RUSSIAN
JOURNAL OF LINGUISTICS

2024 Volume 28 No. 4

**Discourse-pragmatic markers
of (inter)subjective stance in Asian languages:
With special focus on Chinese etymons**

Guest Editors
Seongha RHEE, Bernd HEINE and Wenjiang YANG

**Founded in 1997
by the RUDN University**

Научный журнал

Издается с 1997 г.

RUSSIAN JOURNAL OF LINGUISTICS

Aims and Scope

The Russian Journal of Linguistics is a peer-reviewed international academic journal publishing research in Linguistics and related fields. It is international with regard to its editorial board, contributing authors and thematic foci of the publications.

The aims of the journal:

♦ to promote scholarly exchange and cooperation among Russian and international linguists and specialists in related areas of investigation;

♦ to disseminate theoretically grounded research and advance knowledge pertaining to the field of Linguistics developed both in Russia and abroad;

♦ to publish results of original research on a broad range of interdisciplinary issues relating to language, culture, cognition and communication;

♦ to cover scholarly activities of the Russian and international academia.

As a Russian journal with international character, it aims at discussing relevant intercultural/linguistic themes and exploring general implications of intercultural issues in human interaction in an interdisciplinary perspective. The most common topics include *language and culture, comparative linguistics, sociolinguistics, psycholinguistics, cognitive linguistics, pragmatics, discourse analysis, intercultural communication, and theory and practice of translation*. In addition to research articles, the journal welcomes book reviews, literature overviews, conference reports and research project announcements.

The Journal is published in accordance with the policies of *COPE* (*Committee on Publication Ethics*) **http://publicationethics.org**. The editors are open to thematic issue initiatives with guest editors.

Further information regarding notes for contributors, subscription, open access and back volumes is available at **http://journals.rudn.ru/linguistics**.

E-mail: **lingj@rudn.ru**

Цели и тематика

Журнал Russian Journal of Linguistics – периодическое международное рецензируемое научное издание в области междисциплинарных лингвистических исследований. Журнал является международным как по составу редакционной коллегии и экспертного совета, так и по авторам и тематике публикаций.

Цели журнала:

♦ способствовать научному обмену и сотрудничеству между российскими и зарубежными лингвистами, а также специалистами смежных областей;

♦ знакомить читателей с новейшими направлениями и теориями в области лингвистических исследований, разрабатываемых как в России, так и за рубежом, и их практическим применением;

♦ публиковать результаты оригинальных научных исследований по широкому кругу актуальных лингвистических проблем междисциплинарного характера, касающихся языка, культуры, сознания и коммуникации;

♦ освещать научную деятельность как российского, так и международного научного сообщества.

Будучи международным по своей направленности, журнал нацелен на обсуждение теоретических и практических вопросов, касающихся взаимодействия культуры, языка и коммуникации. Особый акцент делается на междисциплинарные исследования. Основные рубрики журнала: *язык и культура, сопоставительное языкознание, социолингвистика, психолингвистика, когнитивная лингвистика, прагматика, анализ дискурса, межкультурная коммуникация, теория и практика перевода*. Кроме научных статей публикуется хроника научной жизни, включающая рецензии, научные обзоры, информацию о конференциях, научных проектах.

Перечень отраслей науки и групп специальностей научных работников в соответствии с номенклатурой ВАК РФ: Отрасль науки: 10.00.00 – филологические науки; Специальности научных работников: 10.02.01 – русский язык, 10.02.04 – германские языки, 10.02.05 – романские языки, 10.02.19 – теория языка, 10.02.20 – сравнительно-историческое, типологическое и сопоставительное языкознание.

Журнал строго придерживается международных стандартов публикационной этики, сформулированных в документе *COPE* (*Committee on Publication Ethics*) **http://publicationethics.org**.

Правила оформления статей, архив и дополнительная информация размещены на сайте: **http://journals.rudn.ru/linguistics**.

Электронный адрес: **lingj@rudn.ru**

Discourse-pragmatic markers of (inter)subjective stance in Asian languages: With special focus on Chinese etymons

CONTENTS

BOOK REVIEWS

Research article / Научная статья

Discourse-pragmatic markers of (inter)subjective stance in Asian languages

Bernd HEINE[1] ✉, **Wenjiang YANG[2]** and **Seongha RHEE[3]**

[1]University of Cologne, Cologne, Germany
[2]University of Macau, Macao SAR, China
[3]Faculty of Liberal Arts, Mahidol University, Nakhon Pathom, Thailand
[3]Hankuk University of Foreign Studies, Seoul, Republic of Korea
✉srhee@hufs.ac.kr

Abstract

This special issue is concerned with languages belonging to the Sinosphere, a region where China played an important geo-political and cultural leadership role. It aims to trace areal effects that the impact of Chinese had on the languages of the region over centuries. It deals with a number of words of Chinese origin used in Japanese, Korean, Vietnamese and Thai, as well as Chinese and investigates theoretically significant issues related to language contact, discourse-pragmatic aspects of language change, and socio-cultural influence on language development, among others, as exemplified in the development of discourse markers from their earlier lexical expressions originating from Chinese etyma. The nine contributions presented in this special issue have a number of things in common, in particular the following. First, they deal in some way or other with areal effects that the impact of Chinese had on these languages over centuries. Second, their goal is to achieve linguistic reconstruction, tracing present-day patterns of language use back to earlier states of language use. Third, linguistic reconstruction is restricted to linguistic material that was responsible for the rise and development of new patterns of discourse organization. Fourth, the tool most commonly employed for achieving reconstructions is grammaticalization theory. And finally, a central concern of the authors contributing to this special issue is with understanding the role played by discourse markers in linguistic development – how they arose and developed into what they are today. This special issue demonstrates that the languages figuring in it have received substantial influence from Chinese through written texts.

Keywords: *Chinese, cooptation, discourse marker, grammaticalization, language contact, Sinosphere*

Дискурсивно-прагматические маркеры (интер)субъективной позиции в азиатских языках

Бернд ХАЙНЕ¹ ⬡ ✉, **Вэньцзян ЯН²** ⬡, **Санха РИ³** ⬡

¹Кельнский университет, Кельн, Германия
²Университет Макао, Макао САР, Китай
³Факультет гуманитарных наук, Университет Махидол, Накхонпатхом, Тайланд;
Университет иностранных исследований Ханкук, Сеул, Республика Корея
✉srhee@hufs.ac.kr

Аннотация

Этот специальный выпуск посвящен языкам, принадлежащим к синосфере – региону, в котором Китай играл важную геополитическую и культурную роль. Его цель – проследить ареальные последствия влияния китайского языка на языки региона на протяжении веков. В выпуске рассматривается ряд слов китайского происхождения, употребляемых в японском, корейском, вьетнамском и тайском языках, а также их изменения. Исследуются теоретически значимые вопросы, связанные с языковыми контактами, дискурсивно-прагматическими аспектами изменения языка и влиянием социокультурных факторов на развитие языка. Данные процессы показаны на примере развития дискурсивных маркеров из более ранних лексических единиц китайского происхождения. Девять статей, представленных в этом специальном выпуске, имеют ряд общих черт. Во-первых, они так или иначе имеют дело с ареальным влиянием, которое китайский язык оказывал на эти языки на протяжении веков. Во-вторых, их целью является лингвистическая реконструкция и сопоставление современных моделей использования языка с более ранними стадиями их использования. В-третьих, лингвистическая реконструкция ограничивается только тем языковым материалом, на котором возникли и развились новые модели организации дискурса. В-четвертых, наиболее часто используемым инструментом для осуществления реконструкций является теория грамматикализации. И, наконец, основное внимание авторов этого специального выпуска уделяется роли дискурсивных маркеров в развитии языка. Ими ставится задача проследить, как они возникли, как развивались и как функционируют сегодня. Этот специальный выпуск демонстрирует, что языки, рассматриваемые в нем, находились под значительным влиянием китайского языка через письменные тексты.

Ключевые слова: *китайский язык, кооптация, дискурсивный маркер, грамматикализация, языковой контакт, синосфера*

1. Introduction

The present special issue grew out of a workshop at the 18th International Pragmatics Conference that took place in Brussels from July 9 to 14, 2023. Organized by Seongha Rhee, Reijirou Shibasaki and Wenjiang Yang, the objective of the workshop was to study the development of words of the same origin, all going back to Chinese, into discourse markers (DMs) having diverse functions in the Asian languages Chinese, Korean, Japanese, Vietnamese, and Thai. The contributors of this issue were asked to explore the extent of commonalities and differences displayed by developmental scenarios of the DMs with shared characteristics in the language or languages analyzed by them.

In accordance with this objective, the present issue deals with a number of words of Chinese origin used in Japanese, Korean, Vietnamese, and Thai, as well as Chinese. It investigates theoretically significant issues related to language contact, discourse-pragmatic aspects of language change, and socio-cultural influence on language development, among others, as exemplified in the development of DMs from their earlier lexical expressions originating from Chinese etyma.

As has been established in a body of research (e.g., Rhee et al. 2021, Higashiizumi & Shibasaki in preparation), a number of words of Chinese origin are used as DMs in Asian languages. The five Asian languages mentioned all belong to different language families, i.e., Sino-Tibetan (Chinese), Japonic (Japanese), Koreanic (Korean), Austroasiatic (Vietnamese), and Kra-Dai (Thai), and they have widely varied typological profiles in syntax and morphology. Notwithstanding their genealogical and typological differences, they share an important commonality: They all belong to the Sinosphere Space a fact that is reflected in a number of features of individual languages, especially after "Sinosphere" in their lexicon.[1]

The Japanese lexicon, for instance, includes a large inventory of Sino-Japanese words; for example, they account for approximately 60 percent of the total word count in practical and popular science magazines (Shibatani 1990: 142–145). Similarly, Sino-Korean words account for 53 per cent of the headwords in an official dictionary (NIKL, n.d.) – according to Sohn (2001[1999]: 87), the number of Sino-Korean words amounts even to 60 per cent, thus being considerably larger than that of native Korean words, which account only for 35 per cent of the Korean lexicon.[2]

Vietnamese people have long had close interaction with Chinese and used Chinese characters until the current Latin-based script was adopted. Vietnamese has a massive inventory of 'literary' Sino-Vietnamese borrowings, but there are as well

---

[1] Thailand was strongly influenced by China early in its history, but at a later time it was more strongly influenced by India, leading to substantial borrowing of words from Sanskrit and Pali. Thus, even though Thailand is normally regarded as belonging to the Indosphere (Haarmann 2012[1986], Kulke & Rothermund 2004), the influence of Chinese on Thai is strongly felt in the modern Thai lexicon (see below).

[2] The last two figures are adopted from Shibasaki and Higashiizumi (in preparation).

'colloquial' Sino-Vietnamese borrowings (Alves 2007). Sino-Vietnamese words account for 60 to 70 per cent of Vietnamese lexicon (Nguyễn 1961, Alves 2001, 2009). Similar to Vietnamese, Thai people also have had interaction with Chinese for millennia, using Chinese characters before they began to use the current Thai script (Haarmann 2012[1986]). Their language is also known to have a large inventory of Sino-Thai words (SEALANG, n.d.).

The goal of this special issue is to demonstrate the substantial influence that Chinese had on the languages of the Sinosphere through written texts. The paper is structured as follows. Section 2 gives a concise characterization of the key notions of the present collection of contributions which are 'discourse marker', 'grammaticalization' and 'cooptation'. Section 3 then provides a general overview of the contributions making up this special issue. Section 4 presents a discussion of the mechanisms underlying the changes examined, followed by some general observations in Section 5. And finally, the conclusions reached are summarized in Section 6.

2. Discourse markers, grammaticalization and cooptation

For a better understanding of the analyses to be presented in the contributions to follow, three key notions employed there are briefly discussed in this section. The term 'discourse marker' (DM) is used for a wide range of phenomena and is referred to with a variety of different terms. More than forty terms have in fact been identified (e.g., Dér 2010, see also Dér & Markó 2010), and nearly a dozen of them are presently in common use. DMs serve to monitor the production of texts and to provide processing instructions on how to interpret texts. A classic definition of them is the one in (1), taken up in a similar format in more recent work, as in (2).

(1) [Discourse markers are] sequentially dependent expressions which bracket units of talk. (Schiffrin 1987: 31)
(2) By a DM I mean a metatextual marker that signals some kind of relationship between clauses/utterances. (Traugott 2018: 27)

As has been shown more recently, however, text organization is neither the only function associated with DMs, nor is it always really a feature of them, as when we say in English, *John is, **well**, a liar*, where the item *well*, commonly classified as a DM, serves a function other than that of text organization. Instead of (1) and (2), the more complex definition in (3) has been proposed to take care of salient functions of these markers.

(3) Discourse markers are (a) invariable expressions which are (b) semantically and syntactically independent from their environment, (c) set off prosodically from the rest of the utterance in some way, and (d) their function is metatextual, relating a text to the situation of discourse, that is, to the organization of texts, the attitudes of the speaker, and/or speaker-hearer interaction. (Heine et al. 2021: 6)

Of all the terms used instead of 'discourse marker' it is 'pragmatic marker' that has received the widest currency. The latter term can be found in a number of different usages, most of all the following three:[3] (a) as an equivalent of DM, (b) as referring to phenomena other than the ones covered by a DM, and (c) as a more general category that includes DMs. To the extent that the relevant authors do not make it quite clear which of these three usages is intended in their work, the term has at times given rise to confusion.

This does not apply to the study of Higashiizumi and Shibasaki (in preparation), who use 'pragmatic marker' in the sense of (b). For them, 'discourse markers' (in their sense) can connect both preceding and following information textually, whereas 'pragmatic markers' show the speaker's attitude to the preceding or the following utterance, not necessarily connected to the following discourse. On this view, corpus-based and text-based evidence from East Asian languages examined in Higashiizumi and Shibasaki (in preparation) suggests that diachronically, some 'discourse markers' have the potential to become 'pragmatic markers', thereby moving from the textual to the interpersonal level. According to Traugott (1982), a change in the opposite direction seems unlikely in language change in general. Of course, it is hard to determine at some point in history whether or not, or to what extent, a given marker shows or does not show the attitudes of a speaker (or writer). Such examples can be considered as cases of 'layering' (Hopper 1991) or 'bridging context' (Heine 2002).

In contrast, in the study of Traugott (2022: 5, Figure 1.2), reviewed in contribution (11) (see Section 3), 'pragmatic marker' is used in the sense of (c), in that it includes discourse structuring markers, which again include what she calls 'multifunctional DMs'. In accordance with the definition in (3), both pragmatic markers and discourse structuring markers are treated here within the general category of DMs.

While research on DMs is relatively young, that on grammaticalization has a distinctly longer history, going back to the 19th century. Its employment for describing the development of DMs, however, is of more recent origin (see especially Traugott 1982, 1995).

In more general terms, grammaticalization is based on the context-induced manipulation of linguistic expressions in discourse (cf. the context extension model of Heine 2002). A wealth of definitions of it has been proposed, and the one in (4) is in the spirit of most of these definitions. As stipulated in this definition, grammaticalization is a unidirectional process, which means that cases not conforming to the definition are not within the scope of grammaticalization theory, to be accounted for in terms of alternative mechanisms of linguistic change.

    (4) Grammaticalization is defined as the development from lexical to grammatical forms and from grammatical to even more grammatical

---

[3] See, for example, Heine et al. (2021: 11, fn. 17). See Brinton (2017: 2–10) for a comprehensive definition of pragmatic markers.

forms. Since the development of grammatical forms is not independent of the constructions to which they belong, the study of grammaticalization is also concerned with constructions and with even larger discourse segments. (Heine & Kuteva 2002: 2, see also Narrog & Heine 2021)

As for the criteria used to identify instances of grammaticalization, there are especially three proposals, namely those of Lehmann ([1982] 2015: 132), relying on six criteria or parameters, of Hopper (1991: 25–9), defining five principles of change, and of Heine and Kuteva (2007: 33–44), using four parameters for identifying instances of grammaticalization. Since the last proposal addresses most of the concerns of the other two approaches, it is adopted here. The parameters distinguished are listed in (5). With the exception of (5a), they have a focus on loss of linguistic substance in the process; their main purpose is to assist in the identification of instances of grammaticalization rather than defining it; for the latter, see (4). Note, however, that in the same way as there is loss in the process of grammaticalization, there are also gains (cf. (5a)). In fact, the present contributions bear witness to the observation made throughout the contributions and elsewhere that these gains can be and commonly are indeed fairly rich.

(5) The parameters of grammaticalization proposed by Heine and Kuteva (2007: 33–46)
   a Context extension: The rise of new meanings when linguistic expressions are extended to new contexts, leading to context-induced reinterpretation.
   b Desemanticization ('semantic bleaching'): Loss or generalization of meaning content or functions in such contexts.
   c Decategorialization: Loss of morphosyntactic properties characteristic of lexical or other less grammaticalized forms.
   d Erosion ('phonetic reduction'): Loss of phonetic substance, including prosodic features.

The third notion figuring in the contributions to follow is cooptation (e.g., Heine et al. 2013, 2017, 2021: 26–27, Furkó 2014). Introduced only recently, this notion addresses a feature of DMs that is not within the scope of grammaticalization theory, namely their metatextual function of discourse management, addressed in the definitions of (2) and (3). Grammaticalization and cooptation thus have starkly contrasting effects: Whereas the former typically entails increasing integration of linguistic items within the sentence or the word, the latter enables language users to extend the use of such items beyond the confines of a sentence. For example, English *beside(s)* experienced a process of grammaticalization from adverbial phrase to adverb, all happening within a sentence, but was later on coopted as a DM whereby it now serves discourse organization beyond the sentence (Traugott & Trousdale 2013: 109–12, Heine et al. 2021: 29).

Cooptation is described as a cognitive-communicative operation enabling speakers or writers to switch their perspective from the level of reasoning anchored

to the meaning of sentences to the metatextual level of reasoning immediately anchored to the situation of discourse (Heine et al. 2021: 67, cf. Rhee 2013). Its main effects on the text segments coopted are summarized in (6).

(6) Common effects of cooptation (Heine et al. 2021: 68)
   a   Meaning: From meaning as part of the sentence to meaning outside the sentence
   b   Function: From sentence-structuring function to metatextual function
   c   Syntax: From syntactic constituent of the sentence to syntactically unattached status
   d   Prosody: From prosodically integrated to unintegrated or less integrated status
   e   Semantic-pragmatic scope: From more restricted to wider scope
   f   Placement: From positionally constrained to less constrained placement

Beyond the three terms discussed in this section, there are other terms that are equally used commonly for describing the rise and development of DMs, especially terms like pragmaticalization and constructionalization. However, since such terms seem to be less relevant to the analyses proposed in the contributions to follow, we are not discussing them in this section.

Furthermore, it should be emphasized that, on the one hand, the development of DMs from lexical material presumably constitutes the predominant pathway characterizing the history of DMs. On the other hand, this constitutes by no means the only way. Other fairly common sources of DMs are provided, for instance, by straightforward borrowing, involving neither grammaticalization nor cooptation. Discussing a wide range of cases from languages across the world, Heine et al. (2021: 215) note:

> It would seem in fact that DMs are amongst the first grammatical items that speakers borrow or code-switch in situations of intense language contact, frequently but not always from the language of the more dominant or 'prestigious' group involved (Matras 1998, Grant 2012).

Furthermore, there are also two other pathways whereby DMs can arise, namely either via the grammaticalization of interactive categories such as vocatives, interjections, directives and attention signals (Heine 2023: 277–297), or via general extenders (Brinton 2017: 272–283, 2024, where 'pragmatic markers' are used instead of 'discourse markers'). For the former, English provides a few examples, like the interjection *oh*, as observed, for example, by Jucker (2002: 218): "To this extent, the Early Modern English *oh* is a proper discourse marker". Note, however, that in a process of grammaticalization, the earlier function tends to be retained side-by-side with the new grammaticalized function, and this is exactly what appears to have happened with *oh*: While it developed into a DM in certain contexts, in many other contexts its function as an interjection is still alive and well.

For the latter, Brinton (2017: 281–282, see also Brinton 2024, Brinton & Shibasaki forthcoming) identifies the following pathway: the parenthetical use of

or whatever becomes the general extender *or/and/like whatever*, finally giving rise to the stand-alone 'pragmatic marker' *whatever*.[4] However, cases of interactive categories and general extenders are not dealt with in the contributions to follow and are therefore not considered further in this paper.

3. The contributions to this special issue

The contributions to this special issue cover a wide range of topics, and they are now looked at in turn.

The first contribution, by **Wenjiang Yang**, provides a contrastive diachronic analysis of the Chinese form *yuánlái* 'originally, previously' and Japanese *ganrai* 'originally, inherently', tracing their development from temporal adverbials to DMs. The Japanese form was borrowed from the Chinese one through written texts, both having been temporal adverbials meaning 'originally, from the beginning', and this meaning has persisted till today. The development of the two into DMs followed contrasting pathways: Chinese *yuánlái* turned into a DM encompassing mirative, background and justificational functions, whereas Japanese *ganrai* eventually evolved into an elaborative marker. This case study relates to several topics, to be discussed in Section 4. First, borrowing was restricted to lexical material and it involved the written rather than a spoken mode. Second, the shift from sentence grammar into the domain of discourse organization appears to have happened independently in the two languages. And third, the change can be viewed as one of parallel grammaticalization rather than replica grammaticalization, in that it took quite different directions and lines of semantic-pragmatic change in the two languages.

The second contribution deals with an issue of a different kind. In their diachronic analysis of Chinese *shènzhì* 'even' and *bùguò* 'however, but' in their various usages, **Haiping Long and Lei Wang** classify these two forms as discourse structuring markers, that is, as a weakly grammaticalized type of DMs. The authors show that the two forms experienced different ways of development. Whereas *shènzhì* was originally used in clause-initial position, expressing speaker-oriented meanings, it acquired clause-medial usages in later stages of its development. These features are not shared by *bùguò*: While also emerging in clause-initial position, it neither conveyed speaker-oriented meanings nor did it acquire clause-medial usages. Based on this observation, Long and Wang identify an interesting correlation between meaning and syntactic change in grammaticalization, in that change from initial to medial position can be accounted for with reference to the presence of early speaker-oriented meaning. The authors find further support for this hypothesis in developments to be observed in English, also confirmed by similar developments in Korean and Hungarian. This case study differs from the

---

[4] In the same manner, Brinton and Shibasaki (forthcoming) explore another complex case, *no matter (what)*, in the history of English. See Overstreet and Yule (2021) for detailed discussions of general extenders.

other contributions mainly in two respects. First, it does not involve language contact and, second, it focuses on DM-internal developments rather than on the question of how discourse structuring markers or, more generally, DMs arise. But its conclusions are noteworthy, demonstrating that the source meaning of a DM can determine the course of its subsequent grammaticalization and usage.

The topic of the third contribution, by **Seongha Rhee and Lin Zhang**, is devoted to a systematic comparison between a Korean DM and its etymological source in Chinese. The Korean DM *cincca*, literally meaning 'a true thing', is historically a noun phrase which was grammaticalized into an item carrying adjectival and adverbial functions of adding genuineness or excellence in quality to a modified noun, or an adverbial function of adding emphasis to an adjective or a predicate. Turning into a DM, it retained its intensifying function but developed various new functions. Its corresponding cognates in Chinese are provided by the DMs *zhende* and *zhenshi*, having similar functions, but *zhenshi* differs in one important respect, marking the speaker's negative evaluation of the person or event it refers to. The authors demonstrate that the Korean and the Chinese DMs underwent massive grammaticalization processes, resulting in both similarities and differences with regard to their functional distribution, prosody, and relative degree of desemanticization, that is, their loss of earlier meaning features. While both the Chinese and the Korean items are DMs, the evidence available suggests that the rise of Korean *cincca* as a DM was a language-internal process taking place some time after the turn of the 20th century.

In the fourth contribution, **Yuko Higashiizumi, Reijirou Shibasaki and Keiko Takahashi** deal with a pathway of grammaticalization that appears to also have occurred in some other languages, as shown in the *World Lexicon of Grammaticalization* (Kuteva et al. 2019: 443). The authors trace the development of the Japanese adverb *shinni* 'truly', containing the Sino-Japanese noun *shin* 'truth'. The adverb serves, for example, to intensify the illocutionary force of an apology or a regret, akin to English *very*. While the development of Japanese *shinni* was restricted to grammaticalization, the authors note that similar processes can also be observed in Chinese, Korean and Thai but leading one step further, giving rise to DMs, now having interactional functions of discourse processing.

The fifth contribution by **Jiyeon Park** is about semantic-pragmatic manipulation of the Japanese adverb *zenzen* 'completely, entirely, not at all', borrowed from Chinese. The author argues that *zenzen* evolved as a DM ('pragmatic marker') indicating the speaker's epistemic stance and viewpoint. Starting as a lexical item with objective meanings, *zenzen* underwent cooptation and turned into a DM having subjective and intersubjective functions. In spoken discourse, it is also found in a specific construction where it combines with a copula. Functioning as an adjectival noun, its syntactical independence from the sentence gives it the appearance of a complex DM or, more generally, of a thetical (Kaltenböck et al. 2011). Park's main concern is with semantic change in grammaticalization leading to the expression of subjectivity and intersubjectivity,

but she also addresses features in development of *zenzen* that are suggestive of cooptation as it is characteristic of the rise of DMs (see Heine et al. 2017, Heine et al. 2021).

The growth of DMs from lexical sources is also the subject of **Hyun Sook Lee** in the sixth contribution. Originally appearing in the 15th century as a noun for 'the world' or 'the world or society people live in', the item was extended semantically, receiving a range of additional meanings, and it grammaticalized into an adverbial degree modifier ('very') and negative polarity item ('at all'). When combining with the locative particle -*ey* (*seysangey*), it became a DM in Modern Korean, being syntactically independent and enjoying positional freedom.

