo* to report a piece of documented information that entails collective recognition, as adverbials of the kind of 据了解 *jù liǎojiě* ‘it is understood that’, 好像 *hǎoxiàng* ‘apparently’, 众所周知 *zhòngsuǒzhōuzhī* ‘as everyone knows’ would be perfectly idiomatic with this usage. This kind of usage is grounded in interpersonal evidentiality and is significantly associated with text types such as scientific essays or reports, as in the case above.

### 4.3 Evidential vs Experiential Categorisation in Context

Significant data-driven intersections of pragmatic, formal and contextual features are elsewhere defined as illocutional concurrences (IC) (cf. Tantucci, Wang 2018; 2020a; 2020b). IC are crucial to show that grammatical meaning is not independent from the pragmatic stance adopted by the interlocutors as well as the ‘contextual situatedness’ in which the speech event takes place.

This point is particularly evident in the last analysis of this paper below. In this case we plotted a conditional inference tree model (cf. Hothorn, Hornik, Zeileis 2006; Tagliamonte, Baayen 2012) gathering unbiased corpus-driven convergences of form, meaning, context and pragmatic effects, all contributing to the spontaneous encoding of either experiential or evidential usages of V-过 *guo*. We took in to account the function of the construction, the polarity (from table 1 in § 2 we can see how experiential usages of 过 *guo* are generally agreed to occur with negative polarity or in questions), the illocutionary force (whether the speech act occurs as a modalised evaluation – e.g. Tantucci, Wang 2018 – a question or a bare assertion), and the presence of sentence final particles, which could shed light on whether the construction is used in questions, or whether the utterance is characterised by modalised elements of intersubjectivity occurring at sentence periphery (cf. Traugott 2012, 2016; Tantucci 2017a, 2017b, 2020, forthcoming; Tantucci, Wang 2018, 2020a, 2020b).

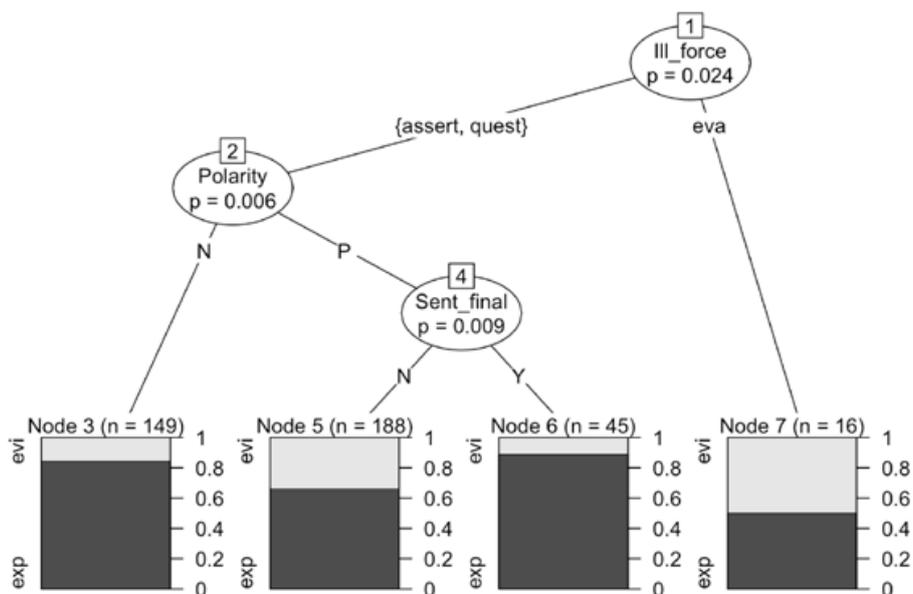


Figure 4 Conditional inference tree IC of evidential vs experiential usages of V-过 *guo*

The plot above is obtained with the ‘ctree’ function of the R package ‘party’ (Levshina 2015, 291). It is important to emphasise that the tree above has nothing to do with a generative one. Conditional dependencies among variables in figure 4 exclusively depend on statistical significance (the higher the node, the more significant the ‘conditional decision’) [fig. 4]. The descending order of each split computationally simulates a conditional ‘decision’ made by the speaker/writer based on degrees of significance of each covariant that comes into play when a speech act including experiential or evidential functions of 过 *guo* is realised. In other words, the plot above is completely usage-based and computes holistically probabilities among semantic, pragmatic together with formal variables. The p-value of each ‘decision’ is reported under each variable before every split (e.g. *ill\_force*  $p = 0.024$  at the top of the tree).