This study of *seysang* and *seysangey* discusses a number of issues, illustrating in particular the growth of subjective and intersubjective meanings and other innovations characteristic of the shift from material of sentence grammar to status as a DM.

Another process leading from lexical item to DM is discussed by **Sunhee Yae** in the seventh contribution. Ultimately a noun in Late Middle Korean, the Sino-Korean item *iltan* 'one morning' was grammaticalized into an adverb and a connective, acquiring new meanings associated with priority, short duration and conditionality. The author goes on to also analyze the Chinese equivalent *yīdàn* of Korean *iltan.* Differences between the two relate most of all to the following: Chinese *yīdàn* seems to be less strongly grammaticalized. While it acquired uses to express brief duration and conditionality, it does not express priority.

The development of *iltan* displays the whole gamut of processes to be expected in the evolution of DMs, that is, grammaticalization and cooptation as well as semantic-pragmatic developments into the expression of subjectification and intersubjectification. At the same time, it does not confirm the hypothesis of a functionally motivated asymmetry between the left periphery and the right periphery of DM placement – an observation also made by Hyun Sook Lee in contribution 6. Korean *iltan* and Chinese *yīdàn* seem to provide an instance of parallel grammaticalization, where two items having the same lexical source underwent a development in the same direction, even though this development was more pronounced in Korean.

Moving from Korea to Thailand, the eighth contribution by **Kultida Khammee** deals with the Thai lexeme *ciŋ* 'true', borrowed from Middle Chinese *cin* 'true, real'. Grammaticalization had the effect that a lexical item denoting 'true, truth, real, genuine' assumed adverbial functions as an intensifier denoting 'surely, certainly, definitely'. But it also appears to have acquired a number of DM functions expressing intersubjective meaning. The development of *ciŋ* was apparently not restricted to grammaticalization; *ciŋ* also must have undergone cooptation, giving rise not only to a DM but also to response signals (Heine 2023), in that it turned into the interactive forms *ciŋ* 'yes' and its negative counterpart *mây ciŋ* 'no', expressing agreement and disagreement, respectively. All evidence that there is

suggests that these changes of Thai *ciŋ* were essentially language-internal, not shared by its Chinese etymological source item *cin.*

The ninth contribution by **Mayumi Adachi** deals with a topic related to several other contributions, but in another language of the Sinosphere, namely, Vietnamese. Vietnamese *thật* has its origin in a Sino-Vietnamese lexeme meaning 'full, rich, fruit', but this meaning was extended to also express 'truth, reality, and fact'. Occurring as an adjective and adverb in the 17th century, it is now found in various spoken and written contexts, including uses as an adjective 'real' and an adverb 'really'. But it is found also in DMs in utterance-initial position in the forms *thật ra* (*thật* + 'out'), *quả thật* ('fruit'+ *thật*), and *kỳ thật* ('its'+*thật*), all of which are glossed 'in fact'. These combinations seem to be vital to attracting the attention of potential interlocutors, thus constituting a case of intersubjectification.The author argues that *thật* experienced a process of grammaticalization leading from denoting obvious facts to expressing the speaker's affective or evaluative stance. Once again, we seem to be dealing with a language-internal change involving grammaticalization but presumably also cooptation as a DM.

The remaining contributions 10 and 11 are of a different kind: They present reviews of books that have appeared more recently and provide new theoretical perspectives on the analysis of DMs and their development.

The first book, which is the subject of the tenth contribution, is *The Rise of Discourse Markers* (Heine et al. 2021), reviewed by **Fangqiong Zhan**. The authors of the book note that some previous works on DMs have suffered from not accounting for salient properties of DMs, such as their relatively independent semantic, syntactic and prosodic status. Studying DMs in languages spoken in various parts of the world, but especially in English, French, Japanese and Korean, they argue that the presence of such properties can be explained in terms of cooptation, a cognitive–communicative operation whereby text segments such as clauses or parts of them are transferred from the level of sentence organization to the domain of discourse organization. The development of DMs thus is explained by a combination of two distinct mechanisms, namely grammaticalization and cooptation. In her careful analysis of the book, Zhan draws attention in particular to the fact that the book could have benefitted from taking into account findings made on constructional change, especially as it is described in the framework of Diachronic Construction Grammar (e.g., Traugott & Trousdale 2013, see also Hilpert 2013 and Barðdal et al. 2015).

The final paper, contribution 11, is a review by **Foong Ha Yap** and **Mikyung Ahn** of the book *Discourse Structuring Markers in English* (Traugott 2022). In the book, the author turns to construction grammar, offering a diachronic perspective on how new functional categories are constructionalized, and suggests how a network model can account for the interlocking patterns in language change. The main theoretical question asked is how pragmatic/contextual factors can best be incorporated in Discourse Construction Grammar, the framework proposed by the author. A constructionalist perspective is provided based on reconstructions of a

number of different types of English DMs, more specifically of discourse structuring markers (see also the review of Heine & Long 2022). Whereas Heine et al. (2021) focus on two specific phenomena in the development of DMs, namely directionality and cooptation, the author of this book provides a wider framework for the reconstruction of DMs. Given the relevance of the framework expounded in Traugott (2022), the reviewers suggest that the book calls for new studies investigating discourse structuring markers also in other languages. Unlike the book reviewed in the preceding contribution (Heine et al. 2021), this book is restricted to looking at one particular language, that is, English.

4. Mechanisms of change

As the preceding summary of the various contributions suggests, the development of DMs that have their origin in Chinese etyma, as observed in Japanese, Korean, Thai and Vietnamese, provides uniquely valuable insights into the role played by conceptual and cognitive operations as well as socio-cultural and discourse-pragmatic strategies in language change. With regard to this development, a couple of more general distinctions need to be made.

First, the development of a number of the DMs examined involve two main stages. At stage 1, there was contact-induced transfer, more precisely borrowing of Chinese lexical material into the other four languages examined. And at stage 2, there was internal development within each of the languages. Obviously, developments within Chinese only involved the second stage (e.g., Long & Wang 2024). The meanings of the borrowed items at stage 1 tended to be close to those of the source language Chinese. Subsequently, at stage 2 then, the borrowed items underwent diverse changes in their form and meaning within each of the languages concerned, a process known as 'internally-motivated' change (Heine & Kuteva 2005: xii). And second, with regard to their status as DMs, two different kinds of development need to be distinguished. On the first development, the relevant item was already a DM in the donor language Chinese and borrowing simply involved transfer of the DM from the donor language to the recipient language. Various lines of research have in fact established that borrowing of DMs is very common in the languages of the world (see Matras & Sakel 2007, Heine et al. 2021: Chapter 7). Developments in the recipient languages Korean, Japanese, Vietnamese, and Thai then would have concerned mainly grammaticalization, whereby a borrowed DM was extended to new contexts and acquired novel grammatical functions.

Some particular meanings appear to have their own potential of triggering grammaticalization independently from language contact and regional areas (e.g., 'fact,' 'result,' 'issue,' to name but a few; see Higashiizumi & Shibasaki in preparation for details). As was shown in various contributions to the present issue, the developments of the DMs of Chinese origin exhibit such scenarios at varying degrees.

On the second development, borrowing entailed the transfer of lexical material from Chinese, and the rise and development of the DM took place subsequently

within the donor language. And here again, two different pathways need to be distinguished: Either the development from lexical item to DM took place simultaneously with its borrowing, that is, the development into a DM coincided with the process of borrowing. Alternatively, a lexical item of Chinese was borrowed as lexical material in the recipient language and the development from lexical item to DM took place subsequently in the recipient language. In the latter case, the development can follow two contrasting pathways: It either involves 'replica grammaticalization', namely a process whereby a grammaticalization process in the recipient language is realized based on a model process in the donor language. Alternatively, grammaticalization takes place independently of one another in the two languages, that is, there is 'parallel grammaticalization' (Heine & Kuteva 2005: 92–100). More recent research suggests that replica grammaticalization is crosslinguistically not highly common.

Another important issue relates to the nature of contact between languages. Typically, DMs are used in spoken discourse (Jucker 2002). Therefore, if a DM is borrowed from one language to another, it would imply borrowing through spoken contact (e.g., *you know* in Clyne 2003). However, spoken language is for the most part not accessible to historical analysis via written documents. Situated in a geographically connected land mass, China, Vietnam and Thailand have long maintained contact. Direct contact between Chinese and Vietnamese and between Chinese and Thai has continued from historical times, involving migration and intermarriage, though the documentation of early contact is scanty.

On the other hand, the heavy lexical borrowing of Chinese words in Japanese and Korean was channeled extensively through written texts, as these peninsular and insular countries are more distantly located for contact as compared to Vietnam and Thailand. The Japanese and Korean cases, thus, present a somewhat special contact situation not widely found in the languages of the world (see, e.g., Norman 1988, Sohn 1999, Irwin & Zisk 2019, Narrog et al. 2018). It is widely assumed that the rise and development of DMs is a matter of the spoken mode of speech. While this is presumably true most of the time across the world, such processes can be channeled as well through the written media, as can be seen in the contributions to follow (see also Shibasaki and Higashiizumi, in preparation).

Another significant theme addressed in this special issue relates to (inter)subjective stance-marking and stance-taking, often heavily influenced by the culture of the speech community (Englebretson 2007). Situated in interactive discourse, DMs signal both a subjective stance, one that is based on the evaluation of a subject, and an intersubjective stance, based on some form of alignment with the interlocutor (Traugott 1982, Du Bois 2007, Kaltenböck et al. 2011). The grammaticalization process of DMs may be widely variable across individual forms and languages, but it has the following shared feature in common: The DMs, by virtue of being markers functioning at the level of metadiscourse, or macrostructure, are formed by discourse-pragmatic strategies to fulfil the speaker's needs to be efficient, persuasive, and expressive (cf. Hyland 1998, Heine et al. 2021).

5. Discussion

As the observations made in Section 3 suggest, there are two main themes in this special issue. One concerns diachronic reconstruction. Based on the analysis of historical text data and linguistic reconstruction work, the contributions to follow all pursue much the same goal, namely contributing to a deeper understanding of grammatical development in the languages of East and Southeast Asia. The primary incentive underlying most of the contributions is to present a thorough description of language-internal processes as they can be accounted for especially with reference to principles underlying grammaticalization.

A topic to be addressed in more detail in future research concerns the stage of transition from lexical coding anchored to the structure of sentences to coding immediately anchored to the situation of discourse. This transition has been described in terms of cooptation (see Section 2) but more historical data are needed to understand the exact nature of the process. For example, what induced speakers or writers of Chinese, Japanese, Korean, Thai and Vietnamese to introduce new ways of structuring their texts, as well as of expressing their attitudes and their relations with their hearers or readers, respectively? And what exactly were the semantic, syntactic, and prosodic effects of this process?

The second theme in the collection of contributions is language contact. All the linguistic items examined can ultimately be traced back in some form or other to Chinese. Their extension from the Chinese donor language to the four recipient languages, i.e. Japanese (Yang 2024, Higashiizumi et al. 2024, Park 2024), Korean (Rhee & Zhang 2024, Lee 2024, Yae 2024), Vietnamese (Adachi 2024) and Thai (Khammee 2024) involved borrowing, that is, transfer of form-meaning units from one language to another; replication, that is, the transfer of meaning and/or structure but not form, seems to have played only a minor role, if at all. As we saw in Section 3, the process following borrowing was apparently for the most part one of internal grammaticalization, no clear case of replica grammaticalization has been reported. That the overall development is one of borrowing of lexical material from Chinese, which subsequently undergoes grammaticalization and cooptation in the recipient language, not involving replica grammaticalization, has also been observed in studies beyond the present contributions (Higashiizumi & Shibasaki in preparation). These studies also confirm another observation made here, namely that this process seems to always have taken place independently in each of the recipient languages.

As Section 4 showed, the nature of borrowed items can be of two kinds: Either that item was already a DM in the donor language and as such underwent grammaticalization and cooptation subsequently in the recipient language (Japanese, Korean, Vietnamese and Thai), or borrowing consisted of lexical material from the donor language (Chinese). While borrowing of DMs is a worldwide common process, there seems to be no clear case of it in the studies examined here; rather, the process seems to have been generally one where Chinese provided lexical material to be changed subsequently in the recipient languages.

In addition to their focus on detailed description and reconstruction work, the contributions also deal with general theoretical issues relating to the growth of discourse structuring material. For example, some authors (Rhee & Zhang 2024, Lee 2024, Yae 2024, Khammee 2024) take issue with the claim made in various earlier studies according to which the development of DMs entails structural asymmetry between two kinds of usages -- a hypothesis that has received quite some attention in work on discourse organization (Beeching et al. 2009, Detges & Waltereit 2014, Beeching & Detges 2014, Beeching et al. 2018, Pons Bordería 2018). On this hypothesis, the left periphery of an utterance tends to be associated with subjective functions while the right periphery is dedicated to intersubjective functions. While being appealing in some way, this hypothesis must be taken with care, as the studies by Hyun Sook Lee on Korean *seysang* and *seysangey* (Lee 2024) and Sunhee Yae on Korean *iltan* (Yae 2024) suggest. That the hypothesis is problematic has also been observed by other researchers (e.g., Pons Bordería 2018).

6. Conclusions

What makes this collection of contributions an invaluable document is especially its richness in the documentation and analysis of processes characterizing the history of the languages of East and Southeast Asia.

The combination of languages featuring in this special volume provides an optimal basis for comparative linguistic work. On the one hand, the languages have a number of features in common. They share a long history of language contact resulting in massive borrowing. This history accounts for the presence of a substantial volume of lexical items being similar in form and meaning to be found in the five languages. What makes this also a noteworthy case is the fact that contact-induced change and its effects on the present-day spoken language strongly involved the written mode – thereby contrasting with situations in most other parts of the world where language contact was restricted to the spoken medium (see also Higashiizumi and Shibasaki in preparation).

On the other hand, the five languages analyzed also differ remarkably from one another. For one thing, they are – as far as we know – all genetically unrelated. And for another, they exhibit highly contrasting typological profiles relating to both their morpho-syntactic organization and their phonology. In particular, Chinese, Thai and Vietnamese have predominantly isolating-analytic structures, and all three are tonal languages. Japanese and Korean, by contrast, are characterized by the presence of agglutinating and inflectional word structures, and neither disposes of distinctions in lexical or grammatical tone. Thus, these Asian languages present a close to ideal testing ground for theories of diachronic linguistics.

The testing achieved in the contributions to follow is most of all of the following kind. The findings presented suggest that there is a general line of linguistic development leading from lexical material, such as nominal and adverbial text segments, to the emergence of grammatical units serving new ways of expressing the attitudes of speakers or writers and functions of interaction between

speakers and hearers, or writers and readers. But the primary function of the units is to organize the overall structure of discourse, that is, the way texts are constructed and anchored to the situation in which discourse takes place. These units, referred to here as discourse markers (DMs), have been described in previous work also in terms of a variety of alternative theoretical notions, such as pragmatic markers or discourse structuring markers.

The present special issue investigates the extent of commonalities and differences displayed by developmental scenarios of DMs with shared characteristics in multiple languages. It offers a good opportunity to broaden our knowledge on various issues on DMs, and it deserves the attention of researchers working on contact linguistics, interactional linguistics, cultural linguistics, and language change at large. It is hoped that the approach adopted will encourage further comparative research in other regions of the world to contribute to a more comprehensive understanding of how our knowledge of discourse processing in the languages as they are spoken and written today can be explained with reference to the way in which this knowledge evolved over time.

References

Adachi, Mayumi. 2024. From truth to discourse marker: The case of *thật* in Vietnamese. *Russian Journal of Linguistics* 28 (4). 966–990. https://doi.org/10.22363/2687-0088-40501

Alves, Mark J. 2001. What's so Chinese about Vietnamese? In Graham W. Thurgood (ed.), *Papers from the ninth annual meeting of the Southeast Asian linguistics society,* 221–242. Tempe, AZ: Arizona State University, Program for Southeast Asian Studies.

Alves, Mark J. 2007. Categories of grammatical Sino-Vietnamese vocabulary. *Mon-Khmer Studies* 37. 217–229.

Alves, Mark J. 2009. Loanwords in Vietnamese. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook,* 617–637. Berlin: Mouton de Gruyter.

Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.). 2015. *Diachronic Construction Grammar*. Amsterdam, Philadelphia: John Benjamins.

Beeching, Kate & Ulrich Detges (eds.). 2014. *Discourse Functions at the Left and the Right Periphery: Crosslinguistic Investigations of Language Use and Language Change.* Leiden, Boston: Brill.

Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers: Pathways of Change.* Cambridge: Cambridge University Press.

Brinton, Laurel J. 2024. The rise of *what*-general extenders in English. *Journal of Historical Pragmatics* 25 (1). 104–136.

Brinton, Laurel J. & Reijirou Shibasaki. (forhcoming). "No matter what, you must think positive": The development of *no matter (what)* as a pragmatic marker. Manuscript, Meiji University, Tokyo.

Beeching, Kate, Chiara Ghezzi & Piera Molinelli. 2018. *Positioning the Self and Others: Linguistic Traces.* Amsterdam, Philadelphia: Benjamins.

Clyne, Michael. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press.

Dér, Csilla Ilona. 2010. On the status of discourse markers. *Acta Linguistica Hungarica* 57 (1). 3–28.

Dér, Csilla Ilona & Alexandra Markó. 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53 (2). 135–180.

Detges, Ulrich & Richard Waltereit. 2014. *Moi je ne sais pas* vs. *Je ne sais pas moi*: French disjoint pronouns in the left vs. right periphery. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the Left and the Right periphery: Crosslinguistic investigations of language use and language changem,* 24–46. Leiden: Brill.

Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction,* 139–182. Amsterdam: John Benjamins.

Englebretson, Robert. 2007. Stancetaking in discourse: An introduction. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction,* 1–25. Amsterdam: John Benjamins.

Furkó, Bálint Péter. 2014. Cooptation over grammaticalization. *Argumentum* 10. 289–300.

Grant, Anthony P. 2012. Contact, convergence, and conjunctions: A cross-linguistic study of borrowing correlations among certain kinds of discourse, phasal adverbial, and dependent clause markers. In Claudine Chamoreau & Isabelle Léglise (eds.), *Cross-linguistic tendencies in contact-induced change: A typological approach based on morphosyntactic studies,* 311–58. Berlin: Mouton de Gruyter.

Haarmann, Harald. 2012[1986]. *Language in Ethnicity: A View of Basic Ecological Relations*. Berlin: De Gruyter Mouton.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 83–101. Amsterdam, Philadelphia: Benjamins.

Heine, Bernd. 2023. *The Grammar of Interactives*. Oxford: Oxford University Press.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2013. An outline of discourse grammar. In Shannon Bischoff & Carmen Jany (eds.), *Functional approaches to language,* 175–233. Berlin : Mouton de Gruyter.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2017. Cooptation as a discourse strategy. *Linguistics* 55 (4). 813–55.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2007. *The Genesis of Grammar: A Reconstruction.* Oxford: Oxford University Press.

Heine, Bernd & Haiping Long. 2022. Review of Elizabeth Close Traugott 2022, Discourse structuring markers in English: A historical constructionalist perspective on pragmatics. *Functions of Language* 29 (3). 338–343.

Higashiizumi, Yuko, Reijirou Shibasaki & Keiko Takahashi. 2024. From truth to truly: The case of *shinni* 'truly' in Japanese compared to Chinese, Korean, and Thai counterparts. *Russian Journal of Linguistics* 28 (4). 843–864. https://doi.org/10.22363/2687-0088-40518

Higashiizumi, Yuko & Reijirou Shibasaki (eds.). (In preparation). *The Emergence of Pragmatic Markers from Chinese Compounds in Chinese, Japanese, and Korean: Perspectives from East Asian Languages and Beyond.* Leiden: Brill.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word-Formation, and Syntax.* Cambridge: Cambridge University Press.

Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth Close Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, Volume 1, 17–35. Amsterdam, Philadelphia: Benjamins.

Hyland, Ken. 1998. Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics* 30. 437–455.

Irwin, Mark & Matthew Zisk. 2019. *Japanese Linguistics*. (The Japanese Language I)*.* Tokyo: Asakura Publishing.

Jucker, Andreas. 2002. Discourse markers in Early Modern English. In Richard Watts & Peter Trudgill (eds.), *Alternative histories of English,* 210–230. London: Routledge.

Kaltenböck, Gunther, Bernd Heine &Tania Kuteva. 2011. On thetical grammar. *Studies in Language* 35 (4). 848–893.

Khammee, Kultida. 2024. From objective to subjective and to intersubjective functions: The case of the Thai 'truth'-lexeme. *Russian Journal of Linguistics* 28 (4). 942–965. https://doi.org/10.22363/2687-0088-40496

Kulke, Hermann & Dietmar Rothermund. 2004. *A History of India*. Fourth edition. New York: Routledge.

Lee, Hyun Sook. 2024. From a noun to a discourse marker: The case of *seysang* 'world' in Korean. *Russian Journal of Linguistics* 28 (4). 891–915. https://doi.org/10.22363/2687-0088-40499

Lehmann, Christian. 2015[1982]. *Thoughts on Grammaticalization.* Third edition. Berlin: Language Science Press.

Long, Haiping & Lei Wang. 2024. Speaker-orientation meaning and positional shift of discourse structuring markers. *Russian Journal of Linguistics* 28 (4). 794–817. https://doi.org/10.22363/2687-0088-40605

Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36. 281–331.

Matras, Yaron & Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31 (4). 829–865.

Narrog, Heiko & Bernd Heine. 2021. *Grammaticalization*. Oxford: Oxford University Press.

Narrog, Heiko, Seongha Rhee & John Whitman. 2018. Grammaticalization in Japanese and Korean. In Heiko Narrog & Bernd Heine (eds.), *Grammaticalization from a typological perspective,* 166–188. Oxford: Oxford University Press.

Nguyễn, Đình Hòa. 1961. *The Vietnamese Language.* Vietnam Culture Series, no. 2. Saigon: Department of National Education.

Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.

Overstreet, Maryann & George Yule. 2021. *General Extenders: The Forms and Functions of a New Linguistic Category*. Cambridge: Cambridge University Press.

Park, Jiyeon. 2024. The evolution of pragmatic marker *zenzen* in Japanese: From objectivity to intersubjectivity. *Russian Journal of Linguistics* 28 (4). 865–890. https://doi.org/10.22363/2687-0088-40516

Pons Bordería, Salvador. 2018. Paths of grammaticalization: Beyond the LP/RP debate. In Salvador Pons Bordería & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers,* 334–383. Leiden, Boston: Brill.

Rhee, Seongha. 2013. "I know I'm shameless to say this": Grammaticalization of the mitigating discourse marker *makilay* in Korean. *Procedia: Social and Behavioral Sciences* 97. 480–486.

Rhee, Seongha, Reijirou Shibasaki & Xinren Chen (eds.). 2021. Grammaticalization of discourse markers in East Asian Languages. *East Asian Pragmatics* 6 (3). [Special issue].

Rhee, Seongha & Lin Zhang. 2024. The way of truth: The case of the Korean discourse marker *cincca* in comparison with Chinese *zhenshi* and *zhende*. *Russian Journal of Linguistics* 28 (4). 818–842. https://doi.org/10.22363/2687-0088-40500

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Shibasaki, Reijirou & Yuko Higashiizumi. (In preparation). An introduction. In Yuko Higashiizumi & Reijirou Shibasaki (eds.), *The emergence of pragmatic markers from Chinese compounds in Chinese, Japanese, and Korean: Perspectives from East Asian languages and beyond.* Leiden: Brill.

Shibatani, Masayoshi. 1990. *The Languages of Japan.* Cambridge: Cambridge University Press.

Sohn, Ho-min. 2001[1999]. *The Korean Language.* Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for Historical linguistics: A symposium,* 245–271. Austin: University of Texas Press.

Traugott, Elizabeth Closs. 1995. The role of the development of discourse markers in a theory of grammaticalization. *Paper presented at the International Conference of Historical Linguistics XII, Manchester 1995.*

Traugott, Elizabeth C. 2018. Modeling language change with constructional networks. In Salvador Pons Bordería & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers: New issues in the study of language change*, 17–50. Leiden: Brill.

Traugott, Elizabeth Closs. 2022. *Discourse Structuring Markers in English: A Historical Constructionalist Perspective on Pragmatics.* Amsterdam, Philadelphia: Benjamins.

Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Yae, Sunhee. 2024. From 'one morning' to a discourse marker: The case of *iltan* in Korean. *Russian Journal of Linguistics* 28 (4). 916–941. https://doi.org/10.22363/2687-0088-40602

Yang, Wenjiang. 2024. From temporal adverbials to discourse markers: The development of Chinese *yuánlái* and its Japanese cognate *ganrai*. *Russian Journal of Linguistics* 28 (4). 771–793. https://doi.org/10.22363/2687-0088-40514

Dictionaries and Internet resources

NIKL (The National Institute of Korean Language). n.d. Dictionary statistics of *Phyocwun Kwuke Taysacen* [A Comprehensive Standard Korean Dictionary]. https://stdict.korean.go.kr, accessed in January 2024.

SEALANG (Southeast Asian Languages Project Library). n.d. sealang.net/thai/chinese/middle.htm, accessed in January 2024.

Bionotes:

Bernd HEINE is Emeritus Professor at the Institute of African Studies (Institut für Afrikanistik und Ägyptologie), University of Cologne, Germany. He has held visiting professorships in Europe, Eastern Asia, Australia, Africa, North America, and South America. His research interests embrace African languages, grammaticalization theory, linguistic typology, language contact, and interactive grammar. He has authored and co-authored numerous books, the most recent of which are *The Rise of Discourse Markers* (CUP, 2021), and *The Grammar of Interactives* (OUP, 2023).
e-mail: bernd.heine@uni-koeln.de
https://orcid.org/0000-0002-9304-7814

Wenjiang YANG is Associate Professor of Japanese linguistics at University of Macau, Macao SAR, China. He received his PhD in Japanese linguistics at Peking University in 2014. His research interests include morphology, syntax, semantics, historical linguistics,

and contrastive linguistics, with a special focus on tense, aspect, evidentiality and discourse markers in Present-Day Japanese as well as their diachronic evolution.
e-mail: wenjiangyang@um.edu.mo
https://orcid.org/0009-0000-1802-0630

Seongha RHEE is Professor of Linguistics at Mahidol University, Thailand and Professor Emeritus at Hankuk University of Foreign Studies, Korea. He received his PhD in linguistics from the University of Texas, Austin in 1996. His primary research interest is to identify cognitive and discursive mechanisms that enable language change from the crosslinguistic and typological perspectives. He published *World Lexicon of Grammaticalization* (co-author, 2019, CUP); book chapters in *The Cambridge Handbook of Korean Linguistics* (2022, CUP); and research articles in *Nature*, *Journal of Pragmatics, Language Sciences, Russian Journal of Linguistics* and *Lingua,* among others.
e-mail: srhee@hufs.ac.kr
https://orcid.org/0000-0003-0312-0975

Сведения об авторах:

Бернд ХАЙНЕ – почетный профессор Института африканских исследований Кельнского университета, Германия. Работал в качестве приглашенного профессора в Европе, Восточной Азии, Австралии, Африке, Северной Америке и Южной Америке. Его исследовательские интересы включают африканские языки, теорию грамматикализации, лингвистическую типологию, языковые контакты и интерактивную грамматику. Является автором и соавтором многих книг, последние из которых *The Rise of Discourse Markers* (Cambridge University Press, 2021) и *The Grammar of Interactives* (Oxford University Press, 2023).
e-mail: bernd.heine@uni-koeln.de
https://orcid.org/0000-0002-9304-7814

Вэньцзян ЯН – доцент кафедры японской лингвистики в Университете Макао, Макао, Китай. Получил степень PhD по японской лингвистике в Пекинском университете, Китай, в 2014 году. Его исследовательские интересы включают морфологию, синтаксис, семантику, историческую лингвистику и контрастивную лингвистику с особым акцентом на времени, аспекте, эвиденциальности и дискурсивных маркерах в современном японском языке, а также в их диахронической эволюции.
e-mail: wenjiangyang@um.edu.mo
https://orcid.org/0009-0000-1802-0630

Сеонгха РИ – профессор лингвистики в Университете Махидол, Тайланд, и почетный профессор в Ханкукском университете иностранных исследований, Корея. Получил степень PhD по лингвистике в Техасском университете в Остине в 1996 году. Сфера его научных интересов – выявление когнитивных и дискурсивных механизмов изменения языка в межъязыковой и типологической перспективе. Автор книги *World Lexicon of Grammaticalization* (Cambridge University Press, 2019), глав в книгах *The Cambridge Handbook of Korean Linguistics* (2022) и научных статей в журналах *Nature*, *Journal of Pragmatics*, *Language Sciences*, *Russian Journal of Linguistics*, *Lingua* и др.
e-mail: srhee@hufs.ac.kr
https://orcid.org/0000-0003-0312-0975

Research article / Научная статья

From temporal adverbials to discourse markers:
The development of Chinese *yuánlái* and its Japanese cognate *ganrai*

Wenjiang YANG ⓘ ✉

University of Macau, Macao SAR, China
✉wenjiangyang@um.edu.mo

Abstract

This study aims to explore and compare the developmental pathways of Chinese *yuánlái* 'originally, previously' and its Japanese cognate *ganrai* 'originally, inherently' from temporal adverbials to discourse markers (DMs). It seeks to reveal how words originating from the same Chinese source evolve into discourse markers with different functions. The data utilized are drawn from 4 electronic corpora and 1 database, which encompass the complete history of Chinese and Japanese. Chinese *yuánlái* developed from the temporal noun *yuán* 'origin,' and the suffix *lái* 'to come' was added in the 8th century CE, driven by the disyllabification trend in traditional Chinese. It entered Japanese through written texts in about the same period. Both *yuánlái* and *ganrai* embarked with the meaning of 'originally, from the beginning.' This study demonstrates that their semantic changes were motivated by the differentiation of pragmatic implicatures, specifically the Q-Principle and the R-Principle proposed by Laurence R. Horn (1984, 2012a, 2012b). *Yuánlái* followed a trajectory starting from the lexical meaning 'previously' with the feature of [+contrast], evolving into mirative, background and justificational markers. On the other hand, DM *ganrai* originated from the lexical meaning of 'from the beginning till now, always' with the feature of [-contrast], later interpreted as "by nature, inherently," and eventually transformed into an elaborative marker. The findings of this study offer fresh insights into the emergence of discourse markers from shared Chinese etyma through language contact within the Sinosphere. Moreover, it is revealed that constraints of lexical meanings could influence the emergence of potential DM functions.

Keywords: *discourse marker*, *temporal adverbial*, *historical development*, *pragmatic implicature*, *Chinese-Japanese language contact*

От обстоятельств времени к дискурсивным маркерам: развитие китайского *yuánlái* и его японского когната *ganrai*

Вэньцзян ЯН ⬭✉

Университет Макао, САР Макао, Китай
✉wenjiangyang@um.edu.mo

Аннотация

Целью данной статьи является сравнение путей развития китайского *yuánlái* (原来) «изначально, ранее» и его японского когната *ganrai* «изначально, по своей сути» от обстоятельств времени до дискурсивных маркеров (ДМ). Задача исследования – выявить, как слова, происходящие из одного и того же китайского источника, эволюционируют в дискурсивные маркеры с разными функциями. Используемые данные взяты из четырех электронных корпусов и одной базы данных, которые охватывают всю историю китайского и японского языков. Китайское *yuánlái* развилось из существительного *yuán* «происхождение» и суффикса *lái* «приходить», который был добавлен в VIII в. н. э. под влиянием тенденции к разделению на слоги в традиционном китайском языке. Оно вошло в японский язык через письменные тексты примерно в тот же период. И *yuánlái*, и *ganrai* изначально имели значения «первоначально, с самого начала». Данное исследование показывает, что их семантические изменения были мотивированы дифференциацией прагматических импликатур, в частности, Q-принципом и R-принципом, предложенными Лоуренсом Р. Хорном (Horn 1984, 2012a, 2012b). *Yuánlái* следовал траектории, начинающейся с лексического значения «ранее» с признаком [+контраст], эволюционируя в маркеры миратива, фона и подтверждения. ДМ *ganrai* произошел от лексического значения «с начала до сих пор, всегда» с признаком [-контраст], позже интерпретируемого как «по природе, по своей сути», и в конечном итоге трансформировался в уточняющий маркер. Результаты этого исследования предлагают новые версии появления дискурсивных маркеров из общего китайского источника через языковой контакт в синосфере. Более того, выявлено, что ограничения лексических значений могут влиять на появление потенциальных функций ДМ.

Ключевые слова: *дискурсивный маркер, обстоятельство времени, историческое развитие, прагматическая импликатура, китайско-японский языковой контакт*

1. Introduction

The language contact between Chinese and Japanese has a long history, dating back to the 5th century CE when the Japanese began to use Chinese characters (*kanji* in Japanese) for writing. The most remarkable aspect of Chinese-Japanese language contact is the extensive borrowing of Chinese vocabulary by Japanese through written texts. According to a survey on 70 magazines published in 1994, about one third of Japanese vocabulary are Sino-Japanese words (NINJAL 2005: 32). Some of the loanwords have evolved into discourse markers (DMs). Their developmental pathways often differ from their Chinese counterparts, but

sometimes share commonalities (see Rhee et al. 2021, Shibasaki & Higashiizumi in preparation).

Chinese *yuánlái* and Japanese *ganrai* serve as a prime example of Chinese-Japanese language contact. They both consist of a morpheme meaning "origin, beginning" (*yuán* and *gan*) and a morpheme meaning "to come" (*lái* and *rai*), and thus they both literally mean "from the beginning, originally." Ever since their first occurrences, *yuánlái* and *ganrai* have been used as temporal adverbials, which eventually gave rise to DM functions. In addition, their nominal uses with the meaning of "the past, origin" have persisted till today.

According to the Center for Chinese Linguistics (CCL) Corpora, Chinese *yuánlái* first appeared as 元來 in the Tang Dynasty (618–907 CE). This written form lasted until the late 14th century when 元 began to be gradually supplanted by 原, a homonym of 元 with the similar meaning, because of the taboo on using the same Chinese character as in the name of the preceding Yuan (元) Dynasty (1271–1368 CE), or the name of the first emperor Yuanzhang Zhu (朱元璋) of the Ming Dynasty (1368–1644 CE) (cf. M. Li 2019, Dong 2020). In Present-Day Chinese, *yuánlái* is written as 原來 or its simplified form 原来.

元來 entered Japanese in the 8th century CE through written texts. It was first read in the Japanese style (*kun'yomi*) as *moto-yori* 'lit. from the origin' or *hajime-yori* 'lit. from the beginning.' According to the investigation of the Corpus of Historical Japanese (CHJ) and *Shinpen Nihon Koten Bungaku Taikei* (SNKBT, the New Complete Collection of Japanese Classic Literature), the Chinese-style reading (*on'yomi*) *gwanrai* first appeared in a Chinese poem included in *Shasekishū* (circa. 13th c. CE) and had been one of the major readings for 元來 until the 19th century when it gradually changed to *ganrai*. The simplified written form 元来 has become the standard way of writing in Present-Day Japanese. On the other hand, during the 17th century, Chinese new written form 原來 [1] was introduced into Japanese, but it had a low frequency and became obsolete in the 20th century (cf. Wang 2022).

The goal of this paper is to investigate the evolution of Chinese *yuánlái* and Japanese *ganrai* from temporal adverbials to discourse markers. It aims to address the research questions regarding why Chinese *yuánlái* and Japanese *ganrai* evolved into DMs with different functions despite of their common origin, and what roles pragmatic implicatures play in the process of their semantic changes.

This paper is structured as follows. Section 2 introduces some fundamental concepts utilized in this paper. Section 3 provides a brief overview of the data and methodology of this study. In Section 4, contemporary uses and historical contexts of Chinese *yuánlái* and Japanese *ganrai* will be outlined. Section 5 delves into

---

[1] 原來 in Japanese was read as *genrai*, which had the same meaning as *ganrai*. In this paper, Japanese *genrai* will not be discussed because as a synonym of *ganrai*, it only existed for a short period of time.

comparing and illustrating their pathways from temporal adverbials to discourse markers. Section 6 is the conclusion.

2. Theoretical background

2.1. Discourse markers

Discourse markers have been referred to with a variety of terms (see the latest review in Heine et al. 2021: Chapter 1, Traugott 2022: Chapter 4), and are often mentioned in the same breath with pragmatic markers (PMs). Fraser (1996, 1999, 2006) restricted DMs to the subset of PMs. Feng (2019: 216) took a similar stance when classifying Chinese PMs and DMs, claiming that "a DM is connective in nature, while a PM is not necessarily so." In a similar vein, Shibasaki and Higashiizumi (in preparation) differentiate DMs and PMs by postulating that DMs are those which connect preceding and following information textually whereas PMs show the speaker's attitude and are not necessarily connected to the following discourse. To highlight the connective nature of such elements, Traugott (2022) proposed the term "discourse structuring markers" to refer to "Connectors that allow the speaker/writer (SP/W) to signal what relationship they wish the addressee/reader (AD/R) to deduce from the linking of discourse segments in a non-subordinate way." (p. 4) On the other hand, for a discourse marker that also signals the speaker/writer's stance of evaluation, such as Chinese *běnlái* 'originally,' Zhan (2022) dubbed it a "stance discourse marker."

As will be demonstrated in this paper, Chinese DM *yuánlái* has three uses, i.e., justificational, mirative, and background. Justificational and background *yuánlái* function to provide additional information to the preceding discourse, and hence are typical discourse connectors. Apart from clause linking function, mirative *yuánlái* also signals the speaker/writer's attitude of unexpectedness towards the situation s/he is facing, much like "stance discourse markers" proposed by Zhan (2022). In order to cover all the non-lexical uses of *yuánlái*, this paper employs "discourse markers" in a broader sense, the same as what is defined in (1) by Heine et al. (2021: 6). In short, DMs in this paper are essentially discourse connectors, and they may or may not express the attitude of the speaker/writer.

(1) Discourse markers are (a) invariable expressions which are (b) semantically and syntactically independent from their environment, (c) set off prosodically from the rest of the utterance in some way, and (d) their function is metatextual, being anchored in the situation of discourse and serving the organization of texts, the attitudes of the speaker, and/or speaker-hearer interaction.

2.2. From temporal adverbials to discourse markers

The pathway from a temporal adverbial to a discourse marker is not uncommon among world languages, such as English *after all* (Traugott 2018: 26–43, Heine et al. 2021: 93–97), French *enfin* 'at last' (Hansen 2005, Heine et al. 2021: 162–165), and Chinese *běnlái* 'originally' (Zhan 2022), to name a few. Temporal adverbials

contribute to the propositional content of its host clause by anchoring the time of the event being described, and therefore are truth-conditional. When they evolve into discourse markers, they may develop justifying and concessive uses such as English *after all*, or reformulative and interjectional uses such as French *enfin*. This study of Chinese *yuánlái* and Japanese *ganrai* will add to our knowledge about what DM functions will arise from temporal adverbials.

The DM uses of Chinese *yuánlái* have received extensive attention from Chinese scholars. Xing (1985), Tang (2006), Yan (2011) and Zhao & Bai (2022) discussed *yuánlái* in Present-Day Chinese, and M. Li (2019), Dong (2020) and Chen (2021) examined its historical development. As for Japanese *ganrai*, Wang (2022) investigated its variation of written forms in different historical periods, along with comparison with Chinese *yuánlái*, but no study has been conducted on its DM use and semantic change. This paper explores and compares the history of Chinese *yuánlái* and Japanese *ganrai*, with an emphasis on how they evolved into discourse markers.

2.3. Semantic types of *yuánlái* and *ganrai*

Paul Grice initially introduced the concept of implicature during the William James Lectures to address situations where a speaker's intended meaning exceeds the literal interpretation of a specific utterance during communication (Grice 1975, 1989). His well-known general maxims of conversation were later reduced by Horn (1984, 2012a, 2012b) and integrated into a bipartite model of principles cited in (2). The Q-Principle (Q stands for quantity) brings together Grice's Maxims of Quantity and Manner, and the R-Principle (R stands for relation) unites the Maxims of Quantity, Relation and Manner.

(2) a. The Q-Principle (Addressee/hearer-based):
      Make your contribution sufficient. Say as much as you can.
    b. The R-Principle (Speaker-based):
      Make your contribution necessary. Say no more than you must.

Since the literal meanings of *yuánlái* and *ganrai* focus on the origin or beginning of a previous situation, there are two possibilities as to whether the same situation still holds true at the time of utterance. Following M. Li (2014, 2019), this paper distinguishes two semantic types of *yuánlái* (as well as Japanese *ganrai*), namely [+contrast] and [-contrast] as illustrated in (3), which is based on the differentiation of pragmatic implicatures indicated in (2).

(3) a. *yuánlái* [+contrast]:
      According to the Q-Principle, since the speaker/writer has said as much as s/he can, the implicature is that what s/he said is only applicable to the time *yuánlái* denotes, i.e., in the past. Hence the situation at the time of utterance is assumed to differ from the previous one, and thus the sense of contrast arises. It can be translated as "previously" or "originally."
    b. *yuánlái* [-contrast]:

According to the R-Principle, the speaker/writer has only given the necessary information as to the time *yuánlái* denotes, so the implicature is that what s/he said may still hold true at the time of utterance. In this sense, *yuánlái* has the implicit meaning of "from the beginning till now" or "always." In the case of Japanese *ganrai*, this meaning is further interpreted as "by nature" or "inherently."

The two semantic types of temporal adverbial *yuánlái* can be demonstrated by examples (4) and (5), both of which are taken from the historical works in the initial stage of its development. In (4), *yuánlái* is used in a context that contrasts with the present time, whereas in (5), it simply signifies "from the beginning till now." Note that in Present-Day Chinese, adverbial *yuánlái* has the default interpretation of [+contrast] although it is defeasible in certain contexts. As for its nominal use, the interpretation of [-contrast] is still feasible (see Example (9) in Section 4.1). By contrast, the default meaning of Present-Day Japanese *ganrai* is [-contrast], and the cases of [+contrast] are rare.

(4) 元來不見，他自尋常；無故相逢，卻交煩惱。

Yuánlái	*bù*	*jiàn,*	*tā*	*zì*	*xúncháng;*
previously	not	meet	it	naturally	normal
wúgù	*xiāngféng,*	*què*	*jiāo*	*fánnǎo.*	
no.reason	meet	then	bring	trouble	

'In the past when we didn't meet, everything was normal. When we happen to meet now, it brings troubles.'

(c700 Zhuo Zhang, *You Xianku* [CCL, see also M. Li 2019: 370])

(5) 唯黃河以北鎮、幽、魏、路²等四節度元來敬重佛法，不毀拆寺舍，不條流僧尼。佛法之事，一切不動之。

Wéi	*Huánghé*	*yǐběi*	*Zhèn,*		*Yōu,*	*Wèi,*
only	Yellow.River	north	PN		PN	PN
Lù	*děng*	*sì*	*jiédù*		***yuánlái***	*jìngzhòng*
PN	and.so.on	four	military.commander		always	respect
fófǎ,	*bù*	*huǐchāi*	*sìshè,*		*bù*	*tiáoliú*
Buddhism	not	demolish	temple		not	drive.away
sēngní.	*Fófǎ*	*zhī*	*shì,*		*yīqiè*	*bù*
monk.and.nun	Buddhism	GEN	matter		all	not
dòng	*zhī.*					
change	it					

'Only the four military commanders of Zhen, You, Wei, and Lu in the north of the Yellow River always respected Buddhism and did not demolish temples or drive away monks and nuns. Matters related to Buddhism remained unchanged.'

(838–848 *Rutang Qiufa Xunli Xingji*, Vol. 4 [CCL, see also M. Li 2019: 370])

3. Data and methodology

---

² The symbol "、" is a common punctuation mark in Chinese writing. It functions similarly to the comma in English but is specifically used to separate nominal items in a list.

The data used for this paper are drawn from 4 electronic corpora and 1 database. The sources for Chinese and Japanese data are summarized in (6) and (7) respectively. The last access date for all the sources was April 2, 2024.

(6) a. The Center for Chinese Linguistics (CCL) Corpora consists of the Classical Chinese Corpus (circa. 500 BCE–1930 CE) and the Modern Chinese Corpus (1930–present). The two corpora comprise selected written text data amounting to 5,841,676,206 tokens of characters. Available online at http://ccl.pku.edu.cn:8080/ccl_corpus/.
b. The Media Language Corpus (MLC) consists of transcribed Contemporary Chinese data from radio and TV broadcasting during the period 2008–2013, and includes 241,316,530 tokens of characters. Available online at https://ling.cuc.edu.cn/RawPub/.

(7) a. The Balanced Corpus of Contemporary Written Japanese (BCCWJ) is comprised of 104.3 million words, covering 11 contemporary genres of written texts during the period of 1971–2008. Available online at https://shonagon.ninjal.ac.jp/.
b. The Corpus of Historical Japanese (CHJ) consists of selected written text data between the 8th century CE and 1925. Available online at https://chunagon.ninjal.ac.jp/chj/.
c. *Shinpen Nihon Koten Bungaku Taikei* (SNKBT) [The New Complete Collection of Japanese Classic Literature] is a database consisting of 88 volumes of books published by Shogakukan which consists of 236 representative works of Japanese Classic Literature between the 8th and 19th centuries CE. Available online at https://japanknowledge.com/contents/koten/.

The methodology is qualitative. Candidates for analysis are manually searched and identified according to their written forms in Chinese characters. The readings for Japanese *kanji* words in historical literary works rely on Japanese syllabaries attached to Chinese characters, if applicable. The DM uses of Chinese *yuánlái* and Japanese *ganrai* are rather easy to identify since they are placed solely in clause-initial positions except for mirative *yuánlái*, whereas for their lexical uses, they tend to appear in clause-medial positions.

Apart from the data extracted from the above corpora and database, a few constructed examples are employed to test the syntactic positions of *yuánlái* in different uses. The constructed data provides insights into language-user's knowledge and intuition that are not accessible through corpora.

4. Chinese *yuánlái* and Japanese *ganrai*

4.1 *Yuánlái* in Present-Day Chinese

As a temporal adverbial, *yuánlái* is typically used in such contexts as contrasted with the time of utterance indicated by *xiànzài* 'now,' as in (8). Even if there is no such a context, the use of *yuánlái* still has the implication that the current situation differs from that of *yuánlái* clause. Its nominal use also has a default

interpretation of [+contrast], but this reading can be cancelled by using adverbs such as *réng* 'still' as in (9).

(8) 我國在知識產權保護方面正面臨著一些新的問題，一些**原來**在門店銷售的假冒偽劣商品，**現在**正從門店銷售轉移到了互聯網上。

Wǒguó	*zài*	*zhīshì*	*chǎnquán*	*bǎohù*	*fāngmiàn*
our.country	in	intellectual	property	protection	aspect
zhèng	*miànlín*	*zhe*	*yīxiē*	*xīn*	*de*
PROG	face	PROG	some	new	GEN
wèntí,	*yīxiē*	***yuánlái***	*zài*	*méndiàn*	*xiāoshòu*
problem	some	previously	at	physical.store	sell
de	*jiǎmào*	*wěiliè*	*shāngpǐn,*	***xiànzài***	*zhèng*
GEN	counterfeit	inferior	product	now	PROG
cóng	*méndiàn*	*xiāoshòu*	*zhuǎnyí*	*dào*	*le*
from	physical.store	sale	transfer	to	PFV
hùliánwǎng	*shang.*				
internet	LOCZ				

'Our country is facing some new problems in the protection of intellectual property. Some counterfeit or inferior products that used to be sold in physical stores are now being sold on the internet.' (2010 *Jiaodian Fangtan*, CCTV [MLC])

(9) 藍翎雖已離休，但仍經常坐在**原來**的辦公室裡讀書、寫作。

Lán Líng	*suī*	*yǐ*	*líxiū,*	*dàn*	*réng*
PN	although	already	retire	but	still
jīngcháng	*zuò*	*zai*	***yuánlái***	*de*	*bàngōngshì*
often	sit	at	the.past	GEN	office
li	*dúshū,*	*xiězuò.*			
LOCZ	read	write			

'Although Ling Lan has already retired, she still frequently sits in her former office to read and write.' (1994 *Baokan Jingxuan* [CCL])

DM *yuánlái* in Present-Day Chinese has two uses. The first use, henceforth "justificational *yuánlái*," is illustrated in (10). *Yuánlái* is placed in clause-initial position, usually followed by a comma in written texts and a pause in spoken language. It functions as a connector, signaling that the following discourse segment provides the justification or explains the reason for what is said or written in the preceding discourse, corresponding roughly to *it turned out that* in English.

(10) 第二天一大早，吳強就搭車趕到拍戲現場。可他一進門就看到現場已是燈火通明，戲已經開拍了，他嚇了一跳，以為自己遲到了。再仔細一看，是在拍那位名角的戲。**原來**，那個女演員昨晚沒有回去，就住在這家提供拍戲場地的酒店，所以一早就趕著先拍她的戲了。

Dì'èr	*tiān*	*yīdàzǎo,*	*Wú Qiáng*	*jiù*	*dā*
second	day	early.morning	PN	then	take
chē	*gǎn*	*dào*	*pāixì*	*xiànchǎng.*	*Kě*
car	rush	to	filming	location	but
tā	*yī*	*jìn*	*mén*	*jiù*	*kàndao*
he	as.soon.as	enter	door	then	see
xiànchǎng	*yǐ*	*shì*	*dēnghuǒtōngmíng,*	*xì*	*yǐjīng*
location	already	be	brightly.lit	filming	already
kāipāi	*le,*	*tā*	*xià.le.yītiào,*	*yǐwéi*	*zìjǐ*
start	PRF	he	be.startled.PRF	think	himself
chídào	*le.*	*Zài*	*zǐxì*	*yīkàn,*	*shì*
be.late	PRF	then	closely	inspect	be
zài	*pāi*	*nà*	*wèi*	*míngjué*	*de*
PROG	film	that	CLF	famous.actress	GEN
xì.	***Yuánlái,***	*nàge*	*nǚyǎnyuán*	*zuówǎn*	*méiyǒu*
scene	turn.out	that	actress	last.night	not
huíqù,	*jiù*	*zhù*	*zài*	*zhè*	*jiā*
return	then	stay	at	this	CLF
tígōng	*pāixì*	*chǎngdì*	*de*	*jiǔdiàn,*	*suǒyǐ*
provide	filming	location	GEN	hotel	so
yīzǎo	*jiù*	*gǎnzhe*	*xiān*	*pāi*	*tāde*
early.morning	then	hurry	first	film	her
xì	*le.*				
scene	PFR				

'Early the next morning, Qiang Wu rushed to the filming location by car. But as soon as he entered, he saw that the set was already brightly lit and the filming had started. He was taken aback, thinking that he was late. But upon closer inspection, he realized that they were filming the scene of a famous actress. As it turned out, the actress had not gone back home the previous night and stayed in the hotel where the filming location was provided, so they started filming her scenes early in the morning.'

(2005 Qingkui Bian, *Zhongguo Beipiao Yiren Shengcun Shilu* [CCL])

While justificational *yuánlái* is usually found in narratives with preceding discourse, written or spoken, DM *yuánlái* of the second use solely appears in colloquial language with or without preceding discourse, cf. (11) and (12). It is typically used in such a context that the speaker finds out something not known to him/her before. This paper follows Tsai & Yang (2022) and refers to this use as "mirative" because it always carries the tone of unexpectedness[3].