From the above we can see that one interesting IC has to do with illocutionary force being either assertive or interrogative, and the polarity being negative. Convergence of these two features is significantly ( $p = 0.006$ ) connected to experiential usages of V-过 *guo*.

12. 婉姐, 我没见过什么世面, 啥也不懂。(UCLA / M: Science fiction)

*wǎn jiě wǒ méi jiàn-guo shénme shìmiàn shá*  
 Wan sister I NEG see-EXP what life what  
*yě bù dǒng*  
 also NEG know  
 ‘Wan, I haven’t seen much of life, and I don’t know anything’.

In (12) is given a negative assertion of the speaker referring to his/her past experience as a specific persona. This usage is distinctly narrative and occurs in a fictional text (M: Mystery fiction).

Another interesting IC has to do with presence of sentence final particles, which is not preponderant in neither of the two usages, yet still significantly more salient when experiences are narrated or enquired by the speaker/writer.

13. 你在海上清理垃圾的时候, 你想过有一天你会死在这件事上吗? (UCLA / L: Mystery detective fiction)

*nǐ zài hǎishàng qīnglǐ lājī de shíhòu nǐ*  
 you at sea clean.up rubbish SP time you  
*xiǎng-guo yǒu yī tiān nǐ huì sǐ zài zhè jiàn*  
 think-EXP have one day you would die at this CLF  
*shì-shàng ma*  
 thing-on Q  
 ‘While cleaning up rubbish at sea, have you ever thought about one day you would die because of this?’

As (13) illustrates, experiential usages of the construction tend to occur in dialogic contexts and thus are more likely to be attracted to sentence final particles such as the interrogative 吗 *ma* above. This IC is significantly absent when evidential grounding is at play, as statements are given assertively as reported, potentially verifiable pieces of information. This underpins a clear division of labour between the two functions, one hinging on affective engagement with an animate subject’s past experiences, the other being distinctively uttered to mark a proposition as an intersubjectively reliable piece of information. This case study has shed light on the holistic and multimodal factors that concur to the differentiation of the evidential vs experiential senses of the V-过 *guo* construction.

What emerged from this analysis is that speakers differentiate experiential and evidential meanings based on the context in which the construction is used, the illocutionary force of the linguistic act, the polarity and the presence of sentence final particles. This entails that meaning disambiguation occurs simultaneously at grammatical, semantic, pragmatic, and situational levels and results from the repeated ascription of a linguistic function to the situation type in which a lexeme is used. The present usage-based analysis of V-过 *guo* is relevant to a broader discussion about linguistic categorisation. In fact, from a usage-based perspective, categorisation is a process that arises as a result of single token instantiations of meaning. What this analysis suggests, is that speakers’ ability to identify analogies and similarities among instantiations of meaning cannot be detached from the physical or sociocultural space in which each occurrence takes place. Put simply, context and conventions of usage inform grammatical categorisation. The role of context is thus a crucial one for conceptualisers’ ability to establish categories at increasing levels of schematicity and grammatical specialisation. In this sense, the diachronic notion of upward strengthening regards the increased abstraction of a linguistic form leading to the progressive formation of grammatical categories (Hilpert 2015; Tantucci, Di Cristofaro 2019). When the latter reaches highly schematic nodes in a constructional network, it is then possible that context and ‘situatedness’ become progressively detached from schematic heuristics. This is the case of very abstract schemas such as transitivity, di-transitivity or resultativity, or even aspect or tense, in which conceptualisations of meaning are almost entirely schematic, and not metonymically attached to contextual state of affairs and sociocultural conventions. However, most linguistic functions are the result of a combination of single instantiations and schematic representation, and context does indeed play a crucial role in the speakers’ ability to identify and express categorial membership. This is precisely the case of the evidential functions of V-过 *guo* in Mandarin Chinese, as speakers’ ability to ascribe the relatively schematic notion of ‘shared