---

[3] Zhao & Bai (2022) referred to DM *yuánlái* as *jie fanyuqi biaoji* (marker of removing counter-expectation) because according to their analysis, the preceding discourse is what the speaker/writer did not expect, and the function of *yuánlái* to state the reason for his/her counter-expectation. Justificational and mirative uses of *yuánlái* are dubbed *shiyin xing* (explanatory) and *xingwu xing* (enlightening) functions respectively in Zhao & Bai (2022).

(11) 當她的目光落在窗臺上那個頑皮的史諾比上時，一下蹦起來，跑過去把它拿在手裡：「呀，我還以為你早在路上隨手丟了呢，**原來**你又把它帶回來了！」

Dāng	*tāde*	*mùguāng*	*luò*	*zai*	*chuāngtái*
when	her	gaze	fall	on	window.sill
shang	*nàge*	*wánpí*	*de*	*Shĭnuòbĭ*	*shang*
LOCZ	that	naughty	GEN	Snoopy	LOCZ
shí,	*yīxià*	*bèng*	*qĭlái,*	*păo*	*guòqù*
when	suddenly	jump	up	run	over
bă	*tā*	*ná*	*zài*	*shŏu*	*li:*
ACC	it	take	in	hand	LOCZ
"*Yā,*	*wŏ*	*hái*	*yĭwéi*	*nĭ*	*zăo*
oh	I	once	thought	you	already
zài	*lù*	*shang*	*suíshŏu*	*diū*	*le*
on	way	LOCZ	casually	throw.away	PRF
ne,	***yuánlái***	*nĭ*	*yòu*	*bă*	*tā*
SFP	unexpectedly	you	again	ACC	it
dài	*huilai*	*le*!"			
bring	back	PRF			

'When her gaze fell on the naughty Snoopy on the windowsill, she suddenly jumped up and ran over to pick it up and take it in her hand, saying, "Oh, I thought you had thrown it away on the way, but you brought it back again!"' (2002 *Dongzhi* [CCL])

(12) 身穿紅衣的楊欣，如火球一樣沖入狼陣。狼見狀嚇得四處逃竄。狼嚇跑後，楊欣喃喃地說：「**原來**狼也怕人。」

Shēnchuān	*hóngyī*	*de*	*Yáng Xīn,*	*rú*	*huŏqiú*
wear	red.clothes	GEN	PN	be.like	fireball
yīyàng	*chōng*	*rù*	*láng*	*zhèn.*	*Láng*
same	rush	into	wolf	pack	wolf
jiànzhuàng	*xià*	*de*	*sìchù*	*táocuàn.*	*Láng*
see.this	be.frightened	ADV	all.directions	run.away	wolf
xiàpăo	*hòu,*	*Yáng Xīn*	*nánnán*	*de*	*shuō,*
flee	after	PN	murmur	ADV	say
"***Yuánlái***	*láng*	*yĕ*	*pà*	*rén.*"	
unexpectedly	wolf	also	fear	human	

'Xin Yang, dressed in red, rushed into the wolf pack like a fireball. The wolves were frightened and ran away in all directions. After the wolves had fled, Yang Xin murmured, "So wolves are also afraid of humans."' (1996 *People Daily* [CCL])

Another difference between mirative *yuánlái* and justificational *yuánlái* lies in their syntactic positions. Mirative *yuánlái* may appear in clause-initial or clause-medial position, whereas justificational *yuánlái* occurs only clause-initially. When *yuánlái* is in medial position, i.e., right after the subject or the topic of the clause, the clause may be ambiguous between temporal and mirative readings, cf. (13a). The latter is often but not obligatorily accompanied by intonation of surprise. When

yuánlái is in initial position, it may have three readings, i.e., temporal, mirative or justicational, cf. (13b). In the case of justificational reading, preceding discourse is indispensable. Not surprisingly, two *yuánlái* may cooccur in one clause, with mirative or justificational *yuánlái* in initial position and temporal *yuánlái* in medial position, as illustrated in (13c)[4].

(13) a. 他原來在北京工作。

Tā	***yuánlái***	*zài*	*Běijīng*	*gōngzuò.*
he	*yuánlái*	in	Beijing	work

  Temporal: 'He used to work in Beijing.'
  Mirative: 'So he is working in Beijing!' (constructed example)

  b. 原來他在北京工作。

Yuánlái	*tā*	*zài*	*Běijīng*	*gōngzuò.*
yuánlái	he	in	Beijing	work

  Temporal: 'He used to work in Beijing.'
  Mirative: 'So he is working in Beijing!'
  Justificational: 'It turned out that he is working in Beijing.' (constructed example)

  c. 原來他原來在北京工作。

Yuánlái	*tā*	***yuánlái***	*zài*	*Běijīng*	*gōngzuò.*
yuánlái	he	*yuánlái*	in	Beijing	work

  Mirative + temporal: 'So he used to work in Beijing!'
  Justificational + temporal: 'It turned out that he had worked in Beijing.'
                                    (constructed example)

4.2. A brief history of *yuánlái*

As the predecessor of *yuánlái*, monomorphemic *yuán*, a noun meaning "origin," had been used as a temporal adverbial with the same lexical meaning as *yuánlái* (M. Li 2019, Dong 2020). *Yuánlái* arose through combining *yuán* and *lái*, which is a recurrent process called "disyllabification" in the history of Chinese. *Lái* has been a temporal suffix from the Middle Chinese (3th c. CE to 10th c. CE), and attached to some monosyllabic temporal nouns and adverbs to create disyllabic words. It originally meant "to come," but as a suffix it can be interpreted as "from" or even meaningless because there was no obvious semantic change when *yuánlái* substituted *yuán*[5].

---

[4] (13c) is grammatical but slightly unnatural to some native speakers because two *yuánlái* are used adjacently and with different meanings. This sentence will be improved if the second *yuánlái* is replaced by *yǐqián* or *zhīqián* both meaning 'previously.'

[5] A similar case of *běnlái* 'originally' developing from monomorphemic *běn* 'root, basis, origin' is discussed by Zhan (2022). For the history of other Chinese disyllabic X-*lái*, where X is a temporal noun or adverb, see Z. Li (2019: 55–62) and Chen (2021).

The development of *yuánlái* can be divided into three stages. At Stage I (from 7th c. CE), *yuánlái* was a temporal adverbial or a noun with the interpretation of [+contrast] and [-contrast], cf. (4) and (5)[6]. At Stage II (from 9th c. CE), mirative *yuánlái* appeared, as exemplified by (14) and (15). It remains a mystery how the adverbial use gave rise to the function of mirativity because no bridging context of the two uses has been discovered (see also M. Li 2019).

(14) 卻是偶然行未到，**元來**有路上寥天。

Què	*shì*	*ǒurán*	*xíng*	*wèi*	*dào,*
but	be	by.chance	go	not.yet	reach
yuánlái	*yǒu*	*lù*	*shàng*	*liáotiān.*	
unexpectedly	have	road	ascend	vast.sky	

'It was by chance that the path had not yet been reached. Contrary to my expectations, there was a way to ascend the vast sky.' (c850 Gan Fang, *Tizeng Li Jiaoshu* [CCL])

(15) 師因半夏上黃蘗，見和尚看經。師雲：“我將謂是個人，**元來**是揹黑豆老和尚。”

Shī	*yīn*	*bànxià*	*shàng*	*Huángbò,*	*jiàn*
master	in	early.summer	go.up	PN	see
héshang	*kàn*	*jīng.*	*Shī*	*yún,*	“*Wǒ*
monk	read	scripture	master	say	I
jiāngwèi	*shì*	*gè*	*rén,*	***yuánlái***	*shì*
thought	be	that	person	unexpectedly	be
ǎn	*hēidòu*	*lǎo*	*héshang.*”		
collect	black.bean	old	monk		

'The Master went up to Huangbo Mountain in early summer and saw a monk reading scriptures. The Master said, "I thought he was another person. I didn't know he is the old monk who collects black beans[7]."' (c860 *Zhenzhou Linji Huizhao Chanshi Yulu* [CCL])

At Stage III (from 16th c. CE)[8], background *yuánlái* and justificational *yuánlái* came into being. Background *yuánlái* offers background information for the referent mentioned in the preceding context, and justificational *yuánlái*, as stated in Section 4.1, provides the reason. The former is illustrated in (16), where *yuánlái* initiates a detailed account of the profile of *bǎozhèng* 'village head.' It is named "background" because the discourse segment prefaced by *yuánlái* deviated from the main slot of the story and thus not indispensable to the storyline. In (16), the narratives about Heng Lei, the local soldiers and the big man resumed after the

---

[6] It is unknown when adverbial *yuánlái* lost the interpretation of [-contrast].

[7] "Black beans" is a metaphor for "scriptures" probably because Chinese characters written in black look like black beans from afar.

[8] M. Li (2019) and Dong (2020) used some earlier examples which were claimed to be taken from works of the Song Dynasty (960–1276 CE), but since the dating of these works they cited is not considered authentic among scholars, this paper sets the starting point of Stage III at a much later period.

inserted introduction to the village head, who was not on stage in this scene. On the other hand, as illustrated in (17), the story will become apparently incomplete if the discourse segment following justificational *yuánlái* is omitted. The background use of *yuánlái* is obsolete in Present-Day Chinese.

(16) [At that time, Heng Lei arrived at the Lingguan Temple and saw a big man sleeping on the altar. The local soldiers approached and tied him up, taking him away from the temple. It was still early in the morning, just before dawn. Heng Lei said, "Let's take this guy to the Chao family's estate to get some food, and then take him to the county office for questioning.]

一行眾人卻都奔這保正莊上來。**原來**那東溪村保正，姓晁名蓋，祖是本縣本鄉富戶。

Yī	*xíng*	*zhòngrén*	*què*	*dōu*	*bèn*
one	group	people	then	all	head
zhè	*bǎozhèng*	*zhuāng*	*shang*	*lái.*	***Yuánlái***
this	village.head	estate	LOCZ	come	actually
nà	*Dōngxī*	*cūn*	*bǎozhèng,*	*xìng*	*Cháo*
that	PN	village	village.head	surname	PN
míng	*Gài,*	*zǔ*	*shì*	*běn*	*xiàn*
given.name	PN	ancestor	be	this	county
běn	*xiāng*	*fùhù.*			
this	town	wealthy.family			

'The group of people headed to the Chao family's estate. Actually, the village head of Dongxi was named Chao Gai, whose ancestors were wealthy in the same county and town.'
[He was always ready to help people in need and was willing to make friends with heroes from all over the world. Anyone who came to him for help, no matter how good or bad, was allowed to stay at his estate… Heng Lei and the local soldiers escorted the big man to the front of the estate and knocked on the door…] (c1524 Nai'an Shi, *Shuihu Zhuan* [CCL])

(17) [When Wu Song saw the tiger flip over and come back, he wielded his stick in both hands and put all his strength into one strike from mid-air. Only a loud sound was heard as the branches and leaves of the tree were chopped off and fell.]

定睛看時，一棒劈不著大蟲，**原來**打急了，正打在枯樹上，把那條哨棒折做兩截，只拿得一半在手裡。

Dìngjīng	*kàn*	*shí,*	*yī*	*bàng*	*pībùzháo*
gaze	look	when	one	stick	miss
dàchóng,	***yuánlái***	*dǎ*	*jí*	*le,*	*zhèng*
tiger	turn.out	hit	hurriedly	PRF	just
dǎ	*zài*	*kūshù*	*shang,*	*bǎ*	*nà*
hit	on	dead.tree	LOCZ	ACC	that
tiáo	*shàobàng*	*shé*	*zuò*	*liǎngjié,*	*zhǐ*
CLF	stick	break	into	two.halves	only
nádé	*yībàn*	*zài*	*shǒu*	*li.*	
hold	one.half	in	hand	LOCZ	

'Upon closer inspection, the strike missed the tiger. It turned out that it was in such a hurry that he hit the dead tree instead, breaking the stick in half and leaving only one half in his hand.'

(c1524 Nai'an Shi, *Shuihu Zhuan* [CCL])

4.3. *Ganrai* in Prensent-Day Japanese

In most cases, Japanese *ganrai* displays the feature of [-contrast], corresponding roughly to *from the beginning till now* or *by nature* in English, as in (18). It is usually not equivalent to Present-Day Chinese *yuánlái* since *yuánlái* has lost this meaning. On the other hand, *ganrai* is also used as a temporal adverbial, meaning "originally" or "previously," as in (19), which is the exact equivalent of *yuánlái*. Like *yuánlái*, *ganrai* also has the nominal use, cf. (20). *Ganrai* in both (19) and (20) bears the feature of [+contrast].

(18) 私は元来怠け者の性格で、仕事や人間関係について消極的です。三十四歳になる今日まで彼女がいません。しかし、結婚はしたいと思っています。

Watasi=wa	**ganrai**	*namakemono=no seikaku=de,*	
I=TOP	by.nature	lazy.person=GEN personality=ESS	
sigoto=ya	*ningen~kankei=ni*	*tuite*	*syookyokuteki=des-u.*
work=ENU	people~relationship=DAT	concerning	passive=COP-NPST
Sanzyuuyon-sai=ni nar-u		*konniti=made*	*kanozyo=ga*
34-years.old=DAT become-NPST		today=LIM	girlfriend=NOM
i-mase-n.	*Sikasi,*	*kekkon=wa*	*si-ta-i=to*
have-POL-NEG	however	marriage=TOP	do-DES-NPST=QUO
omot-te	*i-mas-u.*		
think-GER	be-POL-NPST		

'I am inherently lazy and passive when it comes to work and interpersonal relationship. I'm 34 years old now and I've never had a girlfriend. However, I do want to get married.'

(2005 *Yahoo Chiebukuro* [BCCWJ])

(19) ストレスという言葉は「外力によるひずみ」という意味で、元来は工学用語である。

Sutoresu=to	*i-u*	*kotoba=wa*	*gairyoku=ni*
stress=QUO	say-NPST	word=TOP	external.force=DAT
yor-u	*hizumi=to*	*i-u*	*imi=de,*
cause=NPST	strain=QUO	say-NPST	meaning=ESS
***ganrai**=wa*	*koogaku~yoogo=de*	*ar-u.*	
originally=TOP	engineering~term=ESS	COP-NPST	

'"Stress" means "strain caused by external forces." It is a term that originally came from engineering.' (2004 Shigeta Saito, "*Utsu*" *kara Genki ni Nareru Hon* [BCCWJ])

(20) 山本という姓は、もちろんこの養子先のもので、<u>元来</u>の姓は出口という。

Yamamoto=to	*i-u*	*sei=wa,*	*motiron*
PN=QUO	say-NPST	surname=TOP	of.course
kono	*yoosi-saki=no*	*mono=de*	***ganrai**=no*
this	adapted.son-place=GEN	thing=ESS	origin=GEN
sei=wa	*Deguchi=to*	*i-u.*	
surname=TOP	PN=QUO	say-NPST	

'The surname Yamamoto is, of course, the name of his adopted family, and his original surname was Deguchi.' (2005 Tomohiko Suzuki et al., *Nihon Autorō Retsuden: Oyabun* [BCCWJ])

While lexical *ganrai* can occur both clause-medially and clause-initially, DM *ganrai* is exclusively placed in clause-initial position, providing additional information for the referent mentioned in the preceding discourse. I will refer to this use as "elaborative."[9] As exemplified by (21), *ganrai* is typically followed by the construction N=*to i-u mono=wa* 'N=QUO say-NPST thing=TOP,' which is usually employed to illustrate the essence of N. Sometimes *ganrai* is also followed by "N=TOP," as in (22). In both cases, N is the main topic of the preceding discourse, such as "film" in (21) and "(Seoul's) subway" in (22). This use of elaborative *ganrai* resembles the background *yuánlái* (see Section 4.2). While both serve to offer supplementary details about a preceding referent, the distinction lies in that the background information introduced by the *yuánlái* clause is entirely fresh within the discourse, whereas the *ganrai* clause adds further elaboration.

(21) [Director Sidney Lumet is known for his societal dramas such as *12 Angry Men*, *Serpico*, *Dog Day Afternoon*, and *Network*. However, his talent is also fully utilized in suspense and mystery films. The screenplays are written by British craftsperson Paul Dehn, known for the legendary work *Orders to Kill* and the bizarre work *The Night of the Generals*.]
特筆すべきはミステリ映画にスター・システムを導入した点だ。<u>元来</u>映画というものは小説とは違って、配役ひとつで観客の興味をひくかひかないかが、決定するきらいがある。

Tokuhitu~su=beki=wa	*misuteri~eiga=ni*	*sutaa~sisutemu=o*	*doonyuu~si-ta*
worth.attention~do=DEO=TOP	mystery~film=DAT	star~system=ACC	introduce~do-PST
ten=da.	***Ganrai***	*eiga=to*	*i-u*
point=COP	actually	film=QUO	say-NPST
mono=wa	*syoosetu=to=wa*	*tigat-te,*	*haiyaku*
thing=TOP	novel=QUO=TOP	differ-GER	casting
hitotu=de	*kankyaku=no*	*kyoomi=o*	*hik-u=ka*
one=INS	audience=GEN	interest=ACC	attract-NPST=ITR
hik-ana-i=ka=ga,	*kettei~su-ru*	*kirai=ga*	*ar-u.*
attract-NEG-NPST=ITR=NOM	decide~do-NPST	tendency=NOM	be-NPST

---

[9] What Traugott (2022: Chapter 6) called elaborative markers in English include *also*, *further*, *furthermore*, and *moreover*. They resemble elaborative *ganrai* in that the DM-prefaced discourse segment is the continuation and expansion of the preceding discourse.

'Notably, he introduced the star system to mystery films. Actually, unlike novels, films can often rely heavily on casting to arouse the audience's interest.'

(2003 Nonaka Rodi & Yoshinori Nagata,
Ekkyoosuru Honkaku Misuteri [BCCWJ])

(22) [Anyway, it's convenient that this apartment is right next to the subway station. I have a prepaid card worth 10,000 won (about 1,000 yen), so I can ride without buying a ticket every time.]

つまり、行きたい駅まで運賃はいくらなのか確かめる必要がないのだ。<u>元来</u>、ソウルの地下鉄はとてもわかりやすくできている。

Tumari,	*iki-ta-i*	*eki=made*	*untin=wa*
in.other.words	go-DES-NPST	station=LIM	fare=TOP
ikura=na=no=ka	*tasikame-ru*	*hituyoo=ga*	*na-i=no=da.*
how.much=ADN=NMZ=ITR	check-NPST	necessity=NOM	not.exist-NPST=NMZ=COP
Ganrai,	*Souru=no*	*tikatetu=wa*	*totemo*
actually	PN=GEN	subway=TOP	very
wakari-yasu-ku	*deki-te*	*i-ru.*	
understand-easy-ADV	be.made-GER	be-NPST	

'In other words, I don't need to check how much the fare is to the station I want to go. Actually, Seoul's subway system is very easy to understand.'

(2004 Kōtaro Sawaki: *Sakazuki: World Cup* [BCCWJ])

4.4. A brief history of *ganrai*

According to the investigation of SNKBT and CHJ, 元来 in the history of Japanese was read in multiple ways, including Japanese-style *motoyori* or *hajimeyori*, and Chinese-style *gwanrai. Gwanrai* changed to *ganrai* in the 19th century CE.

Examples (23) and (24) are two of the earliest occurrences of 元来 [10] with the readings of *motoyori* and *hajimeyori* respectively, where *moto* means "origin," *hazime* means "beginning," and *yori* means "from" or "since." At this stage, 元来 only appeared in written texts which were mere imitation of Chinese writing style.

(23) 此地平原、<u>元来</u>無岡。

Kono	*tokoro=wa*	*hara=ni*	*si-te,*
this	place=TOP	plain=ADV	do-GER
motoyori	*oka*	*nakari-ki.*	
from.the.beginning	hill	not.exist-PST	

'This place is a plain, and from the beginning there have been no hills.'

(c713 *Pizen no Kuni Pudoki* [SNKBT])

---

[10] The original written form of *rai* should be 來. It was adjusted by the compliers of SNKBT to the Present-Day Japanese form 来.

(24) 即日、敕曰、<u>元来</u>諸家貯於神府宝物、今皆還其子孫。

Sono	*hi=ni,*	*mikotonori~si-te*	*notamaw-aku,*
that	day=ADV	edict.of.emperor~do-GER	say-NMZ
"***Hazimeyori***	*moromoro=no*	*ie=no*	*hokura=ni*
from.the.beginning	all.sorts.of=GEN	family=GEN	divine.treasury=DAT
tume-ru	*takaramono,*	*imasi*	*mina*
store-NPST	treasures	now	all
sono	*uminoko=ni*	*kapes-e"=to*	*notamap-u.*
their	descendants=DAT	return-IMP=QUO	say-NPST

'At that day, in accordance with the imperial edict, all the treasures of each family stored in the divine treasury from the beginning will be returned to their descendants.'
(720 *Nihon Shoki* [SNKBT])

Gwanrai is attested first in a Chinese poem included in *Shasekishū* (circa. 13th c. CE), a collection of Buddhist tales. In the same period, it occurred in the colloquial language documented in a military chronicle tale *Soga Monogatari*, which demonstrated that *gwanrai* had been completely absorbed into Japanese vocabulary. Like *ganrai* in Present-Day Japanese, *gwanrai* predominantly had the interpretation of [-contrast], as (25) illustrates, and only in a few cases the interpretation of [+contrast] is feasible, cf. (26). This tendency prevailed throughout the history of the Japanese language.

(25) 兼隆、この由を聞くよりも、「伊豆の山は、<u>元来</u>、大衆剛強の所なり。なまじひなることをし出だし、世間の人の口にかからんよりは」とて、今度の軍は留まりぬ。

Kanetaka,	*kono*	*yosi=o*	*kiku=yori=mo,*
PN	this	message=ACC	hear=as.soon.as=FOC
"*Izu=no*	*yama=wa,*	***gwanrai,***	*daisyu*
PN=GEN	mountain=TOP	always	monks
kookyoo=no	*tokoro=nar-i.*	*Namazii=nar-u*	*koto=o*
strong=GEN	place=COP-CONC	unnecessary=COP-ADN	thing=ACC
siidasi,	*seken=no*	*hito=no*	*kuti=ni*
start.doing	society=GEN	people=GEN	mouth=DAT
kakara-n=yori=wa"=tote,	*kondo=no*	*ikusa=wa*	
rise-FUT=CMP=TOP=QUO	this.time=GEN	battle=TOP	
todomari-nu.			
stop-PFV			

'Upon hearing this, Kanetaka thought, "The mountains of Izu have always been a place where the monks are strong. It is better to avoid doing unnecessary things and being ridiculed by others," so he called off the battle.' (c1300 *Soga Monogatari* [SNKBT])

(26) 汝が父、<u>元来</u>、箱根の権現を信じ給ひし故、御事をも箱王と呼ばれたり。

Nandi=ga	*titi,*	***gwanrai,***	*Hakone=no*
you=GEN	father	at.that.time	PN=GEN
gongen=o	*sinzi~tamai-si*	*yuwe,*	*okoto=o=mo*
deity=ACC	believe~HON-PST	reason	you=ACC=FOC
Hakowau=to	*yob-are-tar-i.*		
PN=QUO	call-HON-PST-CONC		

'As at that time your father believed in the deity of Hakone, he named you Hakowau (lit. king of Hakone).' (c1300 *Soga Monogatari* [SNKBT])

The earliest attested instance of DM *gwanrai* is (27) in the 17th century CE. *Gwanrai* appeared clause-initially and was followed by X[11]=*to i-u=wa* which is equivalent to X=*to i-u mono=wa* in Present-Day Japanese explained in Section 4.3. In this case, the reading of lexical *gwanrai* is still available since the writer is explaining the nature of "flattering a man." At the same time, the clause-initial position of *gwanrai* and the structuring of background information evokes a new interpretation of *gwanrai* as a discourse connector as it provides additional information for "flattering" which has just been discussed.

(27) [Therefore, by using such means repeatedly, a woman naturally learns the behavior of a man who is already prone to lying. Although she may also lie later, everything a courtesan says and does when trying to please a man is considered a lie by those who do not understand the true nature of love. However, this does not mean that means such as flattering or adulating do not exist.]
これも、焼くぞ、焼かるるぞといふ根源をたづねみれば、いとしも悪からぬなり。<u>元来</u>男をやくといふは、いつはりをいひて、その男によく思はれんとの心から焼くなり。

Kore=mo	*yak-u=zo,*	*yak-ar-ur-u=zo=to*	
this=FOC	flatter-NPST=SFP	flatter-PASS-NPST-ADN=SFP=QUO	
i-u	*kongen=o*	*tazune~mi-reba,*	*itosimo*
say-NPST	origin=ACC	inquire~try-COND	in.the.least
waru-kar-anu=nar-i.		***Gwanrai***	*otoko=o*
bad-VBZ-NEG=COP-CONC		actually	man=ACC
yaku=to	*i-u=wa,*	*ituwari=o*	*ii-te,*
flatter=QUO	say-NPST=TOP	falsehood=ACC	tell-GER
sono	*otoko=ni*	*yo-ku*	*omow-are-n=to=no*
that	man=DAT	good-ADV	think-PASS-INT=QUO=GEN
kokoro=kara	*yaku=nar-i.*		
heart=ABL	flatter=COP-CONC		

'But if one were to inquire about the origin of flattering or being flattered, one would find it truly valuable. Actually, flattering a man was done out of a desire to make him think he was popular, even if it meant lying.' (c1600–1700, *Kana Sōshishū*, [SNKBT])

---

[11] X is a noun or a clause.

5. Discussion

The development of Chinese *yuánlái* and Japanese *ganrai* can be summarized in Table 1. In present day, they are only partially equivalent in the lexical use. However, since *ganrai* [+contrast] is not so commonly used, and [+contrast] is the default reading for *yuánlái*, in most cases they are not translatable. In the domain of DM functions, they have no commonality except for background *yuánlái* and elaborative *ganrai* which are similar in that they both provide additional information to a referent in the previous discourse.

As mentioned in Section 4.2, it remains unclear how mirative *yuánlái* arose from its adverbial use. Dong (2020) proposed a hypothesis that mirative *yuánlái* derived from background *yuánlái*[12] in that the sense of mirativity may have arisen from the contrast between the preceding discourse and the *yuánlái*-prefaced background information. However, she also admitted that mirativity was the first DM function attested in historical documents, so she further hypothesized that background *yuánlái* might have not been recorded although it was used in the colloquial language of that time. Since no empirical evidence is available, for the time being this hypothesis cannot be authenticated.

Table 1. **Meanings and functions of Chinese *yuánlái* and Japanese *ganrai* through time**

Meaning/Function		Chinese *yuánlái*		Japanese *ganrai*	
		Appeared first in	Present-Day Chinese	Appeared first in	Present-Day Japanese
Lexical	Adverbial/Nominal [+contrast]	8th c. CE (4)[13]	In use (8)	14th c. CE (26)	Rare (19)
	Adverbial/Nominal [-contrast]	9th c. CE (5)	Rare (9)	14th c. CE (25)	In use (18)
DM	Mirative	9th c. CE (14, 15)	In use (11, 12)	n.a.	
	Background	16th c. CE (16)	Obsolete	n.a.	
	Justificational	16th c. CE (17)	In use (10)	n.a.	
	Elaborative	n.a.		17th c. CE (27)	In use (21, 22)

Background *yuánlái* is likely to have directly derived from its adverbial use with [+contrast]. Since adverbial *yuánlái* is temporal in its original meaning, the clause containing adverbial *yuánlái* is restricted to past time. By hypothesis, the background information indicated by *yuánlái* was first restricted to situations in the past, but later extended to situations free of temporal restrictions (see the description

---

[12] In Dong (2020), background *yuánlái* is referred to as *chaxu biaoji* 'interpolated narrative marker' and justificational *yuánlái* is called *jieshi biaoji/shuoming biaoji* 'marker of interpretation/illustration'. Mirative *yuánlái* is called *guanlian fuci* 'conjunctive adverb'.

[13] Bracketed numbers in Table 1 indicate the numbers of corresponding examples quoted in Sections 3 and 4.

of present-day *yuánlái* by Xing 1985). The mechanism of semantic change from temporal to textual is a case of metaphor: things happening in the past is liable to be reinterpreted as the background or the reason. Justificational *yuánlái* seem to have evolved one step further to denote the causal relationship between the preceding and the following discourse segments. Compared with lexical *yuánlái*, justificational and mirative uses are freed from temporal restriction because the background information or the new findings are not necessarily what happened in the past.

Another difference between lexical *yuánlái* and DM *yuánlái* lies in that the former can appear in interrogative sentences, but the latter cannot. This echoes the findings that DMs are syntactically unattached (cf. Heine et al. 2021: 50). In terms of syntactic positions, adverbial *yuánlái* and mirative *yuánlái* can be either clause-initial or clause-medial immediately following the subject or the topic of the clause, whereas background and justificational *yuánlái* are basically positioned clause-initially, which is the typical position for a clause linker.

Japanese *ganrai* exhibits a different scenario on its pathway to a DM. Lexical *ganrai* has been primarily used with the interpretation of [-contrast], which acts as the starting point of its DM use. Lexical *ganrai* can appear in clause-initial position and more frequently in clause-medial position, but DM *ganrai* is solely placed clause-initially. Its discourse function is to provide elaborative information which is typically anchored in non-past tense although past tense is also plausible.

The difference between DM *yuánlái* and DM *ganrai* originated from their meanings as temporal adverbials. Both mirative and justificational uses of *yuánlái* stem from the inherent feature of [+contrast]. Mirative *yuánlái* emphasizes the disparity between the situation witnessed by the speaker and its underlying cause, while justificaional *yuánlái* offers a previously undisclosed explanation. On the other hand, ever since Japanese *ganrai* was borrowed from the Middle Chinese, it has been carrying the interpretation of [-contrast], with the strong tendency of denoting "from the beginning till now" or "by nature." In this background, it is unlikely for *ganrai* to develop mirative or justificational DM functions.

Another potential reason for the limited versatility of *ganrai* may stem from its conservative historical development. Unlike *yuánlái*, *ganrai* has predominantly been used in written language and only rarely found in colloquial contexts (for this characteristic of *ganrai* in Present-Day Japanese, also see Hida & Asada 2018: 131). This is a common feature of Sino-Japanese vocabulary in contrast to native Japanese words. In fact, *ganrai* faces competition from synonymous native Japanese words such as *motomoto* 'from the beginning, originally, always, by nature, naturally' and *motoyori* 'from the beginning, originally,' which are often interchangeable with *ganrai* but are applied in a broader range of discourse genres. There exists a tendency for expressions predominantly used in written language, as opposed to those used in everyday speech, to evolve more slowly in terms of their semantic and pragmatic extensions (see Higashiizumi et al. 2024 for relevant discussion).

6. Conclusion

This study examined the developmental pathways of Chinese *yuánlái* and Japanese *ganrai* from temporal adverbials to discourse markers. They both started their journeys as temporal adverbials signifying "originally, from the beginning," but later diverged into distinct trajectories. *Yuánlái* evolved along the pathway characterized by [+contrast], giving rise to mirative, background and justificational markers. On the other hand, *ganrai* started with the feature of [-contrast], and developed into an elaborative marker. The distinction between [+contrast] and [-contrast] could be attributed to the application of Q-principle and R-principle proposed by Horn (1984, 2012a, 2012b).

The findings of this study provide novel perspectives on the emergence of discourse markers derived from shared Chinese etyma through language contact within the Sinosphere. It is revealed that constraints stemming from lexical meanings can influence the emergence of potential DM functions. Furthermore, the contextual genre may also have impact on the versatility of DM uses.

Abbreviations

ABL = ablative; ACC = accusative; ADN = adnominal; ADV = adverbial; CLF = classifier; CMP = comparative; CONC = conclusive; COND = conditional; COP = copula; DAT = dative; DEO = deontic; DES = desiderative; ENU = enumerative; ESS = essive; FOC = focus; FUT = future; GEN = genitive; GER = gerund; HON = honorative; IMP = imperative; INS = instrumental; INT = intentional; ITR = interrogative; LIM = limitative; LOCZ = localizer; N = noun; NEG = negation; NMZ = nominalization; NOM = nominative; NPST = non-past tense; PST = past tense; PASS = passive; PFV = perfective; PN = proper name; POL = politeness; PRF = perfect; PROG = progressive; QUO = quotative; SFP = sentence-final particle; TOP = topic; VBZ = verbalization

References

Chen, Changlai. 2021. *Hanyu "X-lai"shi shuangyinci cihuihua ji yufahua yanjiu [Lexicalization and Grammaticalization of Chinese Disyllabic Words X-lai]*. Beijing: The Commercial Press.

Dong, Xiufang. 2020. Hanyu yupian zhong de chaxu biaoji ji qi yanbian [Markers for narration interposed in Chinese texts and its historical changes]. *Hanyu Xuebao [Chinese Linguistics]* (1). 33–42.

Feng, Guangwu. 2019. Pragmatic markers in Chinese discourse. In Chris Shei (ed.), *The Routledge handbook of Chinese discourse analysis*, 216–243. London: Routledge.

Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6. 167–190.

Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31. 931–952.

Fraser, Bruce. 2006. Towards a theory of discourse markers. In Kerstin Fischer (ed.), *Approaches to discourse particles*, 189–204. North Holland: Elsevier.

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and Semantics, Volume 3: Speech Acts*, 41–58. New York: Academic Press.

Grice, H. Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Hansen, Maj-Britt Mosegaard. 2005. From prepositional phrase to hesitation marker: The semantic and pragmatic evolution of French *enfin*. *Journal of Historical Pragmatics* 6 (2). 37–68.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press.

Hida, Yoshifumi & Hideko Asada. 2018. *Gendai Fukushi Yōhō Jiten [A Usage Guide to Adverbs in Contemporary Japanese]*. (new edition). Tokyo: Tōkyōdō Shoten.

Higashiizumi, Yuko, Reijirou Shibasaki & Keiko Takahashi. 2024. From truth to truly: The case of *shinni* 'truly' in Japanese compared to Chinese, Korean, and Thai counterparts. *Russian Journal of Linguistics* 28 (4). 00–00. https://doi.org/10.22363/2687-0088-40518

Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Deborah Schiffrin (ed.), *Meaning, form, and use in context: Linguistic applications*, 11–42. Washington, D.C.: Georgetown University Press.

Horn, Laurence R. 2012a. Implying and inferring. In Keith Allan & Kasia M. Jaszczolt (eds.), *The Cambridge handbook of pragmatics*, 69–86. Cambridge: Cambridge University Press.

Horn, Laurence R. 2012b. Implicature. In Gillan Russell & Delia Graff Fara (eds.), *The Routledge companion to philosophy of language*, 53–66. New York: Routledge.

Li, Ming. 2014. Shi tan yuyong tuili ji xiangguan wenti [Pragmatic inference and related issues]. *Guhanyu Yanjiu [Research in Ancient Chinese Language]* (4). 42–52.

Li, Ming. 2019. *Ben*, *yuan* lei fuci de yanbian [Evolution of adverbs with the sense of origin-tracing]. *Lishi Yuyanxue Yanjiu [Historical Linguistic Research]* 13. 366–382.

Li, Zongjiang. 2019. *Jindai Hanyu Yuyong Biaoji Yanjiu [Pragmatic Markers in Modern Chinese]*. Shanghai: Shanghai Educational Publishing House.

NINJAL (National Institute for Japanese Language and Linguistics). 2005. *Gendai zasshi no goi chōsa: 1994-nen hakkō 70-shi* [A Survey of Vocabulary in Contemporary Magazines (1994)]. Tokyo: National Institute for Japanese Language and Linguistics.

Rhee, Seongha, Reijirou Shibasaki & Xinren Chen. 2021. Grammaticalization of discourse markers in East Asian Languages: Introduction. *East Asian Pragmatics* 6 (3). 271–281.

Shibasaki, Reijirou & Yuko Higashiizumi. in preparation. The emergence of pragmatic markers in East Asian languages: An introduction. In Yuko Higashiizumi & Reijirou Shibasaki (eds.), *The emergence of pragmatic markers from Chinese compounds in Chinese, Japanese, and Korean*. Leiden: Brill.

Tang, Weiqun. 2006. Fuci *yuánlái* de duojiaodu kaocha [Examining adverbial *yuánlái* from multiple perspectives]. *Changjiang Xueshu [Yangtze River Academic]* (4). 117–123.

Traugott, Elizabeth C. 2018. Modeling language change with constructional networks. In Salvador Pons Bordería & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers: New issues in the study of language change*, 17–50. Leiden: Brill.

Traugott, Elizabeth C. 2022. *Discourse Structuring Markers in English: A Historical Constructionalist Perspective on Pragmatics*. Amsterdam: John Benjamins.

Tsai, Wei-Tien Dylan & Ching-Yu Helen Yang. 2022. On the syntax of mirativity: Evidence from Mandarin Chinese. In Andrew Simpson (ed.), *New explorations in Chinese theoretical syntax: Studies in honor of Yen-Hui Audrey Li,* 431–444. Amsterdam: John Benjamins.

Wang, Yunjiao. 2022. *Ganrai* yu *yuánlái* de lishixing hanri duibi yanjiu [Contrastive study of diachronicity of Japanese *ganrai* and Chinese *yuánlái*]. *Mingguwu Daxue/Pingdong Daxue Jiaoliu ji Lunwen Fabiaohui Lunwenji [Proceedings of the Conference on Literature Communication between Nagoya University and National Pingtung University]* 2. 115–122.

Xing, Fuyi. 1985. Cong *yuánlái* de cixing kan ci de guilei wenti [On the issue of syntactic categorization from the syntactic category of *yuánlái*]. *Hanyu Xuexi [Chinese Language Learning]* (6). 1–4.

Yan, Yi. 2011. *Yuánlái de duoshijiao yanjiu [Study of yuánlái from multiple perspectives].* Master thesis, Nanjing Normal University, Nanjing, China.

Zhan, Fangqiong. 2022. A constructional account of the development of the Chinese stance discourse marker *běnlái*. *Journal of Historical Pragmatics* 23 (2). 245–284.

Zhao, Yu & Xuefei Bai. 2022. *Yuánlái* de fanyuqi xingzhi ji qi xiangguan wenti [The counter-expectation nature of *yuánlái* and related issues]. *Yuyan Yanjiu Jikan [Bulletin of Linguistic Studies]* (1). 53–67.

Bionote:

Wenjiang YANG is Associate Professor of Japanese linguistics at University of Macau, Macao SAR, China. He received his PhD in Japanese linguistics at Peking University, China in 2014. His research interests include morphology, syntax, semantics, historical linguistics and contrastive linguistics, with a special focus on tense, aspect, evidentiality and discourse markers in Present-Day Japanese as well as their diachronic evolution.
e-mail: wenjiangyang@um.edu.mo
https://orcid.org/0009-0000-1802-0630

Сведения об авторе:

Вэньцзян ЯН – доцент кафедры японской лингвистики в Университете Макао, Макао, Китай. Получил степень PhD по японской лингвистике в Пекинском университете, Китай, в 2014 г. Его исследовательские интересы включают морфологию, синтаксис, семантику, историческую лингвистику и контрастивную лингвистику, с особым акцентом на времени, аспекте, эвиденциальности и дискурсивных маркерах в современном японском языке, а также в их диахронической эволюции.
e-mail: wenjiangyang@um.edu.mo
https://orcid.org/0009-0000-1802-0630

Research article / Научная статья

Speaker-orientation meaning and positional shifts of discourse structuring markers

Haiping LONG[1] ⓘ ✉ **and Lei WANG**[1,2] ⓘ

[1] Sun Yat-sen University, Guangzhou, P.R. China
[2] University of Konstanz, Konstanz, Germany
✉lei.4.wang@uni-konstanz.de

Abstract

Discourse structuring markers, and, more generally, discourse markers, are known to be sensitive to their positions, often favoring the clause-initial position. On a diachronic dimension, some discourse structuring markers emerge at the clause-initial position and move to occur the clause-medial position at a later stage of their development, whereas some do not exhibit such positional shifts. The factors that enable such positional shifts have not yet been identified, and this study intends to fill the gap. This study collected discourse structuring markers attested in Chinese historical data from various sources and traced their developmental paths, focusing on their occurrence positions. Two groups of markers, i.e., the *shènzhì* (甚至) 'even' group and the *bùguò* (不过) 'however, but' group, exhibited a marked contrast in terms of their meanings and developmental patterns. An analysis led to the conclusion that the shift from the clause-initial to clause-medial positions is correlated with the presence of the speaker-oriented meanings; those with the speaker-oriented meanings (the *shènzhì*-group) shifted their positions, whereas those without such meanings (the *bùguò*-group) did not. This hypothesized correlation between the speaker-oriented meaning and the positional shift was examined with discourse structuring markers in English, which supported our hypothesis. Further research is needed to ascertain the crosslinguistic validity of the hypothesis, but at the current level of analysis, there is a strong indication that the presence or absence of the speaker-oriented meanings in the discourse structuring markers is the semantic determinant of their positional shift.

Keywords: *discourse structuring marker, speaker-orientation, initial-to-medial change, shènzhì, bùguò*
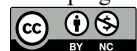
Значение с позиции говорящего и позиционные сдвиги маркеров структурирования дискурса

Хайпин ЛУН[1] 🆔 ✉, Лэй ВАН[1, 2] 🆔

[1]Университет Сунь Ятсена, Гуанчжоу, КНР
[2]Университет Констанца, Констанц, Германия
✉lei.4.wang@uni-konstanz.de

Аннотация

Маркеры структурирования дискурса (в более общем плане – дискурсивные маркеры) чувствительны к своим позициям и часто отдают предпочтение позиции в начале высказывания. В диахроническом измерении некоторые маркеры структурирования дискурса, которые появляются в начале предложения, на более поздней стадии развития перемещаются в середину предложения, в то время как другие не демонстрируют таких позиционных сдвигов. Цель данного исследования – выявить факторы, которые способствуют данным позиционным сдвигам, что ранее не было предметом исследования. Нами были собраны маркеры структурирования дискурса, засвидетельствованные в китайских исторических источниках. Мы проследили пути их развития, сосредоточившись на их позициях. Две группы маркеров, а именно группа *shènzhì* (甚至) 'даже' и группа *bùguò* (不过) 'однако, но', продемонстрировали заметный контраст с точки зрения их значений и моделей развития. Анализ позволяет сделать вывод о том, что сдвиг от начальных позиций высказывания к серединным позициям коррелирует с наличием значений с позиций говорящего. В частности, маркеры со значениями с позиции говорящего (группа *shènzhì*), продемонстрировали позиционный сдвиг, в то время как те, у которых это значение отсутствовало (группа *bùguò*), сохранили ее. Эта предполагаемая корреляция между значением и позиционным сдвигом была исследована с помощью маркеров структурирования дискурса и в английском языке, что подтвердило нашу гипотезу. Необходимы дальнейшие исследования для выяснения обоснованности данной гипотезы в других языках, однако на текущем этапе анализа есть веские основания полагать, что наличие или отсутствие в маркерах структурирования дискурса значений с позиции говорящего является семантическим детерминантом их позиционного сдвига.

Ключевые слова: *маркер структурирования дискурса, ориентация на говорящего, позиционный сдвиг, shènzhì; bùguò*

1. Introduction

Discourse markers (DMs) in general have long been known to have strong preference of occurrence position, favoring the clause-initial position (Traugott 2018, Heine et al. 2021). The preference of initial position is often so robust that Keller (1979: 222) lists it among the three defining characteristics of DMs. However, numerous recent studies report that DMs show greater positional variability (Heine et al. 2021: 89–90). Similarly, discourse structuring markers (DSMs), a subset of pragmatic markers consisting of monofunctional DSMs and

multifunctional DMs (Traugott 2022: 5), the "connectors that allow the speaker/writer to signal what relationship they wish the addressee/reader to deduce from the linking of discourse segments in a non-subordinate way" (Traugott 2022: 4), also typically occur at the clause-initial position (Traugott 2022: 16).

Of particular interest is that while it is true that DSMs typically emerge at the clause-initial position, some of them also occur at the clause-medial position, whereas some of them only occur at the clause-initial position. Furthermore, the positional variability is not arbitrary, i.e., DSMs do not randomly occur at any morphosyntactic position preserving their functions (see section 2 for more). From this state of affairs, there arises a question as to whether certain DSMs have indeed undergone positional shifts from their clause-initial position in the course of their development. A large body of research addressing positionality of DMs and DSMs notwithstanding, the issue of what enables positional variability has not been studied in earnest to date. Thus, the present study intends to fill this research gap by answering the following two research questions: (a) Is there indeed an initial-to-medial change attested in the development of some DSMs?, and (b) If the answer to question (a) is "yes", how do we account for the change, i.e., what enables such positional shifts? The answers to these questions will be sought by investigating the diachronic changes of some Chinese DSMs in this study. The goal of this research is to identify the factors involved in DSM positionality from the diachronic investigation of the DSMs and their contemporary distributional patterns.

In order to pursue the answers to the two key questions, this paper collected historical and contemporary data from various sources in Chinese, paying special attention to the data used by the DM and DSM researchers. Similarly, for discussion of comparable phenomena in English, the data were collected from historical and contemporary corpora as well as earlier research. The data cited in discussion of potential crosslinguistic validity were taken from the reliable native-speaker informants from field work (see section 3 for more).

This study is structured as follows: Section 2 discusses the theoretical background. Section 3 briefly describes data collection. Based on the collected data, Section 4 illustrates two groups of Chinese DSMs that developed in the clause-initial position. It argues that the *shènzhì* (甚至) 'even' group expressed speaker-oriented meanings (and thus are also speaker-oriented adverbials) and acquired the clause-medial usage in later stages of development; the *bùguò* (不过) 'however, but' group did not indicate speaker-oriented meanings (and thus are not speaker-oriented adverbials) and did not acquire the usage in later stages of development. Based on the results of the analysis of Chinese DSMs in Section 4, Section 5 examines the enabling factors of positional shifts and expands the scope to other languages in order to ascertain the potential crosslinguistic applicability of the analysis. Drawing upon Long et al. (2022), we propose a hypothetical initial-to-medial change to account for the similar positional shifts of DSMs expressing speaker-oriented meanings in Chinese and English, further making brief reference to other languages. Section 6 offers the conclusion.

2. Theoretical background

In her seminal work, Traugott (2022) offers a comprehensive analysis on the development of DSMs. Among a number of DSMs in English addressed in the work, *but*, *parenthetically*, and *back to your point* are exemplified in (1):

(1)   a   *Countries separate, they break relations, they … leaders don't speak to each other.* **But** *there's always a coming back together.*
           (ABC Nightline 1990 [COHA BLOG], quoted in Traugott 2022: 4)

      b   *I just typed way too much so I'll stop here.* **Parenthetically**, *I saw a short interview you did.*
           (Blog maverick 2012 [COCA], quoted in Traugott 2022: 151)

      c   *He was designing the B–1 bomber at Wright-Patterson Air Force Base. But* **back to your point**. *So customers are demanding of these authentic flavors.*
           (NPR Here and Now 2014 [COCA], quoted in Traugott 2022: 160)

In the present research, the notion of 'speaker-oriented adverbials' (SpOAs) plays a crucial role in DSM semantics. Following Ernst (2009: 497–502), we use SpOAs to refer to three groups of adverbials:[1] (i) discourse-oriented adverbials that "modify propositions and can be paraphrased by 'I say ADV that P'" (p. 498; see English *honestly* in (2a)), (ii) modal adverbials "which indicate, roughly speaking, speakers' degree of commitment to the truth of P" (p. 498; see English *probably* in (2b)), and (iii) evaluative adverbials "which represent speakers' evaluation of the fact represented by P" (p. 498; see English *luckily* in (2c)) (for other similar classifications of SpOAs, see Ifantidou (1994: 130–155) and Brinton (2008: 8)).

(2)   a   **Honestly**, *I don't know what you mean.*
           (Quoted in Ernst 2009: 498)
      b   *Karen is* **probably** *going to dance a tango.*
           (Quoted in Ernst 2009: 498)
      c   **Luckily**, *Aaron did not fall off his bicycle.*
           (Quoted in Ernst 2009: 498)

Brief remarks on DSMs and SpOAs are in order. These two closely-related terms are not mutually exclusive. One may find, for example, that some linguistic structures may serve to connect two linguistic segments (and thus are DSMs) and at the same time indicate speaker-oriented meanings (and thus are SpOAs); see (3a–c) for the examples of English *even*, Chinese *shènzhì* (甚至) 'even', and Korean *simcie* 'even'.

---

[1] Adverbial structures are notoriously amorphous, falling somewhere between being adverbial phrases and being adverbs. In this study, following Traugott (1995, 2022), Bussmann (2006: 1060), Swan & Breivik (2011), and Long et al. (2022), we use the term *adverbials* as a cover term for all forms of adverbial structures.

(3) a　English
　　　*Mary got a full fellowship from State U. They're **even** paying her way out to visit the Department.*
　　　(Quoted in Kay 1990: 74)

　　b　Chinese
　　　圣如孔子，还假装生病，哄走了儒悲，孟子<u>甚至</u>对齐宣王也撒谎装病。

Shèng	*rú*	*Kǒngzǐ,*	*hái*	*jiǎzhuāng*	*shēngbìng,*	*hǒng*
sage	like	Confucius	still	pretend	be.sick	coax

zǒu	*le*	*Rúbēi,*	*Mèngzǐ*	***shènzhì***	*duì*	*Qí*
go	PFV	Rubei	Mencius	even	to	Qi

Xuānwáng	*yě*	*sāhuǎng*	*zhuāng*	*bìng.*
King.Xuan	also	lie	pretend	be.sick

　　　'A sage like Confucius still pretended to be sick and coaxed Rubei to go; Mencius **even** lied to King Xuan of Qi Nation and pretended to be sick.'
　　　(*Wéi Chéng* [*Fortress Besieged*], quoted in Liu 2012: 256)

　　c　Korean

Ku-nun	*yelsimhi*	*ilha-y*	*sungcin-ul*	*ha-yss-ko*	*hwuey*	*ku-nun*
he-TOP	seriously	work-CSL	promotion-ACC	to-PST-and	later	he-TOP

simcie	*hoycang-i*	*toy-ess-ta.*
even	CEO-NOM	become-PST-DECL

　　　'He worked hard and was promoted, and later he **even** became the CEO.'
　　　(Seongha Rhee, personal communication)

　　Take English *even* (see (3a)) as an example. It is a DSM because it serves to connect two clauses and indicate the relationship between clauses. It is also a SpOA because it expresses a meaning that violates the expectation of the addressee/reader (see Kay 1990: 82–84). Without *even*, there will be no expectation-violating meaning, and it will also be difficult for the addressee/reader to deduce the exact relationship between the two clauses (see (4), a modified example of (3a)).[2]

(4) *Mary got a full fellowship from State U. They're paying her way out to visit the Department.*

　　The literature has conducted a lot of studies on the grammatical changes of linguistic structures that may be seen both as DSMs and SpOAs; see Traugott (1982, 1989, 1995, 2020a, 2020b, 2020c, 2022), Lewis (2000, 2007, 2011), Traugott & Dasher (2002: 152–175), Leuschner (2006), Brinton (2008: 203–218, 2017: 251–283), Lenker (2010, 2014), Heine (2013), Haselow (2015), Heine et al. (2021), and others. However, there are few authors discussing the correlation

---

[2] English *even* (see (3a)) and its counterparts in Chinese (i.e., *shènzhì* 'even', see (3b)) and Korean (i.e., *simcie* 'even', see (3c)) are "largely contentful/semantic" DSMs according to Traugott's (2022: 3–4) classification. English *parenthetically* (see (1b)) also falls into this category according to the author.

between position differences (see the examples of clause-initial DSMs in (1a–c) and examples of clause-medial DSMs in (3a–c)) and the expression of speaker-oriented meanings in these structures.