knowledge’ to the construction is inherently determined by the register and the sociocultural context in which those utterances occur (cf. text types of the kind of J, E, F, B in figure 3). This clearly entails that some degree of entrenchment (e.g. Langacker 1987; Schmid 2017; Tantucci, Culpeper, Di Cristofaro 2018; Tantucci, Di Cristofaro 2019) underpins the recurrent usage of 过 *guo* specifically in connection with text types that allow speakers to infer an evidential meaning rather than an experiential one. In turn, this means that entrenchment as such is also a process that is inherently context-driven and socioculturally situated, and not simply arising as the result of frequent co-occurrence of two or more items independently from contextual situatedness and pragmatic conventions (cf. Terkourafi 2015).

## 5 Conclusions

In this paper we argued that polysemy and categorial membership cannot be detached from ‘contextual situatedness’. While we maintain that, at very high levels of abstraction, sociocultural context does not play a role for the identification of grammatical categories, we also suggest that the progressive formation of those categories is inherently determined by the sociocultural instantiations in which a particular form tends to occur. Entrenchment is therefore experienced as a socioculturally situated phenomenon, and the contextual and co-textual environment where a particular form occurs is a crucial factor for identifying a division of labour among its usages. In this paper we provided a detailed case-study centred on the V-过 *guo* construction in Mandarin Chinese. We showed that a clear division of labour is at stake among experiential and evidential usages of the construction. This categorial separation occurs as a result of features underpinning form, usage and ‘contextual situatedness’. Evidentiality in Mandarin is therefore a category that emerges significantly from specific intersections among these three dimensions and from distinctive illocutional concurrences of conventionalised behaviour.

## Bibliography

- Aikhenvald, A.Y. (2004). *Evidentiality*. Oxford: Oxford University Press.
- Arslan, S. et al. (2014). “Finite Verb Inflections for Evidential Categories and Source Identification in Turkish Agrammatic Broca’s Aphasia”. *Journal of Pragmatics*, 70, 165-81. <https://doi.org/10.1016/j.pragma.2014.07.002>.
- Austin, J.L. (1962). *How to Do Things with Words*. Oxford: Oxford University Press.

- Benveniste, E. [1958] (1971). "Subjectivity in Language". Transl. by M.E. Meek. *Problems in General Linguistics*. Coral Gables: University of Miami Press, 223-30.
- Brugman, C.M.; Macaulay, M. (2015). "Characterizing Evidentiality". *Linguistic Typology*, 19(2), 201-37. <https://doi.org/10.1515/lingty-2015-0007>.
- Bybee, J.; Perkins, R.; Pagliuca, W. (1994). *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Cao G. 曹广顺 (1995). *Jindai Hanyu Zhuci 近代汉语助词 (The Auxiliary Particles of Modern Chinese)*. Beijing: Yuwen chubanshe.
- Capone, A. (2005). "Pragmemes (a Study with Reference to English and Italian)". *Journal of Pragmatics*, 37(9), 1355-71. <https://doi.org/10.1016/j.pragma.2005.01.013>.
- Capone, A. (2010). "On the Social Practice of Indirect Reports (Further Advances in the Theory of Pragmemes)". *Journal of Pragmatics*, 42(2), 377-91. <https://doi.org/10.1016/j.pragma.2009.06.013>.
- Carey, K. (1994). *Pragmatics, Subjectivity and the Grammaticalization of the English Perfect* [PhD Dissertation]. San Diego: University of California.
- Chao, Y.R. (1968). *A Grammar of Spoken Chinese*. Oakland: University of California Press.
- Chappell, H. (2001). "A Typology of Evidential Markers in Sinitic Languages". Chappell, H. (ed.), *Chinese Grammar. Synchronic and Diachronic Perspectives*. New York: Oxford University Press, 56-85.
- Chen Q. 陈前瑞 (2008). *Hanyu timao yanjiu de leixingxue shiye 汉语体貌研究的类型学事业 (A Study on the Aspectual System of Mandarin from a Typological Perspective)*. Beijing: The Commercial Press.
- Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.
- Dahl, Ö. (ed.) (2000). *Tense and Aspect in the Languages of Europe*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110197099>.
- Dahl, Ö.S. (1985). *Tense and Aspect Systems*. Blackwell: Oxford.
- Dahl, Ö.S.; Hedin, E. (2000). "Current Relevance and Event Reference". Dahl 2000, 385-402. <https://doi.org/10.1515/9783110197099.3.385>.
- Dai Y. 戴耀晶 (1997). *Xiandai Hanyu shiti xitong yanjiu 汉语时体系统研究 (A Study of Aspect in Modern Chinese)*. Hangzhou: Zhejiang Educational Press.
- Faller, M. (2002). *Semantics and Pragmatics of Evidentials in Cuzco Quechua* [PhD Dissertation]. Stanford (CA): Stanford University.
- Fludernik, M. (2006). *Towards a 'Natural' Narratology*. New York: Routledge.
- Formato, F.; Tantucci, V. (2020). "Uno. A Corpus Linguistic Investigation of Intersubjectivity and Gender". *Journal of Language and Discrimination*, 4(1), 51-73. <https://doi.org/10.1558/jld.40129>.
- Guardamagna, C. (2017). "Reportative Evidentiality, Attribution and Epistemic Modality. A Corpus-Based Diachronic Study of Latin Secundum NP ('According to NP')". *Language Sciences*, 59, 159-79. <https://doi.org/10.1016/j.langsci.2016.09.001>.
- Gries, S.T. (2010). "Behavioral Profiles. A Fine-Grained and Quantitative Approach in Corpus-Based Lexical Semantics". *The Mental Lexicon*, 5(3), 323-46. <https://doi.org/10.1075/ml.5.3.04gri>.
- Haugh, M. (2012). "Conversational Interaction". Allan, K.; Jaszczolt, K.M. (eds), *The Cambridge Handbook of Pragmatics*. Cambridge: Cambridge University Press, 251-74. <https://doi.org/10.1017/CB09781139022453.014>.