Traugott (1982, 1989, 1995, 2020a, 2020b, 2020c, 2022), Lewis (2000, 2007, 2011), Traugott & Dasher (2002: 152–175), Leuschner (2006), Brinton (2008: 203–218, 2017: 251–283), Lenker (2010, 2014), Heine (2013), Haselow (2015), Heine et al. (2021), and others study the development of a number of English DSMs from circumstance adverbials and argue that the development typically took place in the initial position of the clauses that they connect.[3] According to Traugott (1995: 9–10), for example, English *in fact* was originally a circumstance adverbial meaning 'in actuality' (see (5a)). It developed into a SpOA expressing counter-expectation meaning in the clause-initial position (see (5b)), and later into a DSM that is "primarily elaborative in function" in the same position (see (5c)).[4]

(5) a *But it is evident **in fact** and experience that there is no such universal Judge, appointed by God over the whole World, to decide all Cases of temporal Right.*
        (*Sermons Tillotson.* 1671: 445 [HC], quoted in Traugott 1995: 9–10)

     b *Whence comes it then, that, **in fact**, the utmost a wise magistrate can propose with regard to popular religions, is, as far as possible, to make a saving game of it ...?*
        (*Hume Dial.* Nat. Relig. Pt.12, publ. 1779: 223, quoted in Traugott 1995: 10)

     c *I should not have used the expression. **In fact**, it does not concern you – it concerns only myself.*
        (Austen, *Emma*, Vol. 3, Chap. 10, 1815: 393, quoted in Traugott 1995: 10)

Dik (1997: 406–408) has suggested a Relator Principle (see the following) to account for the positional preferences of grammatical elements "which serve to link two constituents together, and/or to mark the function(s) of a constituent". Chu & Tao (2008), Dong (2012), Peng (2012), Fan (2015), Gao (2016), Shi & Hao (2018), Li (2020), Yu (2020), and others have extended the principle to account for the positional preference of DSMs in Chinese and the languages in the region of China.[5]

Relators have their preferred position (i) in between two relata; (ii) at the periphery of the relatum with which they form one constituent (if they do so). (Dik 1997: 406)

---

[3] Following Traugott (2022: 70), we use *circumstance adverbial* to refer to an adverbial "with semantics of space, time, manner, instrument, etc."

[4] It is important to note that the development of the DSM *in fact*, as illustrated from (5a) to (5c) and even to (9), is fundamentally based on Traugott (1995). We suspect that, although not explicitly stated, Traugott (2022) may have viewed this development differently. This is a significant issue, and we will reserve it for discussion in future publications.

[5] We thank Prof. Bernd Heine for rightfully pointing out that Dik's (1997: 406–408) Relator Principle was not intended to account for the positional preference of DSMs. These authors' arguments are actually extensions of the principle.

The Relator Principle may explain the positional preference of the English DSM *in fact* in (5c) because its preferred position (i.e., the initial position of one of the clauses that it connects) may also be explained as between two clauses (or two relata). What it cannot explain, however, is the fact that apart from the clause-initial position, some English DSMs may also appear in a clause-medial position; see contrasted pairs in (6a–b), (7a–b), and (8a–b).

(6) a *The umpires' decision is final. **Therefore**, so far as the ICC is concerned, the matter is closed.*
(*The Daily Mirror*. 5437 s-units [BNC])

b *This necessarily entails longer term assistance in comparatively stable situations. We **therefore** particularly value our partnership with SCF through TRANSAID ...*
(*Transport*. u.p., n.d., pp. ??. 645 s-units [BNC])

(7) a *We have the additional problem that the shops in Saint Mary's Street, with the majority of shops in Saint Mary's Street, that is the main street, have no rear accesses. **Consequently** heavy lorries make deliveries and other heavy lorries passing through cause congestion.*
(Suffolk County Council Highways and Transport Committee meeting (Pub/instit). Rec. on 17 Nov 1992 with 2 partics, 746 utts [BNC])

b *In February 1934 Nizan was in the Soviet Union. He **consequently** witnessed the February days, the potential civil war at a distance.*
(*Paul Nizan communist novelist*. Scriven, Michael. London: Macmillan Press Ltd, 1988: 1–120. 1550 s-units [BNC])

(8) a *As a result of all this, Members know that any effort which goes into the preparation of a draft Bill prior to the ballot is likely to be wasted. **In consequence**, many Members who put their names down will have only ill-formed ideas about what they want and it frequently happens that Bills presented are ill-prepared or misconceived.*
(An introduction to British constitutional law. Calvert, Harry. London: Blackstone Press, 1985: 43–148. 1543 s-units. [BNC])

b *Employers will be given a yardstick against which to measure teachers. Parents will, **in consequence**, have more faith in state schools.*
(*What is happening to our primary schools*. Pluckrose, Henry. Oxford: Blackwell, 1987: 1–138. 1757 s-units. [BNC])

Interestingly, even a developed DSM like *in fact*, which preferably occupies a clause-initial position according to the Relator Principle, may also appear in a clause-medial position (see (9)).[6]

---

[6] One may wonder if the clause-medial *in fact* in (9) could be construed as a DSM. We argue that it conforms to Traugott's (2022) definition of a DSM (see Section 1) and should be seen as a DSM; see Cliff (2001: 247), Aijmer (2015: 94), Alonso-Almeida & Álvarez-Gil (2021: 12), and others for arguments supporting clause-medial occurrence of English DSMs such as *in fact*, *actually*, etc., and

(9)   *I should not have used the expression. It **in fact** does not concern you – it concerns only myself.*

One may notice that one reason why Traugott (2022) rejected the terms including *pragmatic marker* (see Brinton 1996: 33, 2017: 2–8, Fraser 1996: 167, Aijmer 2013: 4, and others) and *discourse marker* (see Blakemore 1987: 141, Schiffrin 1987: 31, Fraser 1999: 931, Heine 2013: 1206–1213, and others) in favor of the term DSM was to emphasize the significance of the positional differences of these structures (see Traugott 2022: 61–63). Adopting a different nomenclature, however, does not solve the obvious problem, i.e., Traugott (2022) still cannot fully account for the clause-medial occurrence of some DSMs like English *in fact* in (9). To account for it, one may need to postulate that there has occurred an initial-to-medial change in the course of the development.

In this study, by examining the historical developments of some Chinese and English DSMs, as well as some incidental examples of DSMs in other languages, we take on an unresolved issue in Traugott (2022, n.d.), that is, for some DSMs that developed in the clause-initial position, why do they occupy the clause-medial position in later stages of development?

3. Data sources

The present study has adopted data (including historical data) from Chinese, English, and other languages. Era divisions of Chinese data have followed Sun (2006: 15–20), which are: Old Chinese (771 BCE to 220), Middle Chinese (220 to 960), Early Modern Chinese (960 to 1900), and Modern Chinese (1900 to present). We use the label "historical Chinese" to refer to any period of Chinese that is not Modern Chinese. All the data on historical Chinese cited in this study are based on previous studies on the same topics.

English data cited in this study has five sources: (a) previous studies on the same topics, (b) the British National Corpus (BNC), which contains 100 million words of text from a wide range of genres (e.g., spoken, fiction, magazines, newspapers, and academic), (c) the Corpus of Contemporary American English (COCA), which contains more than one billion words of text (over 25 million words each year from 1990 to 2019) from eight genres (spoken, fiction, popular magazines, newspapers, academic texts, TV and movies subtitles, blogs, and other web pages), (d) the Corpus of Historical American English (COHA), which contains more than 475 million words of text from the 1820s to the 2010s, and (e) the Oxford English Dictionary (OED), which contains over 600,000 entries of the English language. It provides detailed definitions, word origins, and historical usage examples, and is an essential resource for understanding the evolution and nuances of English.

---

see Lenker (2014: 21–25) and Traugott (2022: 212–213) for general arguments supporting clause-medial occurrence of English DSMs.

Unless specifically indicated otherwise, data on DSMs in languages other than Chinese and English were collected by the first author through fieldwork. All the informants for the data are experienced linguists who have had more than five years of experience studying the languages consulted. The names of the informants are indicated in parentheses after each sentence example. The sentence example of historical Korean (see (23a)) should be particularly mentioned because it was not collected through fieldwork; rather, it was adopted from AKORN (https://akorn.bab2min.pe.kr), an online resource platform dedicated to providing Korean literature and ancient texts. The site includes several key features: digitized versions of numerous ancient Korean texts, making them accessible for researchers and enthusiasts; modern Korean dictionaries, grammar books, and other resources to aid in the learning and understanding of the Korean language; multiple language support, facilitating use by international users; and a wealth of scholarly documents and materials, beneficial for academics engaged in Korean studies.

4. Result: Two groups of Chinese DSMs in clause-initial position

The data collection from the historical sources as well as existing literature and their analysis renders significant results as to the development of DSMs. A number of DSMs in Modern Chinese may have developed in the clause-initial position. According to Fang & Jiang (2009) and Liu (2012), the Modern Chinese DSM *shènzhì* (甚至) 'even' developed from a predicate meaning 'so hard/serious as to reach (the level/position of)'; see (10a) and (10b). It originally took an NP as its complement in Old Chinese (see (10a)), and later developed to take a predicate as its complement in Early Modern Chinese (see (10b)).

(10) a 肾不生，则髓不能满，故寒**甚至**骨也。

Shèn	*bù*	*shēng,*	*zé*	*suǐ*	*bù*	*néng*	*mǎn,*
kidney	NEG	live	thus	bone.marrow	NEG	can	be.full

gù	*hán*	***shèn***	***zhì***	*gǔ*	*yě.*
so	cold	hard	reach	bone	FP

'If there are no kidneys, there will be inadequate marrow in the bones, and the coldness will be **so severe as to reach the position of** the bones.'
(*Huángdì Nèijīng*, Sùwèn [*Huangdi Neijing*, Suwen], quoted in Liu 2012: 250; no later than the first century CE)

b 若或父母坚不从所谏，**甚至**怒而挞之流血。

Ruò	*huò*	*fùmǔ*	*jiān*	*bù*	*cóng*	*suǒ*	*jiàn,*
if	sometimes	parents	firmly	NEG	follow	NMLZ	advise

shèn	***zhì***	*nù*	*ér*	*tà*	*zhī*	*liúxuè.*
hard	reach	be.angry	and	beat	him	bleed

'Sometimes the parents insistently rejected his advice. (The situation became) **so difficult as to reach the level** that they beat him until he bled.'
(*Zhūzǐ Yǔlèi*, Dì Èrshíqī Juàn [*Zhuzi's Language Category*, Vol. Twenty-Seven], quoted in Liu 2012: 250; 1270)

In the clause-initial position, it developed into a DSM with speaker-oriented meanings (thus also a SpOA) in Early Modern Chinese no later than the early 17th century; see (11). In an example like (11), the speaker (narrator) uses *shènzhì* to indicate that the condition described by the following clause (i.e., *cháotíng gōngwéi zhīzhōng yǒushí zhàoyòng* 'he was sometimes summoned to the court and inner palace') is deemed to be against the principle of conduct by the speaker/writer (for similar arguments, see Liu (2012: 252–253)).

(11)　　所以公卿大夫都有信着他的，**甚至**朝廷宫闱之中有时召用。

Suǒyǐ	*gōngqīng dàfū*	*dōu*	*yǒu*	*xìnzhe*	*tā*	*de,*
therefore	high-ranking.official	all	have	trust	him	NMLZ

shènzhì	*cháotíng*	*gōngwéi*	*zhī*	*zhōng*	*yǒushí*	*zhàoyòng.*
even	Court	inner.palace	POSS	inside	sometimes	summon

'Therefore, the nobles and high-ranking officials all trusted him, and he was **even** sometimes summoned to the court and inner palace.'
(*Chūkè Pāi'àn Jīngqí*, Dì Sānshíjiǔ Juàn [*First Series of Marvelous Tales at the Pavilion*, Chapter Thirty-Nine], quoted in Liu 2012: 253; the early 17th century CE)

In Modern Chinese, the DSM and SpOA *shènzhì* has developed to occupy a clause-medial position; see (12) (= (3b)).

(12)　　圣如孔子，还假装生病，哄走了儒悲，孟子**甚至**对齐宣王也撒谎装病。

Shèng	*rú*	*Kǒngzǐ,*	*hái*	*jiǎzhuāng*	*shēngbìng,*	*hǒng*
sage	like	Confucius	still	pretend	be.sick	coax

zǒu	*le*	*Rúbēi,*	*Mèngzǐ*	***shènzhì***	*duì*	*Qí*
go	PFV	Rubei	Mencius	even	to	Qi

Xuānwáng	*yě*	*sāhuǎng*	*zhuāng*	*bìng.*
King.Xuan	also	lie	pretend	be.sick

'A sage like Confucius still pretended to be sick and coaxed Rubei to go; Mencius **even** lied to King Xuan of Qi Nation and pretended to be sick.'
(*Wéi Chéng* [*Fortress Besieged*], quoted in Liu 2012: 256)

This kind of change is rather common in Chinese. Our investigation reveals that the following DSMs express speaker-oriented meanings (thus are also SpOAs); see (13), as indicated by their glosses. They have developed in the clause-initial position, and have all occupied a clause-medial position in later stages of development.

(13) a *gùrán* (固然)　　　　　'no doubt'　　　　　　　　(Li 2017: 106)
　　 b *guàibùdé* (怪不得)　　　'no wonder'　　　　　　　 (Jiang & Luo 2019: 60–62)
　　 c *guǒbùqírán* (果不其然) 'it really happens, as expected' (Ye 2016: 193–200)
　　 d *guǒrán* (果然)　　　　　'it really happens, as expected' (Long et al. 2022)
　　 e *hékuàng* (何况)　　　　 'not to say'　　　　　　　　(Li 2014)
　　 f *jìrán* (既然)　　　　　　'now that'　　　　　　　　　(Jiang 2010: 103)

g *nánguài* (难怪)　　　　'no wonder'　　　　　(Xie & Zuo 2009: 30–31)
h *suīrán* (虽然)　　　　　'although'　　　　　　(Xu & Jiang 2010)

Notice that not all the Chinese DSMs developed in a clause-initial position may later develop to occupy a clause-medial position. For instance, according to Shen (2004: 34) and He (2016), Chinese *bùguò* (不过) 'however, but' was first a verbal phrase meaning 'not exceed'. It originally took an NP as a complement (see (14a)), and later developed to take a predicate as a complement (see (14b)) in Old Chinese.

(14)　a　内官**不过**九御，外官**不过**九品。

Nèiguān **bù** **guò** *jiǔyù,* *wàiguān* **bù** **guò** *jiǔpǐn.*
inner.official NEG exceed ninth.grade out.official NEG exceed ninth.rank

'The inner officials **do not exceed** the ninth grade, and the outer officials **do not exceed** the ninth ranks.'

(*Guóyǔ*, Zhōuyǔ [*Discourses of the States*, Discourse of Zhou State], quoted in He 2016: 104; the 5th century BCE to the 4th century BCE)

　　b　公输子之意，**不过**欲杀臣。

Gōngshūzǐ *zhī* *yì,* **bù** **guò** *yù* *shā* *chén.*
Gongshuzi POSS intention NEG exceed want.to kill me

'Gongshuzi's intention **does not exceed** killing me (is merely to kill me).'

(*Mòzǐ*, Gōngshū [*Mozi*, Gongshu], quoted in He 2016: 104; the 5th century BCE to the 4th century BCE)

Shen (2004: 36) argues that *bùguò* developed into a DSM meaning 'but, however' in the clause-initial position in Chinese around the early 20th century (see (15)). It is noteworthy, however, that throughout the development, *bùguò* rarely expressed speaker-oriented meanings (thus was not a SpOA).

(15)　我也没有长策。**不过**这种事情，其势已迫，不能计出万全的。

Wǒ *yě* *méiyǒu* *chángcè.* **Bùguò** *zhè* *zhǒng* *shìqíng,*
I also not.have long-term.plan but DEM kind thing

qí *shì* *yǐ* *pò,* *bùnéng* *jì* *chū* *wànquán de.*
its condition already urgent can't solution come.out perfect NMLZ

'I also do not have a long-term plan, **but** in this kind of situation, the matter is urgent, and it is impossible to come up with a completely foolproof solution.'

(*Lǎocán Yóujì*, Dì Shíliù Huí [*The Travels of Laocan*, Chapter Sixteen], quoted in Shen 2004: 36; 1903)

In contrast with *shènzhì* and the other DSMs listed in (13), the DSM *bùguò* has never developed to occupy a clause-medial position in either historical Chinese or Modern Chinese. This pattern is common in Chinese. Our investigation reveals that the following DSMs, listed in (16), do not express speaker-oriented meanings (thus are not SpOAs). They have all developed in the clause-initial position but have not developed a clause-medial usage in either historical Chinese or Modern Chinese:

(16) a *dànshì* (但是)   'however, but'   (He 2016)
      b *kěshì* (可是)   'however, but'   (Long & Zhou 2024)
      c *rán'ér* (然而)   'however, but'   (Zhu 2012)
      d *zhǐshì* (只是)   'however, but'   (He 2016)
      e *bùrán* (不然)   'otherwise'   (Wu 2021)
      f *érqiě* (而且)   'and, what's more' (Li, Liu & Jiang 2015)
      g *nàme* (那么)   'then'   (Liu 2024: 21–22)

The collected data of DSMs in two different categories, i.e., the *shènzhì*-type and the *bùguò*-type, raise a theoretically important issue on DSM positionality. In other words, if the *shènzhì*-type DSMs exhibit clause-initial and clause-medial occurrences whereas the *bùguò*-type DSMs are restricted to the clause-initial position, one may wonder what factors may contribute to the positional differences of these two groups of Chinese DSMs. We will come to this in the next section.

5. Discussion: Initial-to-medial change of speaker-oriented DSMs

In Section 4, we have described the diachronic changes of two groups of Chinese DSMs developed in the clause-initial position: one group, exemplified by the DSM *shènzhì*, expressed the speaker-oriented meanings (thus are also SpOAs) and may occupy the clause-medial position in later stages of development; the other group, exemplified by the DSM *bùguò*, did not express the speaker-oriented meanings (thus are not SpOAs) and have not occupied the clause-medial position in later stages of development. If our arguments in Section 4 are right, one may naturally hypothesize that DSMs developed in the clause-initial position may occupy the clause-medial position in later stages of development if they express speaker-oriented meanings (thus are SpOAs).

If the expression of speaker-oriented meanings (i.e., thus also being SpOAs) is indeed an important factor triggering the initial-to-medial changes of DSMs developed in the clause-initial position, one might wonder why this is the case. To address this issue, one may need to re-examine the typical positions for the occurrence of SpOAs. The literature generally assumes that SpOAs can occur in multiple positions within a clause, with the clause-initial and clause-medial positions being the most typical (see examples (17a) and (17b); also refer to Quirk et al. (1985: 490–491), Hoye (1997: 148), Biber et al. (1999: 892), Clift (2001), Lenker (2014), Suzuki & Fujiwara (2017: 831), and others, for similar arguments).

(17) a ***Possibly*** *they may have been sent to London*.
         (Quoted in Hoye 1997: 148)

      b *They **possibly** may have been sent to London.*
         (Quoted in Hoye 1997: 148)

The literature also generally argues that SpOAs developed in a clause-initial position; see Traugott (1995, 2022), Traugott & Dasher (2002: 152–175), Brinton (2008: 203–218, 2017: 251–283), Haselow (2015), and Heine et al. (2021), among

others. If this is indeed the case, one might naturally wonder whether SpOAs developed in this typical position should further evolve to occupy the other typical position, namely, the clause-medial position. Long et al. (2022) argue that they indeed do. The authors argue that, cross-linguistically speaking, SpOAs developed in the clause-initial position may also occupy the clause-medial position in later stages of development. Take the English SpOA *perhaps* as an example (see (18a–c)). According to Molencki (2021: 417), it developed from a circumstance adverbial meaning 'by chance' in Middle English (see (18a)), and the whole change took place in the clause-initial position (see (18b)). In Modern English, it may also occupy a clause-medial position (see (18c)).

(18) a ***Parhap*** *y wrussen fye.*
'*Perhaps* I might flee.'
(*Origo Mundi in Norris Anc. Corn. Drama* (Bod 791)1352, ? a1450 (? 1350–1375), quoted in Molencki 2021: 417 and Long et al. 2022: 231)

  b ***Perhaps*** *some Merchant hath inuited him.*
(Shakespeare *Com. Err.* ii.i.4, 1590, quoted in Molencki 2021: 417 and Long et al. 2022: 231)

  c *Horatio **perhaps** lost his mind.*
(Quoted in Long et al. 2022: 231)

We have argued in Section 4 that the first group of Chinese DSMs all expressed speaker-oriented meanings and should also be seen as SpOAs.[7] Considering the SpOA status of these Chinese DSMs, one may naturally expect them to follow the hypothetical initial-to-medial pathway like the other SpOAs discussed in Long et al. (2022). This also explains why the English DSM *in fact*, which typically occupies a clause-initial position (see (19a) = (5c)), may also occupy a clause-medial position (see (19b) = (9)) in later stages of development.[8]

(19) a *I should not have used the expression. **In fact**, it does not concern you – it concerns only myself.*
(Austen, *Emma*, Vol. 3, Chap. 10, p. 393, 1815, quoted in Traugott 1995: 10)

  b *I should not have used the expression. It **in fact** does not concern you – it concerns only myself.*

On the other hand, one may also find that there are English DSMs that do not express speaker-oriented meanings. If they developed in a clause-initial position, following the arguments above, one will naturally expect them not to acquire the clause-medial usage in later stages of development. Traugott (2022: 159–160), for

---

[7] In fact, when discussing the development of these DSMs, most linguists cited in (13) viewed them as SpOAs or as a hybrid of DSM and SpOA.

[8] Our arguments here may give the impression to the authors that we endorse the view that the DSM *in fact* developed in a clause-initial position and later moved to a clause-medial position. We admit that this is a plausible pathway, but we would like to emphasize that the issue calls for further investigation following the publication of Traugott (2022).

example, argues that the English *back to X point* (see (20b)) developed from the phrase *to come back to the point* (see (20a)) in the clause-initial position. Confirming our expectation, this DSM has not acquired a clause-medial usage either in historical English or Modern English.

(20) a *The great body of people in every country are idolaters. They worship the image or form rather than God, the living principle of goodness. But, **to come back to the point**, I believe God does mark ('note') the conduct of men.*
(New Englander and Yale Review [COHA], Traugott 2022: 159, 1843)

b *He was designing the B–1 bomber at Wright-Patterson Air Force Base. But **back to your point**. So customers are demanding of these authentic flavors.*
(NPR Here and Now [COCA], Traugott 2022: 159, 2014)

In the present research, we have looked into the developments of other English DSMs that developed in a clause-initial position and have found that if they did not express speaker-oriented meanings, they usually did not acquire a clause-medial usage in the later stages of development. The interesting part is with English DSMs like *incidentally* and *parenthetically* (see (21a) and (21b)). According to Traugott (2022: 150), these two DSMs have "very sparse" examples of being used to express speaker-oriented meanings, and they almost exclusively appear in the clause-initial position in Modern English (see (21a) and (21b)).

(21) a *"They talked about refrigerators, clearly." Misha rolled his eyes at Arkady. "**Incidentally**, you don't happen to know any murderer repairmen who owe you a favor?"*
(Smith, Gorky Park [COHA], quoted in Traugott 2022: 150, 1981)

b *I just typed way too much so I'll stop here. **Parenthetically**, I saw a short interview you did.*
(Blog maverick [COHA BLOG], quoted in Traugott 2022: 151, 2012)

As indicated above in Section 2, numerous researchers assert that the clause-initial position is a typical position for the development of DSMs. Also as briefly alluded to in Section 2, according to the Relator Principle, the clause-initial position is the typical position for the occurrence of DSMs; similar arguments can be found in Lehmann (1988), Kortmann (1998: 545), Ghezzi & Molinelli (2014), Tsunoda (2018), and others. It seems safe to assume that cross-linguistically speaking, the clause-initial position may be a preferred position for both the development and the occurrence of DSMs.

On the other hand, DSMs occupying the clause-medial position are also attested in a number of languages; see (22a–g) for examples of DSMs occupying a clause-medial position in German, Italian, Greek, Hungarian, Korean, Liujiang Zhuang, and Jieyang dialect of Southern Min.