- Hilpert, M. (2015). "From *Hand-Carved* to *Computer-Based*. Noun-Participle Compounding and the Upward Strengthening Hypothesis". *Cognitive Linguistics*, 26(1), 113-47. <https://doi.org/10.1515/cog-2014-0001>.
- Hopper, P.J. (1982). "Aspect Before Discourse and Grammar". Hopper, P.J. (ed.), *Tense-Aspect. Between Semantics and Pragmatics*. Amsterdam: John Benjamins, 3-18.
- Hornik, K.; Zeileis, A.; Meyer, D. (2006). "The Strucplot Framework. Visualizing Multi-Way Contingency Tables with VCD". *Journal of Statistical Software*, 17(3), 1-48. <https://doi.org/10.18637/jss.v017.i03>.
- Hothorn, T.; Hornik, K.; Zeileis, A. (2006). "Unbiased Recursive Partitioning. A Conditional Inference Framework". *Journal of Computational and Graphical Statistics*, 15(3), 651-74. <https://doi.org/10.1198/106186006x133933>.
- Jarque, M.J.; Pascual, E. (2015). "Direct Discourse Expressing Evidential Values in Catalan Sign Language". *eHumanista. Journal of Iberian Studies*, 8, 421-45. [https://www.ehumanista.ucsb.edu/sites/secure.lsit.ucsb.edu/span.d7\\_eh/files/sitefiles/ivitra/volume8/4.monograficIV/5\\_JarquePascual.pdf](https://www.ehumanista.ucsb.edu/sites/secure.lsit.ucsb.edu/span.d7_eh/files/sitefiles/ivitra/volume8/4.monograficIV/5_JarquePascual.pdf).
- Johanson, L. (2000). "Viewpoint Operators in European Languages". Dahl 2000, 27-188. <https://doi.org/10.1515/9783110197099.1.27>.
- Langacker, R.W. (1987). *Foundations of Cognitive Grammar. Theoretical Prerequisites*, vol. 1. Stanford (CA): Stanford University Press.
- Langacker, R.W. (2008). *Cognitive Grammar. A Basic Introduction*. Oxford: Oxford University Press.
- Levshina, N. (2015). *How to Do Linguistics with R. Data Exploration and Statistical Analysis*. Amsterdam: John Benjamins.
- Li, C.; Thompson, S.A. (1981). *Mandarin Chinese. A Functional Reference Grammar*. Berkeley: University of California Press.
- Li, D.C.S. (2011). "'Perfective Paradox': A Cross-linguistic Study of the Aspectual Functions of *-guo* in Mandarin Chinese". *Chinese Language and Discourse*, 2(1), 23-57. <https://doi.org/10.1075/cld.2.1.02li>.
- Lin, J.-W. (2006). "Time in a Language without Tense. The Case of Chinese". *Journal of Semantics*, 23, 1-53. <https://doi.org/10.1093/jos/ffh033>.
- Lin, J.-W. (2007). "Predicate Restriction, Discontinuity Property and the Meaning of the Perfective Marker *Guo* in Mandarin Chinese". *Journal of East Asian Linguistics*, 16(3), 237-57. <https://doi.org/10.1007/s10831-007-9013-5>.
- Lin X. 林新年 (2004). "Shixi Tang Song shiqi de 'guo' yufahua jincheng chihuan de yuanyin" 试析唐宋时期的“过”语法化进程迟缓的原因 (An Analysis of the Slowdown in the Grammaticalisation Process of *guo* During the Tang and Song Periods). *Yuyan Kexue*, 6, 42-52.
- Liu J. 刘坚 (2009). "Shitai zhuci de yanjiu yu VO guo" 时态助词的研究与 VO 过 (A Study on Tense Particles and the VO *guo* Construction). Feng, L 冯力; Yang, Y. 杨永龙; Zhao, C. 赵长才 (eds), *Hanyu shiti de lishi yanjiu* 汉语时体的历时研究 (Diachronic Study on the Tense and Aspect System of Chinese). Beijing: Yuwen Chubanshe, 229-34.
- McCawley, J.D. (1971). "Tense and Time Reference in English". Fillmore, C.; Langendoen, T. (eds), *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston, 96-113.
- McEney, A.; Xiao, Z. (2004). "The Lancaster Corpus of Mandarin Chinese. A Corpus for Monolingual and Contrastive Language Study". *Religion*, 17, 3-4.
- Mey, J.L. (2001). *Pragmatics. An Introduction*. 2nd ed. Oxford: Blackwell.