(22) a German
   Er   *ist*   *reich,*   *das*   *ist*   ***aber***   *nicht*   *wahr.*
   3SG   COP.3SG   rich   DEM   COP.3SG   but   NEG   true
   'He is rich, this **however** is not true.' (Bernd Heine, fieldwork)

  b Italian
   È   *ricco.*   *Questo*   ***peró***   *forse*   *non*   *è*   *vero.*
   be.3SG   rich   this   however   maybe   not   be.3SG   true
   'He's rich. This **however** maybe is not true.' (Francesco Ursini, fieldwork)

  c Greek
   Aftos   *ine*   *plusios,*   *ine*   ***omos***   *poli*   *filikos.*
   3SG   be.3SG   rich   be.3SG   however   very   friendly
   'He is rich; he is **however** very friendly.' (Brian Joseph, fieldwork)

  d Hungarian
   Hatalmas   *zivatar*   *kerekedett*   *és*   *mi*   ***persze***   *bőrig*   *áztunk.*
   large   storm   arose.3SG   and   we   of.course   skin.till   soaked.1PL
   'A fierce storm arose and we **of course** got wet through.'
   (Quoted in Vaskó 2012: 471)

  e Korean
   Ku-nun   *yelsimhi*   *ilha-y*   *sungcin-ul*   *ha-yss-ko*   *hwuey*   *ku-nun*
   he-TOP   seriously   work-CSL   promotion-ACC   to-PST-and   later   he-TOP

   simcie   *hoycang-i*   *toy-ess-ta.*
   even   CEO-NOM   become-PST-DECL
   'He worked hard and was promoted, and later he **even** became the CEO.'
   (Seongha Rhee, personal communication)

  f Liujiang Zhuang (Tai, Kra-Dai)
   Lak⁸   *ka:m¹*   *ni⁴*   *ʔi³*   *ʔdi¹*   *kun¹,*   *mun²*   ***tsi⁶***   *ka:i⁵*   *kun¹*   *le⁵.*
   CLF   citrus   DEM   NEG   good   eat   you   thus   don't   eat   FP
   'The citrus is not delicious. You **thus** don't eat it.' (Fengyu Qin, fieldwork)

  g Jieyang dialect of Southern Min (Min, Sinitic, Sino-Tibetan)
   I³³   ***sui³³zian⁵⁵⁻²²***   *hoʔ²⁻⁵*   *no²¹³⁻⁵³lak⁵,*   *tan³⁵⁻²¹si³⁵⁻²¹*   *sen⁵⁵⁻²²tseʔ²*
   3SG   although   very   endeavor   but   score

   tiam²²⁻²¹ tiam²²⁻²¹   *mo⁴².*
   always   not.good
   '**Although** he is very diligent, his scores are still not good.' (Yanxuan Huang, fieldwork)

  Based on the observations above, we may expand our scope further and attempt to obtain a crosslinguistically valid generalization. In other words, we may ask if crosslinguistically some of the DSMs that developed in the clause-initial position also express speaker-oriented meanings (and thus are also SpOAs), they would acquire a clause-medial usage in later stages of development. Considering that Chinese and English are genetically unrelated languages, if the above argument is

true for Chinese and English, we may also expect it to be (at least partly) true for some other languages.

Confined by the limited data, we currently cannot give a definite answer to the question, but whenever there are historical data supporting the argument that a DSM from a language other than Chinese and English developed in a clause-initial position and at the same time indicated speaker-oriented meanings (and thus are also SpOAs), it indeed seems that we can always expect it to acquire a clause-medial usage in later stages of development. Take the Korean DSM *simcie* 'even' cited in (22e) (= (3c)) as an example. Our investigation of historical Korean reveals that it developed from an adverbial meaning 'such that' (see (23a)) into a DSM meaning 'even' (see (23b)) in the clause-initial position. It was also a SpOA because it exhibited a speaker-oriented meaning indicating that the condition described by the clause is beyond regular people's assessment, as indicated by the translations of (23a) and (23b). *Simcie* has followed the hypothetical initial-to-medial change and has developed a clause-medial usage in Modern Korean (see (22e) = (3c)).

(23) a *Syeyson-i   cangin-ul   tyoha-si-myen   chyengwen-ul   sayo-ko,*
crown.prince-NOM   father-in-law-ACC   like-HON-if   [title]-ACC   hate-and

simcie   *syeyson-i*   Song   *sansak-si-nola,*
such.that   crown.prince-NOM   *Songsa*   dissipate.eliminate-HON-PURP

paskuy   naka-si-myen   Song-ul   ta   sayo-ni…
outside   go.out-HON-if   *Songsa*-ACC   all   hate-as
'When the prince loved his father-in-law, she hated him. (**Her jealousy was) such that** if the prince goes outside in order to neutralize the bad effects from the bad words or characters occurring in *Songsa* (*The History of Song Dynasty*), (she) even hated all the volumes in *Songsa*, so....' (*Hancwunglok* 384 [AKORN]; 1801–1805)

b *Ku-nun   yelsimhi   ilha-y   sungcin-ul   ha-yss-ta.   **Simcie** ku-nun*
he-TOP   seriously   work-CSL   promotion-ACC do-PST-DECL   even   he-TOP

hoycang-i   toy-ess-ta.
CEO-NOM   become-PST-DECL
'He worked hard and was promoted. He **even** became the CEO.' (Seongha Rhee, fieldwork)

Another piece of supporting data comes from Hungarian. The Hungarian SpOA *persze* may also be used as a DSM indicating concessive (see (24a)) or elaborative (see (24b)) meanings; see Furkó (2011) and Vaskó (2012). According to Vaskó (2012), the DSM and SpOA *persze* developed from the particle *persze* that serves to give an affirmative response to the previous utterance (see (25)), and this development occurred in the clause-initial position.

(24) a *Nem   olyan   jó,   mint a könyv, **persze**   azért   izgalmas.*
not   so   good   as   the book of.course   still   exciting
'It is not as good as the book, (but) still exciting, **of course**.'
(Quoted in Vaskó 2012: 475)

b *Ez   a   letiltott csekk. Nem az igazi **persze**,   csak   másolat.*
that   the   canceled check   not   the real   of.course   only   copy
'That's the canceled check. Not the real one, **you know**, it's a copy.'
(House M. D. © NBC Universal Television; quoted in Furkó 2011: 103)

(25) A: *Ismered azt   a   fickót?*
         know   that   the   guy
     B: ***Persze**.*
         of.course
     'A: Do you know that guy?'
     'B: **Of course**.'
     (Quoted in Vaskó 2012: 468)

Furkó (2011: 100) and Vaskó (2012: 467) argue that the Hungarian DSM and SpOA *persze* may also appear in the clause-medial position (see (22d)). Although the two authors did not explicitly state it, we can still infer from their arguments that they support a hypothetical initial-to-medial change for the development of Hungarian DSM and SpOA *persze.*

If the above descriptions of Korean *simcie* and Hungarian *persze* are correct, it seems reasonable to assume that our arguments concerning positional shifts of clause-initial speaker-oriented DSMs may also be true in Korean and Hungarian. In any case, this is an issue that deserves further study, especially more thorough studies focusing on languages other than Chinese and English. However, at the current level of our understanding based on available data from a few languages, the correlation between speaker-oriented meaning of DSMs and their positional shift from the clause-initial to clause-medial position seems to be robust.

As a final remark, we would like to stress that our discussions have left one issue undiscussed, namely that the clause-initial position is not the only position where DSMs may develop. DSMs may also develop in other positions, including the medial position of a clause. For example, according to our investigations, the three English DSMs occupying both a clause-initial and a clause-medial position that we mentioned in Section 2 first acquired the discourse structuring usage in the clause-medial position; see (26a–c), respectively, for the early examples of the clause-medial DSMs *therefore*, *consequently*, and *in consequence*.

(26)   a   *Ysmael pleide hard gamen; Sarra was **ðor-fore** often wroð.*
           'Ysmael played hard games. Sarra was **therefore** often angry.'
           (Gen. & Ex. 1215 [OED], c 1250)

       b   *Vaccination had not been very extensively adopted among the poor, and there were **consequently** fewer cases in proportion, of the varioloid disease among the vaccinated, than in Scotland.*
           (Small-Pox and Vaccination, *North American Review*: October 1821: 286–310 [COHA], 1821)

    c   *In short, a thousand things may take place to prevent Maria from ever again seeing you. She, **in consequence**, sinks into an agonizing suspense concerning your fate.*
(The Wilderness; or, Braddock's Times. *A Tale of the West*, Volume 1 [COHA], 1823)

These DSMs do not express speaker-oriented meanings, either. They have apparently developed the clause-initial usage in later stages of development. Since they did not develop from the clause-initial position, their developments do not contradict the generalizations drawn in this study. For a fuller understanding of DSM positionality more in-depth studies on multiple languages are warranted.

6. Conclusion

In this study, we have discussed the diachronic developments of two groups of Chinese discourse structuring markers (DSMs) that developed in the clause-initial position. We have shown that one group which expressed speaker-oriented meanings (and thus are SpOAs), exemplified by the DSM *shènzhì* (甚至) 'even', has acquired a clause-medial usage in later stages of development. On the other hand, the other group which did not express speaker-oriented meaning (and thus are not SpOAs), exemplified by the DSM *bùguò* (不过) 'however, but', has not acquired a clause-medial usage in later stages of development.

Based on our discussions of the developments of Chinese DSMs, we argued that the expression of speaker-oriented meaning is an important factor triggering the initial-to-medial changes of DSMs that developed in the clause-initial position, an analysis also borne out with the English DSMs. Furthermore, drawing upon the historical data in Korean and Hungarian, we have shown, albeit within limited scope of the data, that the generalization is applicable to the DSMs in these languages. Admittedly, the positionality of DSMs involves multifarious issues that cannot be easily resolved, and more concrete and crosslinguistically robust generalizations call for both microscopic and macroscopic studies across languages.

issue possible. The research reported in this article was funded by the major funding project of the National Social Science Fund of China (The Construction of the Lexicon for Discourse Markers in historical Chinese and Relevant Theoretical Studies; No. 24&ZD250).

References

Aijmer, Karin. 2013. *Understanding Pragmatic Markers: A Variational Pragmatic Approach*. Edinburgh: Edinburgh University Press.

Aijmer, Karin. 2015. Analysing discourse markers in spoken corpora: *Actually* as a case study. In Paul Baker & Tony McEnery (eds.), *Corpora and discourse studies: Integrating discourse and corpora*, 88–109. New York: Palgrave Macmillan.

Alonso-Almeida, Francisco & Francisco J. Álvarez-Gil. 2021. The discourse markers *indeed, in fact, really* and *actually* and their Spanish equivalents in economy. *Revista de Lingüística y Lenguas Aplicadas* 16. 11–23. https://doi.org/10.4995/rlyla.2020.14397

Biber, Douglas, Stig Johansson, Geoffrey Leech & Susan Conrad. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson.

Blakemore, Diane. 1987. *Semantic Constraints on Relevance*. New York: Blackwell.

Brinton, Laurel J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin: Mouton de Gruyter.

Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development.* Cambridge: Cambridge University Press.

Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: Cambridge University Press.

Bussmann, Hadumod. 2006. *Routledge Dictionary of Language and Linguistics*. London/New York: Routledge.

Chu, Zexiang (储泽祥) & Fuping Tao (陶伏平). 2008. Hànyǔ yīnguǒ fùjù de guānlián biāojì móshì yǔ liánxìxiàng jūzhōng yuánzé [The conjunction distributional patterns of Chinese causal complex sentences and the relator principle]. *Zhōngguó Yǔwén* [*Studies of the Chinese Language*] 326 (5). 410–422.

Clift, Rebecca. 2001. Meaning in interaction: The case of *actually*. *Language* 77 (2). 245–491. https://doi.org/10.1353/lan.2001.0074

Dik, Simon C. 1997. *The Theory of Functional Grammar* (Part 1): *The Structure of the Clause* (Second, Revised Edition). Berlin/New York: Mouton de Gruyter.

Dong, Xiuying (董秀英). 2012. Hànyǔ hòuzhì jiǎshè biāojì yǔ liánxìxiàng jūzhōng yuánzé [The postpositional hypothetical markers of Chinese and the relator principle]. *Huáwén Jiàoxué yǔ Yánjiū* [*TCSOL Studies*] 47 (3). 60–65.

Ernst, Thomas. 2009. Speaker-oriented adverbs. *Natural Language and Linguistic Theory* 27 (3). 497–544. https://doi.org/10.1007/s11049-009-9069-1

Fan, Lijun (范丽君). 2015. Liánxìxiàng jūzhōng yuánzé zài zàngmiǎnyǔ jiǎshè fùjù zhōng de fēnbù [Distributional patterns of the Relator Principle in complex sentences of supposition in Tibeto-Burman Languages]. *Mínzú Yǔwén* [*Minority Languages of China*] (3). 67–74.

Fang, Yixin (方一新) & Xinglu Jiang (姜兴鲁). 2009. *Shènzhì* de cíhuìhuà lìchéng [The lexicalization process of *shenzhi*]. *Jiāngnán Dàxué Xuébào (Rénwén Shèhuì Kēxué Bǎn)* [*Journal of Jiangnan University (Humanities & Social Sciences)*] 8 (1). 137–140.

Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6 (2). 167–190. https://doi.org/10.1075/prag.6.2.03fra

Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31 (7). 931–952. https://doi.org/10.1016/S0378-2166(98)00101-5

Furkó, Bálint Péter. 2011. A contrastive study of English *of course* and Hungarian *persze.* In Katalin Balogné Bérces, Kinga Földváry & Rita Mészárosné Kóris (eds.), *HUSSE10-Linx: Proceedings of the HUSSE10 conference: Linguistics volume*, 95–106. Debrecen: Hungarian Society for the Study of English.

Gao, Zailan (高再兰). 2016. Jiè, lián de cícháng tèzhēng yǔ liánxìxiàng qīngzhòng yuánzé — Lèixíngxué shìjiǎo yǔ kuàyǔyán diàochá [The difference between the length of adpositions and conjunctions and Relator Principle of light-heavy: A typology study and cross-language survey]. *Wàiguóyǔ [Journal of Foreign Languages]* 39 (3). 31–42.

Ghezzi, Chiara & Piera Molinelli. 2014. *Discourse and Pragmatic Markers from Latin to the Romance Languages*. Oxford: Oxford University Press.

Haselow, Alexander. 2015. Left vs. right periphery in grammaticalization: The case of *anyway*. In Andrew Smith, Graeme Trousdale & Richard Waltereit (eds.), *New directions in grammaticalization research*, 157–186. Amsterdam: John Benjamins.

He, Xiao (何潇). 2016. Shìlùn xiàndìngxìng fànwéi fùcí jiān zhuǎnzhé liáncí de lìshí yǎnbiàn – Yǐ *dànshì, bùguò, zhǐshì* wéilì [The analysis of diachronic development of limiting-adversative conjunction: The case of *danshi, zhishi* and *buguo*]. *Hànyǔ Xuéxí [Chinese Language Learning]* (6). 103–112.

Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6). 1205–1247. https://doi.org/10.1515/ling-2013-0048

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press.

Hoye, Leo. 1997. *Adverbs and Modality in English*. London: Longman.

Ifantidou, Ely. 1994. *Evidentials and Relevance*. London: University College London dissertation.

Jiang, Lili (姜礼立) & Yaohua Luo (罗耀华). 2019. Huísù tuīlǐ zài yǔyì shíjiě zhōng de yìngyòng – Yǐ *guài bùdé* de yǎnbiàn wéilì [Application of retrospective ratiocination into the process of meaning construal: A case study of the evolution of *guai bude*]. *Gǔhànyǔ Yánjiū [Research in Ancient Chinese Language]* 123 (2). 57–66.

Jiang, Tao (姜涛). 2010. *Jìrán* de yǔfǎhuà guòchéng [The grammaticalization of *jiran*]. *Yǔwén Xuékān [Journal of Language and Literature Studies]* (13). 102–105.

Kay, Paul. 1990. Even. *Linguistics and Philosophy* 13. 59–111. https://doi.org/10.1007/BF00630517

Keller, Eric. 1979. Gambits: Conversational strategy signals. *Journal of Pragmatics* 3 (3/4). 219–238.

Kortmann, Bernd. 1998. Adverbial subordinators in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, Part 1. 457–562. Berlin/New York: Mouton de Gruyter.

Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam/Philadelphia: John Benjamins.

Lenker, Ursula. 2010. *Argument and Rhetoric: Adverbial Connectors in the History of English*. Berlin/New York: Mouton de Gruyter.

Lenker, Ursula. 2014. Knitting and splitting information: Medial placement of linking adverbials in the history of English. In Simone E. Pfenninger, Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja, Marianne Hundt & Daniel Schreier (eds.), *Contact, variation, and change in the history of English*, 11–38. Amsterdam/Philadephia: John Benjamins.

Leuschner, Torsten. 2006. *Hypotaxis as Building-Site: The Emergence and Grammaticalization of Concessive Conditionals in English, German and Dutch*. Lincom Europa.

Lewis, Diana M. 2000. *Some Emergent Discourse Connectives in English: Grammaticalization via Rhetorical Patterns*. Oxford: Oxford University dissertation.

Lewis, Diana M. 2007. From temporal to contrastive and causal: The emergence of connective *after all*. In Agnès Celle & Ruth Huart (eds.), *Connectives as discourse landmarks*, 88–99. Amsterdam: John Benjamins.

Lewis, Diana M. 2011. A discourse-constructional approach to the emergence of discourse markers in English. *Linguistics* 49 (2). 415–443. https://doi.org/10.1515/ling.2011.013

Li, Qianqian (李茜茜). 2017. *Gùrán* de cíhuìhuà jí yǔfǎhuà [The lexicalization and grammaticalization of *guran*]. *Húzhōu Shīfàn Xuéyuàn Xuébào [Journal of Huzhou University]* 39 (7). 105–109.

Li, Sixu (李思旭). 2020. Cóng jiǎshè liáncí wèizhì fēnbù kàn liánxìxiàng jūzhōng yuánzé [On the relator principle from the perspective of the position distribution of hypothetical conjunctions]. *Shìjiè Hànyǔ Jiàoxué [Chinese Teaching in the World]* 34 (2). 215–231.

Li, Yongchun (李永春), Dandan Liu (刘单单) & Lili Jiang (姜莉丽). 2015. Shuāngyīnjié liáncí *ér qiě* shǐyòng qíngkuàng lìshǐ kǎochá [Diachronic investigation into the uses of disyllabic conjunction *er qie*]. *Mǔdānjiāng Dàxué Xuébào [Journal of Mudanjiang University]* 24 (11). 15–17.

Li, Zongjiang (李宗江). 2014. Liáncí *hékuàng* hé *qǐkuàng* shì zěnyàng xíngchéng de? [The evolution of conjunctions *hekuang* and *qikuang*]. *Hànyǔ Xuébào [Chinese Linguistics]* 46 (2). 63–70.

Liu, Chenyang (刘晨阳). 2024. *Nàme* de huàyǔ biāojì gōngnéng [The function of discourse marker *name*]. *Yǔyán Yánjiū [Studies in Language and Linguistics]* 44 (2). 17–24.

Liu, Hongni (刘红妮). 2012. *Shènzhì* de cíhuìhuà yǔ duōzhǒng gōngnéng de xíngchéng [Lexicalization and formation of multiple functions of *shenzhi*]. *Dāngdài Yǔyánxué [Contemporary Linguistics]* 14 (3). 246–259.

Long, Haiping, Francesco Ursini, Bernd Heine & Yaohua Luo. 2022. From separate clause to epistemic adverbial, the neglected source construction and initial-to-medial pathway: Chinese *guoran* 'it really happens'. *Australian Journal of Linguistics* 42 (3–4). 226–250. https://doi.org/10.1080/07268602.2022.2149256

Long, Haiping & Weihua Zhou. 2024. From affirmation to concession: Diachrony of Modern Chinese concessive connective *kě shì* 'but' and its implications to connective formation. *Journal of Historical Pragmatics* Online.

Molencki, Rafał. 2021. The grammaticalization of the epistemic adverb *perhaps* in late Middle and Early Modern English. *Studia Anglica Posnaniensia* 56 (s1). 411–424. https://doi.org/10.2478/stap-2021-0005

Peng, Xiaoqiu (彭小球). 2012. Yìyáng fāngyán yǒubiāo jiǎshè fùjù yánjiū [Study on the marked hypothesis complex sentences of Yiyang Dialect]. *Qíqíhā'ěr Dàxué Xuébào* (Zhéxué Shèhuì Kēxué Bǎn) *[Journal of Qiqihar University (Philosophy & Social Sciences)]* (2). 34–37.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Harlow: Longman.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Shen, Jiaxuan (沈家煊). 2004. Shuō *bùguò* [Remarks on *buguo*]. *Qīnghuá Dàxué Xuébào* (Zhéxué Shèhuì Kēxué Bǎn) *[Journal of Tsinghua University (Philosophy and Social Sciences)]* 19 (5). 30–36.

Shi, Xiuju (史秀菊) & Jingjing Hao (郝晶晶). 2018. Shānxī píngyáo fāngyán fùjù guānlián biāojì *mén* de yǎnbiàn — Cóng hòuzhì dào qiánzhì [From postposition to preposition:

Evolution of the conjunctive marker *məŋ²¹³* in complex sentences in the Shanxi Pingyao dialect]. *Zhōngguó Yǔwén [Studies of the Chinese Language]* (6). 672–680.

Sun, Chaofen. 2006. *Chinese: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Suzuki, Daisuke & Takashi Fujiwara. 2017. The multifunctionality of 'possible' modal adverbs: A comparative look. *Language* 93 (4). 827–841. https://doi.org/10.1353/lan.2017.0059

Swan, Toril & Leiv Egil Breivik. 2011. English sentence adverbials in a discourse and cognitive perspective. *English Studies* 92 (6). 679–692. https://doi.org/10.1080/0013838X.2011.604917

Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics*, 245–271. Amsterdam: John Benjamins.

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meaning in English: An example of subjectification in semantic change. *Language* 65 (1). 31–55. https://doi.org/10.2307/414841

Traugott, Elizabeth Closs. 1995. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at *International Conference of Historical Linguistics XII. University of Manchester,* 13–18 August. https://www.researchgate.net/publication/228691469_The_role_of_discourse_markers_in_a_theory_of_grammaticalization (accessed 12 August 2024).

Traugott, Elizabeth Closs. 2020a. Expressions of stance-to-text: Discourse management markers as stance markers. *Language Sciences* 82. 1–14. https://doi.org/10.1016/j.langsci.2020.101329

Traugott, Elizabeth Closs. 2020b. Is "back to my point" a pragmatic marker? An inquiry into the historical development of some metatextual discourse management markers in English. [Special Issue]. *Catalan Journal of Linguistics*. 13–29. https://doi.org/10.5565/rev/catjl.307

Traugott, Elizabeth Closs. 2020c. The development of "digressive" discourse-topic shift markers in English. *Journal of Pragmatics* 156. 121–135. https://doi.org/10.1016/j.pragma.2019.02.002

Traugott, Elizabeth Closs. 2022. *Discourse Structuring Markers in English: A Historical Constructionalist Perspective on Pragmatics*. Amsterdam/Philadelphia: John Benjamins.

Traugott, Elizabeth Closs. n.d. How is *after all* used in medial position? (unpublished manuscript).

Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Tsunoda, Tasaku (ed.). 2018. *Levels in Clause Linkage: A Crosslinguistic Survey*. Berlin/Boston: Mouton de Gruyter.

Vaskó, Ildikó. 2012. Pragmatic particles indicating expectation: The case of *persze*. *Acta Linguistica Hungarica* 59 (4). 465–486.

Wu, Yuanjia (武远佳). 2021. Liáncí *bùrán* de yǔfǎhuà [Grammaticalization of the conjunction *buran*]. *Hēilóngjiāng Jiàoshī Fāzhǎn Xuéyuàn Xuébào [Journal of Heilongjiang Institute of Teacher Development]* 40 (4). 126–128.

Xie, Xiaoming (谢晓明) & Shuangju Zuo (左双菊). 2009. *Nánguài* de yǔfǎhuà [On the Grammaticalization of *nanguai*]. *Gǔhànyǔ Yánjiū [Research in Ancient Chinese Language]* 83 (2). 30–35.

Xu, Zhaohong (徐朝红) & Jicheng Jiang (蒋冀骋). 2010. Cóng jùfǎ jiǎodù kàn liáncí *suīrán* de chǎnshēng [Emergence of syntactical conjunction *suiran*]. *Húnán Shīfàn Dàxué Shèhuì Kēxué Xuébào [Journal of Social Science of Hunan Normal University]* (3). 122–125.

Ye, Jianjun (叶建军). 2016. *Guǒbùqírán* de xíngchéng jíqí yǎnbiàn [The formation and evolution of *guobu(qi)ran*]. *Zhōngguó Yǔwén [Studies of the Chinese Language]* 371 (2). 192–201.

Yu, Zechao (余泽超). 2020. Cóng hànyǔ língxíng xiàzhǐjù kàn liánxìxiàng jūzhōng yuánzé [Dik's (1997) relator principle and Chinese cataphora]. *Zhèjiāng Shīfàn Dàxué Xuébào* (Shèhuì Kēxué Bǎn) *[Journal of Zhejiang Normal University (Social Sciences)]* 45 (3). 66–71.

Zhu, Min (朱敏). 2012. *Ránér* de yǔfǎhuà yánjiū [Studies on the grammaticalization of *raner*]. *Běifāng Wénxué [Northern Literature]* (4). 112.