- Mey, J.L. (2010). "Reference and the Pragmeme". *Journal of Pragmatics*, 42(11), 2882-8. <https://doi.org/10.1016/j.pragma.2010.06.009>.
- Mushin, I. (2001). *Evidentiality and Epistemological Stance*. Amsterdam: John Benjamins.
- Narrog, H. (2009). *Modality in Japanese. The Layered Structure of the Clause and Hierarchies of Functional Categories*, vol. 109. Amsterdam: John Benjamins.
- Nenadic, O.; Greenacre, M. (2007). "Correspondence Analysis in R, with Two- and Three-Dimensional Graphics. The ca Package". *Journal of Statistical Software*, 20(3). <https://doi.org/10.18637/jss.v020.i03>.
- Norrick, N.R. (2013a). "Narratives of Vicarious Experience in Conversation". *Language in Society*, 42(4), 385-406. <https://doi.org/10.1017/S0047404513000444>.
- Norrick, N.R. (2013b). "Stories of Vicarious Experience in Speeches by Barack Obama". *Narrative Inquiry*, 23(2), 283-301. <https://doi.org/10.1075/ni.23.2.04nor>.
- Oakes, M.P. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University.
- Olsen, M.B. (1997). *A Semantic and Pragmatic Model of Lexical and Grammatical Aspect*. New York: Garland Press.
- Peterson, T. (2010). *Epistemic Modality and Evidentiality in Gitksan at the Semantics-Pragmatics Interface* [Phd Dissertation]. Vancouver (BC): University of British Columbia.
- Plungian, V.A.; van der Auwera, J. (2006). "Towards a Typology of Discontinuous Past Marking". *Language Typology and Universals*, 59(4), 317-49. <https://doi.org/10.1524/stuf.2006.59.4.317>.
- Portner, P. (2003). "The (Temporal) Semantics and (Modal) Pragmatics of the Perfect". *Linguistics and Philosophy*, 26(4), 459-510. <https://doi.org/10.1023/A:1024697112760>.
- Rubovitz, T. (1999). "Evidential-Existentials. The Interaction Between Discourse and Sentence Structure". *Journal of Pragmatics*, 31(8), 1025-40. [https://doi.org/10.1016/S0378-2166\(99\)00038-7](https://doi.org/10.1016/S0378-2166(99)00038-7).
- Rooryck, J. (2001a). "Evidentiality, Part I". *GLOT International*, 5(4), 125-33. <https://bit.ly/3qHhaTZ>.
- Rooryck, J. (2001b). "Evidentiality, Part II". *GLOT international*, 5(4), 161-8. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.461.2545&rep=rep1&type=pdf>.
- Schmid, H.-J. (2017). "A Framework for Understanding Linguistic Entrenchment and Its Psychological Foundations". Schmid, H.J. (ed.), *Entrenchment and the Psychology of Language Learning. How We Reorganize and Adapt Linguistic Knowledge*. Washington, DC: American Psychological Association; Berlin: Walter de Gruyter, 9-38. <https://doi.org/10.1515/9783110341423-002>.
- Searle, J.R. (1976). "A Classification of Illocutionary Acts". *Language in Society*, 5(1), 1-23. [https://sites.duke.edu/conversions/files/2014/09/Searle\\_Illocutionary-Acts.pdf](https://sites.duke.edu/conversions/files/2014/09/Searle_Illocutionary-Acts.pdf).
- Searle, J. (2010). *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford University Press.
- Smith, C. (1997). *The Parameter of Aspect*. 2nd ed. Dordrecht: Kluwer Academic Publishers.
- Squartini, M. (2012). "Evidentiality in Interaction. The Concessive Use of the Italian Future between Grammar and Discourse". *Journal of Pragmatics*, 44(15), 2116-28. <https://doi.org/10.1016/j.pragma.2012.09.008>.

- Tagliamonte, S.A.; Baayen, R.H. (2012). "Models, Forests, and Trees of York English. Was/Were Variation as a Case Study for Statistical Practice". *Language Variation and Change*, 24(2), 135-78. <https://doi.org/10.1017/S0954394512000129>.
- Tantucci, V. (2013). "Interpersonal Evidentiality. The Mandarin V-过 *guo* Construction and Other Evidential Systems beyond the 'Source of Information'". *Journal of Pragmatics*, 57, 210-30. <https://doi.org/10.1016/j.pragma.2013.08.013>.
- Tantucci, V. (2015a). "Traversativity and Grammaticalization. The Aktionsart of 过 *guo* as a Lexical Source of Evidentiality". *Chinese Language and Discourse*, 6(1), 57-100. <https://doi.org/10.1075/cld.6.1.03tan>.
- Tantucci, V. (2015b). "Epistemic Inclination and Factualization. A Synchronic and Diachronic Study on the Semantic Gradience of Factuality". *Language and Cognition*, 7(4), 590-90. <https://doi.org/10.1017/lang-cog.2014.34>.
- Tantucci, V. (2016a). "Textual Factualization. The Phenomenology of Assertive Reformulation and Presupposition during a Speech Event". *Journal of Pragmatics*, 101, 155-71. <https://doi.org/10.1016/j.pragma.2016.06.007>.
- Tantucci, V. (2016b). "Towards a Typology of Constative Speech Acts. Actions beyond Evidentiality, Epistemic Modality and Factuality". *Intercultural Pragmatics*, 13(2), 181-209. <https://doi.org/10.1515/ip-2016-0008>.
- Tantucci, V. (2016c). "The Multimodal Marking of Evidentiality. Pragmemes of Circumstantial Inference and Mandarin Written News Report". *Pragmemes and Theories of Language*, (6), 477-91. [https://doi.org/10.1007/978-3-319-43491-9\\_24](https://doi.org/10.1007/978-3-319-43491-9_24).
- Tantucci, V. (2017a). "From Immediate to Extended Intersubjectification. A Gradient Approach to Intersubjective Awareness and Semasiological Change". *Language and Cognition*, 9(1), 88-120. <https://doi.org/10.1017/lang-cog.2015.26>.
- Tantucci, V. (2017b). "An Evolutionary Approach to Semasiological Change. Overt Influence Attempts through the Development of the Mandarin 吧 *-ba* Particle". *Journal of Pragmatics*, 120, 35-53. <https://doi.org/10.1016/j.pragma.2017.08.006>.
- Tantucci, V. (2020). "From Co-actionality to Extended Intersubjectivity. Drawing on Language Change and Ontogenetic Development". *Applied Linguistics*, 41(2), 185-214. <https://doi.org/10.1093/applin/amy050>.
- Tantucci, V. (forthcoming). *Language and Social Minds. The Semantics and Pragmatics of Intersubjectivity*. Cambridge: Cambridge University Press.
- Tantucci, V.; Culpeper, J.; Di Cristofaro, M. (2018). "Dynamic Resonance and Social Reciprocity in Language Change. The Case of Good Morrow". *Language Sciences*, 68, 6-21. <https://doi.org/10.1016/j.langsci.2017.09.004>.
- Tantucci, V.; Di Cristofaro, M. (2019). "Entrenchment Inhibition. Constructional Change and Repetitive Behaviour Can Be in Competition with Large-Scale 'Recompositional' Creativity". *Corpus Linguistics and Linguistic Theory*, 16(3), 547-79. <https://doi.org/10.1515/cllt-2019-0017>.
- Tantucci, V.; Wang, A. (2018). "Illocutional Concurrences. The Case of Evaluative Speech Acts and Face-Work in Spoken Mandarin and American English". *Journal of Pragmatics*, 138, 60-76. <https://doi.org/10.1016/j.pragma.2018.09.014>.

- Tantucci, V.; Wang, A. (2020a). "Diachronic Change of Rapport Orientation and Sentence-Periphery in Mandarin". *Discourse Studies*, 22(2), 146-73. <https://doi.org/10.1177/1461445619893777>.
- Tantucci, V.; Wang, A. (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. <https://doi.org/10.1016/j.pragma.2020.05.011>.
- Tao, H.; Xiao, R. (2012). *The UCLA Chinese Corpus*. 2nd ed. Lancaster: University Centre for Computer Corpus Research on Language.
- Terkourafi, M. (2015). "Conventionalization. A New Agenda for Im/politeness Research". *Journal of Pragmatics*, 86, 11-18. <https://doi.org/10.1016/j.pragma.2015.06.004>.
- Traugott, E.C. (1999). "From Subjectification to Intersubjectification". Paper presented at the *Fourteenth International Conference on Historical Linguistics* (Vancouver, 9-13th August 1999).
- Traugott, E.C. (2003). "From Subjectification to Intersubjectification". Hickey, R. (ed.), *Motives for Language Change*. Cambridge: Cambridge University Press, 124-39. <https://doi.org/10.1017/cbo9780511486937.009>.
- Traugott, E.C. (2012). "Intersubjectification and Clause Periphery". *English Text Construction*, 5(1), 7-28. <https://doi.org/10.1075/bct.65.02trau>.
- Traugott, E.C. (2016). "On the Rise of Types of Clause-Final Pragmatic Markers in English". *Journal of Historical Pragmatics*, 17(1), 26-54. <https://doi.org/10.1075/jhp.17.1.02tra>.
- Traugott, E.C.; Dasher, R.B. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Van Olmen, D. (2019). "A Diachronic Corpus Study of Prenominal zo'n 'so a' in Dutch. Pathways and (Inter) Subjectification". *Functions of Language*, 26(2), 216-47. <https://doi.org/10.1075/foL.16017.van>.
- Vendler, Z. (1967). *Linguistics in Philosophy*. Ithaca (NY): Cornell University Press.
- Yeh, M. (1996). "An Analysis of the Experiential *guo*<sub>exp</sub> in Mandarin: A Temporal Quantifier". *Journal of East Asian Linguistics*, 5, 183-215. <https://doi.org/10.1007/BF00215072>.
- Xiao, R.; McEnery, T. (2004). *Aspect in Mandarin Chinese. A Corpus-Based Study*. Amsterdam: John Benjamins.
- Wu, J.-S. (2008). "Terminability, Wholeness and Semantics of Experiential *Guo*". *Journal of East Asian Linguistics*, 17(1), 1-32. <https://doi.org/10.1007/s10831-007-9018-0>.