Bionotes:

Haiping LONG is a professor at the School of Foreign Languages at Sun Yat-sen University, specializes in grammaticalization, linguistic typology, historical linguistics, and discourse grammar. He has authored over 50 papers in journals such as *Studies in Language*, *Functions of Language*, *Linguistics*, *Journal of Historical Pragmatics* and *Language and Linguistics*. He has led four national social science projects, including a major funding project, and has been awarded the Lü Shuxiang Linguistics Second Prize as well as a provincial award for Philosophy and Social Sciences.
e-mail: longhp@mail.sysu.edu.cn
https://orcid.org/0000-0003-2465-8723

Lei WANG is a PhD student at the School of Foreign Languages, Sun Yat-sen University. She is sponsored by the China Scholarship Council (CSC) and is currently a visiting student in the Department of Linguistics at the University of Konstanz. Her research interests include discourse grammar, Chinese dialects, grammaticalization, and the syntax-prosody interface. She has co-authored six papers in journals such as *Functions of Language*, *Language and Linguistics* and *SAGE Open*.
e-mail: lei.4.wang@uni-konstanz.de
https://orcid.org/0000-0001-6622-5163

Сведения об авторах:

Хайпин ЛУН – профессор Школы иностранных языков Университета Сунь Ятсена, специализируется на грамматикализации, лингвистической типологии, историче- ской лингвистике и грамматике дискурса. Является автором более 50 статей в таких журналах, как *Studies in Language, Functions of Language, Linguistics, Journal of Historical Pragmatics* и *Language and Linguistics*. Руководил четырьмя националь- ными проектами в области социальных наук и был удостоен второй премии имени Люй Шусяна по лингвистике, а также региональной премии по философии и соци- альным наукам.
e-mail: longhp@mail.sysu.edu.cn
https://orcid.org/0000-0003-2465-8723

Лэй ВАН — аспирант факультета иностранных языков Университета Сунь Ятсена. Стипендиат Китайского совета по стипендиям, в настоящее время – приглашенный студент на кафедры лингвистики в Университете Констанца. Ее исследовательские интересы включают грамматику дискурса, китайские диалекты, грамматикализацию и интерфейс синтаксиса и просодии. Соавтор шести статей в журналах *Functions of Language, Language and Linguistics* и *SAGE Open*.
e-mail: lei.4.wang@uni-konstanz.de
https://orcid.org/0000-0001-6622-5163

Research article / Научная статья

The way of *truth*: The case of the Korean discourse marker *cincca* in comparison with Chinese *zhenshi* and *zhende*

Seongha RHEE[1,2] ✉ **and Lin ZHANG[3]**

[1]Faculty of Liberal Arts, Mahidol University, Nakhon Pathom, Thailand
[2]Hankuk University of Foreign Studies, Seoul, Republic of Korea
[3]Seoul National University, Seoul, Republic of Korea
✉srhee@hufs.ac.kr

Abstract
Korean has a number of discourse markers (DMs) of Chinese origin, which carry similar and different functions as compared to Chinese source lexemes. Despite their significance, they have not received much attention, hence the rationale of the present study. The goal of the study is to compare DMs of the same origin in Korean and Chinese to identify similarities and differences, based on the data taken from historical and contemporary sources. The Korean DM *cincca* 'a true thing', composed of *cin* 'tru(th)' and the nominalizer *cca* 'thing, person', presents an interesting grammaticalization scenario into diverse discourse functions. The findings demonstrate that in addition to its original nominal function, the DM also carries an adjectival function of adding genuineness or excellence in quality to a modified noun or an adverbial function of adding emphasis to an adjective or a predicate. From this intensifying function there arise diverse DM functions through the interaction of the source meaning of 'truthfulness' and diverse inferences from the discourse contexts. The Chinese DMs involving the same etymon are *zhende* (from *zhen* 'true' and *de* 'NOMINALIZER') and *zhenshi* (from *zhen* 'true' and *shi* 'be so'). The functions of these two DMs are similar to those of the Korean *cincca*, but the Chinese DM *zhenshi* is negatively-biased by marking the speaker's negative evaluation of the referenced person or event. An exploration on grammaticalization processes and functions in the two languages reveals much commonality but some differences in terms of functional distribution, prosody, and the extent of desemanticization.
Keywords: *grammaticalization, discourse marker, truth, Chinese, Korean*

Путь *истины*: корейский дискурсивный маркер *cincca* в сравнении с китайскими *zhenshi* и *zhende*

Сеонгха РИ[1] (iD) ✉, **Линь ЧЖАН**[2] (iD)

[1]Факультет гуманитарных наук, Университет Махидол, Накхонпатхом, Таиланд
[2]Ханкукский Университет иностранных исследований, Сеул, Республика Корея
[3]Сеульский национальный университет, Сеул, Республика Корея
✉srhee@hufs.ac.kr

Аннотация

В корейском языке есть ряд дискурсивных маркеров (ДМ) китайского происхождения, которые выполняют как схожие, так и различные функции по сравнению с исходными китайскими лексемами. Несмотря на свою значимость, им не уделялось особого внимания в литературе. Цель данной статьи – сравнить дискурсивные маркеры одного и того же происхождения в корейском и китайском языках для выявления их сходств и различий. Исследование проведено на основе данных, взятых из исторических и современных источников. В центре внимания – корейский ДМ *cincca* 'истинная вещь', состоящий из *cin* 'истина' и номинализатора *cca* 'вещь, человек', который представляет интересный случай грамматикализации и приобретения способности выполнять различные функции в дискурсе. Результаты исследования демонстрируют, что в дополнение к первоначальной номинальной функции этот ДМ также выполняет адъективную функцию, добавляя модифицированному существительному значение подлинности или качественности, также адвербиальную функцию добавления акцента к прилагательному или сказуемому, усиливая прилагательное или сказуемое. Из этой усиливающей функции возникли различные функции ДМ посредством взаимодействия исходного значения правдивости и значений, выводимых из различных контекстов. Китайские ДМ, включающие тот же этимон, – *zhende* (от *zhen* 'истинный' и *de* 'НОМИНАЛИЗАТОР') и *zhenshi* (от *zhen* 'истинный' и *shi* 'быть должным'). Функции этих двух китайских ДМ схожи с функциями корейского *cincca*, но китайский ДМ *zhenshi* имеет отрицательную коннотацию, то есть содержит отрицательную оценку говорящим упомянутого человека или события. Исследование процессов и функций грамматикализации в двух языках выявляет много общего, а также выделяет некоторые различия в плане функционального распределения, просодии и степени десемантизации.

Ключевые слова: *грамматикализация, дискурсивный маркер, истина, китайский язык, корейский язык*

1. Introduction

Lexemes denoting 'truth, right' seem to be among the common sources of discourse markers (DMs). In an extensive lexicon of grammaticalization, Kuteva et al. (2019: 374) exemplify the grammaticalization of DMs that originated from the lexemes denoting 'right, true, correct', e.g., English *right*, German *richtig* 'right', Hebrew *naxon* 'right', Turkish *doğru* 'right', etc. The lexicon also lists intensive

markers developed from the lexemes denoting 'true, real', e.g., Archaic Chinese *qing* 'truth', Early Modern Chinese *zhen* 'true', Old French *verai* 'tru(ly), truthful(ly)' (borrowed into Middle English *verray*, Modern English *very*), Hungarian *igaz* 'true', *igaz-án* 'really', Baka *ko* 'truly, really, completely', and American Sign Language TRUE (see individual references cited in the lexicon). Since the intensive markers, also commonly called as 'intensifiers', closely resemble and are often treated as DMs, the list suggests a widespread grammaticalization of DMs from the lexemes meaning 'right, real, true, correct'.

Many Asian states have been influenced by China for millennia, due to China's 'cultural prestige or attractivity' (Bisang 2006: 89), and this influence is often clearly visible in language, which is particularly true with Korean. For instance, even though the exact time is unknown, Koreans used Chinese characters or their modified versions to suit Korean, taking the character's meaning (semantogram) or sound (phonogram), prior to the invention of the Korean script *Hangeul* (= *Hangul, Hankul*) in 1443. Identifiably Korean texts written in Chinese characters date back to the 5th century CE (Whitman 2015). Chinese characters are no longer in common use in Korea, but the Korean lexicon has a large proportion of Sino-Korean words. In *Phyocwun Kwuke Taysacen* (1999), an authoritative dictionary containing 364,000 head words, 53.2 per cent of the entry is of Chinese origin, mostly nouns. Some of these Sino-Korean words grammaticalized into DMs after they were borrowed into Korean. The morpheme *cin* 'tru(th)' is one such borrowed word, from which the DM *cincca* was formed in Modern Korean.

Despite its recency in development, the Korean DM *cincca* presents an interesting grammaticalization scenario. It is a hybrid form consisting of the Sino-Korean bound morpheme *cin* 'truth' and the nominalizer *cca* 'thing, person', thus denoting 'a true thing' (see 4.1 for more). A noun phrase in form, *cincca,* from the early days of its appearance, carries an adjectival function of adding genuineness or excellence in quality to a modified noun (e.g., *cincca yayki* 'a true story', *cincca pomwul* 'a genuine treasure') or of an adverbial function of adding emphasis to an adjective or a predicate (e.g., *cincca aphu-* 'be truly painful', *cincca hayngwuni-* 'be truly fortunate'). In Modern Korean *cincca* displays a wide range of discourse functions. The DM *cincca* has been studied by Kim and Jeong (2012) and Rhee (2021), but since it was analyzed along with other DMs involving various 'truth'-related DMs, the analyses were necessarily brief.

Chinese has two DMs that developed from the same etymon *zhen* (眞) (note the notational variation in *zhen* and *cin*), i.e., *zhende* and *zhenshi*. These DMs have received attention in a number of studies, e.g., Li (2003), Yan (2006), Chen (2010), Fang (2012), Han (2016), Liu (2021) for *zhende* and Zhang (2004), Dong (2005), Wang (2010), Li (2011), Guo (2015), Li (2015) for *zhenshi*. These studies have addressed each DM largely with respect to syntactic and semantic features, not involving comparative or crosslinguistic perspectives. A historical investigation shows that the development of these two DMs is rather recent, i.e., from around the 20th century, and that their functions have substantial commonalities as well as

relatively minor differences with the Korean counterpart (henceforth all these DMs are collectively called 'the TRUTH-DMs'). This state of affairs presents theoretically significant implications in grammaticalization studies but there has been no research comparing and contrasting these DMs to date. Given this background, this research intends to fill the gap. It aims to compare them and answer the following research questions:

(i)   how did the truth-DMs in Korean and Chinese develop?

(ii)   what are the similarities and differences between their functions in the two languages?

(iii) what are the major changes that occurred in the development of these DMs?

The organization of the paper is as follows. Section 2 briefly presents a theoretical background; section 3 explains research methods; section 4 presents discourse functions of Korean DM *cincca* (in 4.1) and Chinese DMs *zhende* and *zhenshi* (in 4.2 and 4.3); section 5 discusses a few issues of theoretical importance, i.e., notable changes (in 5.1), and commonalities and differences (in 5.2); and section 6 summarizes the major findings and concludes the paper.

2. Theoretical background

Languages come into contact with other languages for diverse reasons, and such contact often leads to change. Grammaticalization is a ubiquitous process in the scenes of language contact, which may affect any part of language structure (Heine & Kuteva 2011: 291). In recent linguistic research traditions, the role of language contact in grammaticalization, and, more generally, language change, has attracted the attention of researchers, notably Heine and Kuteva (2005), Adamou and Matras (2020) and Grant (2020). Notable contributions of these studies include identifying the influence of the donor language with respect to the borrowed lexemes' original semantics, the language's typological profile, and the limits of the transfer, i.e., borrowability. Furthermore, by comparing the donor and recipient language situations, a better understanding about the contact-induced and evolutive (i.e., language-internal) grammaticalization has been made possible.

3. Data and research methods

DMs arise from an interaction, either real or imagined, in discourse and their functions are highly variable depending on the context they occur in. Therefore, authentic discourse data are of crucial importance in DM research. Also, it is important to have historical data in order to trace the change of the DMs and their earlier forms. For this reason, Korean data are taken from various historical and contemporary resources, including the Sejong Historical Corpus for diachronic investigation, and the Drama Corpus for investigation of contemporary usage. The Sejong Historical Corpus is a 15-million-word, historical section of the Sejong Corpus, developed as part of the government-led 21st Century Sejong Project

(1998–2006). The texts in the historical section date from 1446 through 1913. The Drama Corpus is a 24-million-word contemporary corpus, a collection of 7,454 scenarios of dramas and movies dating from 1992 through 2015, compiled by Min Li while she was at Seoul National University (currently at Tsinghua University). UNICONC, a concordance program developed by Jinho Park, was used as a data search engine. The data for analyses of Chinese DMs is taken from the CCL (The Center for Chinese Linguistics at Peking University) online corpus, consisting of 581,794,456 characters, without word segmentation or POS tagging.[1] The data obtained from corpus search are hand screened for identification of their grammatical status and function.

4. Results

4.1. Korean *cincca*

Korean has a number of TRUTH-related adverbs involving the Sino-Korean morpheme *cin* 'tru(th), e.g., *cinsillo, cintis* (obsolete), *cinceng, cincengulo, cincengkho*, and *cincca*. Of these, only *cincca* functions as a full-fledged TRUTH-DM in contemporary Korean, in view of syntactic behavior and discourse functions, whereas others either carry lexical meanings or are of marginal distribution (cf. Kim & Jeong 2012, Rhee 2021). The meaning of *cin* has been extended over time to a great extent, e.g., '(religious) truth' (15th c.), 'substance, genuineness, reality (counter-expectation), factuality, extraordinariness' (15th c.), 'sincerity' (16th c.), 'depth, strength, truthfulness, true identity' (19th – 20th c.), etc. (Rhee 2021: 468–469).

The word *cincca* first occurs as *cinscA* in a Korean-English bilingual dictionary *Hanyengcacen* (1897), but dictionaries (e.g., *Wulimalsaym* and *Koesacen*) indicate that it is a modern form of the earlier forms, *cintis* (15th c. –19th c.), *cincis* (15th c. – 20th c.), *cincit* (18th c.).[2] It is not clear, however, how the modern *cincca* and the earlier forms *cintis/cincis/cincit* are related, especially phonologically. The modern form, *cincca*, is likely to have developed from a Sino-Korean morphological derivation of *cin* 'true' and *ca* 'thing', of which the last consonant /c/ became tensed to /cc/ for emphasis (see Koo 2009 for discussion of force-dynamic representation of tensed and aspirated sounds).[3]

---

[1] Special thanks go to the developers of these corpora and search programs for generously granting their use for our research.

[2] Two Chinese-Korean historical dictionaries (*Yekeyuhaypo* 1775 & *Pangenyusek* 1778) show that Chinese *zhen* and *zhende* are equal to *cincis* in Korean. Even though *cincca* is first attested in 1923, the DM function is first observed in a 1964 drama, *Culkewun wulicip* [Our sweet home].

[3] If *cca* is indeed related to the Sino-Korean morpheme *ca* (子; 'son, thing, or a NOMINALIZER' or 者 'person, thing, or a NOMINALIZER'), it is a fully Koreanized morpheme in contemporary Korean, as found in *kacca* 'a fake thing', *chocca* 'a novice', *kongcca* 'something free of charge', *sayngcca* 'uncooked food; an unexperienced person', and many other similarly formed words. *Hanyengcacen* (1897) indicates *zhende* (眞的) as its Chinese counterpart, with the meaning in English 'what is real, genuine quality, a sincere person'.

The predecessor forms *cintis/cincis/cincit* had adjectival, adverbial, and nominal functions. They had such meanings as 'pure' (e.g., pure gold); 'truthful' (e.g., truthful principle); 'genuine' (e.g., genuine sap), in the 15th c. data; 'truly' (e.g., truly resembling an army general) in the 17th c. data; and 'very' (e.g., very rich), 'truth' (e.g., truth and falsity) in the 18th c. data. The modern form *cincca* is a noun phrase in form but from the early days of its appearance, it carried an adjectival function of adding genuineness or excellence in quality to a modified noun (e.g., 'a true story', 'a genuine treasure'), or an adverbial function of adding emphasis to an adjective or a predicate (e.g., 'truly fortunate', 'truly painful'). The modern form, *cincca*, inherited the semantic features from its predecessors, and from the emphatic adverbial function there arose diverse DM functions through the interaction of the source meaning of truthfulness and diverse inferences from the discourse contexts.

4.1.1. Emphasis (*cincca*)

Among the primary, and presumably first, DM function is the emphatic function, paraphrasable as 'I am serious/What I say is true!', modifying a proposition, whereby the propositional content can carry stronger, emphatic force. The emergence of this function from a lexeme denoting 'true' is conceptually straightforward. This usage is exemplified in the following excerpt:[4]

(1) [Hearing knocks on the 2nd floor window, C is scared and tries to stay in the arms of his father. His older siblings A and B look at him pathetically.]
A: [You are in your father's arms because you don't want to study (in your room).]
B: [No doubt. I, too, used to say I had stomachache to avoid studying when I was young.]
C: 아니라니깐요! 진짜 누가 창문을 두드렸다구요!
ani-*lanikkan-yo!* ***cincca*** *nwuka*　　*changmwun-ul*　*twutuly-ess-takwu-yo!*
be.not-END -POL DM　　someone:NOM　window-ACC　knock-PST-END-POL
'No! **DM (= I'm serious!)** Someone knocked at the window (and that's why I am so scared!)'

(2008, Sitcom *Khokkili* Episode #116)

One interesting aspect of this development is that the 'truth' is conceptualized as opposed to non-truth in (deceptive) appearance, whereby the addressee is not yet cognizant of the 'truth' and thus the speaker is about to reveal. In other words, in a situation as in (1), the speaker is claiming that the truth (that someone knocked on the 2nd floor window) is in opposition to the appearance (that such an event is impossible). In this line of thought, the 'truth' is often considered as hidden or at least not easily accessible, whose manifestation is considered extraordinary. In the development of the emphasis function, it can be said that the intensifying a property

---

[4] In examples, the lines that do not require morphemic glosses are given in English translation within square brackets and supplementary information not present in the original text but is helpful for understanding the meaning is given in parenthesis in translation.

of an entity (as a quality/degree modifying adjective or adverb) in the subjective world has carried over to emphasizing the value of a proposition in the interactional/intersubjective world. From a crosslinguistic perspective, the development of a 'truth' lexeme into a pause-filling DM is analogous to the development of the 'certainty' lexemes (e.g. *actually, really, no doubt, indeed,* etc.) into the DMs for pause-filling, as well as uncertainty and mitigation (Aijmer 1986, Čapková 2015).

4.1.2. Discontent (*cincca*)

The DM *cincca* has further acquired the emotive function of marking discontent, frustration or even annoyance. This function can be paraphrased as 'I am unhappy with the situation' or 'It's so frustrating'. The DM used in this function typically occurs stand-alone and often precedes or follows ellipsis. This seems to be closely related to the emphatic function elaborated above since such an utterance is often necessitated by someone or a situation not compatible with the speaker's 'true' intention or desire. However, the DM in this function tends to be not other-directed and thus to occur typically in monologues, i.e., the speaker contains the discontented feelings to themselves. This function is illustrated in the following:

(2) Frustration *cincca* (Modern Korean)
[The speaker runs after his girlfriend, who saw him with a woman. His girlfriend gets in a taxi and goes away before he can stop her.]
A: 선주씨이이   아 진짜...!
    sencwu-ssiii   (taxi drives away; A watches it blankly)   *a*   *cincca...!*
    [name]-HON   INTJ   DM...
    'Sunjoo~! Alas, **DM (= it's so frustrating!).**'
                              (2007, Drama *Talcauy pom* Episode #8)

The relationship between 'truth' and discontent is rather subtle. The claimed truth is that what the speaker says or feels is 'truthful'.

4.1.3. Challenge (*cincca*)

Closely related to the discontent function exemplified above is that of marking challenge, paraphrasable as 'I am challenging you', also often occurring in elliptical contexts. Obviously, this is a stronger form of negative emotion than discontent, and its major difference from discontent is that it is other-directed, and thus interactional. As the DM is directed toward the addressee, it is typically spoken in a louder voice than the discontent DM. This function is illustrated by the following excerpt:

(3) Challenge *cincca* (Modern Korean)
    (A and B are traveling companions without mutual acquaintance. A says he is indebted to B, because B helped him get released from the prison. A wants to be of B's help in return, but B is a royal envoy incognito, who does not want A to follow him.)

A: [Even beasts know how to pay back for the favor. They pluck their own hair to make warm clothes for their owner.]

B: [(plucking hair from his nostril,) Here's one. Make a fur jacket with it.]

A: 이 양반이 진짜 보자 보자 하니까...

(in fury) *i* *yangpan-i* **cincca** *po-ca* *po-ca* *ha-nikka...*
        this fellow-NOM DM see-HORT see-HORT say-CONN

'(Look at) this fellow **DM (= I'm challenging you!)** I've been patiently swallowing insult but...' (Lit. 'This fellow, while (I've) been saying to myself, 'let's wait, let's wait',...')

B: 뭐? 이 양반? 이게 죽을라구 진짜...

(in fury) *mwe?* *i* *yangpan?* *ike-y* *cwuk-ullakwu* **cincca**...
        what? this fellow? this.thing-NOM die-PURP DM

'What? This fellow? (Look at) this damn fella. (He's) begging to die, **DM (= I'm challenging you!).**'

(2011, Movie *Cosen myengthamceng*)

In (3) the DM *cincca* in the challenge function occurs twice. It signals that their offended feeling or their intention of performing a punitive action is 'true' and strong.

4.1.4. Surprise (*cincca*)

The DM *cincca* has also acquired a neutral or positive function as an interjection-like marker of surprise, paraphrasable as 'I am surprised'. Unlike the negative functions of discontent and challenge, shown in 4.1.2 and 4.1.3 above, the surprise marked by *cincca* can be neutral, positive, or negative in nature. The positive and negative surprise is exemplified in the following examples:

(4) a. Positive surprise *cincca* (Modern Korean)

(The speaker is impressed by the luxury of the house in which his friend used to live.)

[Is this the house you used to live?]

와~ 너 진짜 좋은 데 살았구나.

wa~ *ne* **cincca** *coh-un* *tey* *sal-ass-kwuna*
INTJ you DM be.good-ADN place live-PST-EXCL

'Wow... you **DM (= what a surprise!)** lived in a wonderful place!'

(2010, Drama *Kutay, wuseyo* Episode #6)

b. Negative surprise *cincca* (Modern Korean)

(A, while interrogating a witness-suspect, coaxes him to confess, promising that he will beprotected by the witness protection laws and that his charges would be exonerated, a promise that he cannot guarantee. A's supervisor (B) stops him and reprimands him for his reckless promises.)

B: [(How can you act so arbitrarily?) A witness protection might be alright, but do you think exoneration is what you can do?]

A: [What's important is to make him talk no matter what. Why should I keep my promises?]

B: (dumbstruck)   강민기... 너 진짜...
kangminki...   *ne*   *cincca...*
[name]   you   DM
'Kang Minki... you **DM (= How surprising! How can you say that?)**'
<div align="right">(2007, Drama *Kaywa nuktayuy sikan* Episode #10)</div>

The DM carrying the interjective surprise function also tends to occur in elliptical structures, often involving elongated trailing of the last syllable. The ellipsis strongly suggests that the speaker is emotionally overwhelmed and is unable to complete the utterance.

4.1.5. Sudden remembrance (*cincca*)

Still another function is to mark sudden remembrance or realization, which can be paraphrased as 'Oh, I almost forgot' or 'Oh, that reminds me'. The DM in this function occurs at the left-periphery, and thus prefaces a statement or question that has been nearly forgotten. This is exemplified in the following example:

(5)   Sudden remembrance *cincca* (Modern Korean)
      A: [Why are you so down? What's the matter?]
      B: [No, nothing.]
      A: [Tell me. What is it? You look so down. Did he cause trouble?]
      B: 아니에요. 진짜... 저기요
         ani-ey-yo.   *cincca...*   *ceki-yo*
         be.not-END-POL   DM   DM-POL
         'No. **DM (= Oh, that reminds me)**... well..'
      A: [Yes. What is it?]
      B: [If a man does not come when a woman's family invites him, it means that he is not serious about the woman, right?]
<div align="right">(2016, Drama *1%-uy etten kes* Episode #8)</div>

The surprise function elaborated above (in 4.1.4) is closely related to this sudden remembrance function, because remembering or realizing forgotten 'truth' is typically instantaneous and often leads to surprise. For this reason, DMs used for these two functions resemble interjections.

4.1.6. Pause-filling (*cincca*)

The last function to be elaborated is that of filling a pause. Pause-filling, in general, may be necessitated by the speaker's need for word search to earn time for formulation of an utterance, which the speaker may either succeed or fail, or by the speaker's desire to signal that the situation is such that any suitable utterance cannot be found. Thus, the DM *cincca* carrying the pause-filling function can be paraphrased as 'I am not yet finished', or 'I am speechless' or rhetorically 'What can I say?!' The acquisition of the filler function suggests that the 'truth' meaning originally associated with the source lexeme has become significantly bleached, with no discernible semantic vestige in the form. This state of affairs is similar to

the English TRUTH-DM *actually* carrying the filler function (cf. Aijmer 1986). This function of *cincca* is illustrated in the following:

(6)   Pause-filling *cincca* (Modern Korean)
      a. (A learns that his girlfriend's (B's) father remarried. Now in reference to B's younger stepsister he met the other day:)
      B: [So, I live with my stepmother's daughter, and the girl you met the other day is...]
      A: [Aha, I got it. She is the collaborative production of my would-be father-in-law and his new wife!]
      B: 아... 진짜...
         (siling)   *aa...*   ***cincca...***
                      INTJ   DM
         'Oh, **DM (= Well, I don't know what to say...?!)**'
      A: 진짜 뭐요?
         cincca   *mwe-yo?*
         real   what-POL
         'Real what?'
                              (2009, Drama *Chenmanpen salanghay* Episode #22)

      b. (Mother (A) is trying to win sympathy from her daughter (B) and son (C) by refusing to eat and feigning to be sick in order to get money from them. B and C know that A is not really sick.)
      C: [Mom... (handing A an envelope with money).. Mom... this much is all I can do for you.]
      A: (Her eyes twinkling all of a sudden) [uh..?]
      C: [My wife does not know this. Just keep that in mind.]
      A: (pretending to be struggling to sit up) [I understand... thank you... I will surely pay it back to you...]
      C: [Now come out of the room ... Eat with us... (and exits the room)]
      A: [... OK... I should try to eat for you guys even if I've lost appetite...]
      B: ...아우 우리 엄마 진짜...
         ... *awu*   *wuli*   *emma*   ***cincca...***
                INTJ   our   mother   DM
         'Oh! Our mother (is) **DM (= what can I say?!)**'
      A: [Me? Me, what?]
                              (2011, Drama *Naysalang naykyethey* Episode #3)

   Note that in (6a) B uses *cincca* as a DM with vacuous semantic content and A (tauntingly) copies the utterance and uses it as a lexical modifier, a commonly observed pun based on lexical and grammatical meanings of *cincca*. In (6b), A (the mother) knows that *cincca* used by B (her daughter) was a pause-filler indicating that she was speechless because of her senseless, childish behavior, but she twists it by pretending that her daughter was looking for words to describe her mother and that she wants to know what they were.

4.2. Chinese *zhende*

The modern Chinese lexeme *zhen* (眞, simplified 真)'tru(th)' is of a long history, found as early as in Archaic Chinese (also known as Old Chinese, 8th–3rd centuries BCE) (see Long et al. 2012). The word is found in a number of TRUTH-adverbs, e.g., *zhen* 'tru(th)', *zhende* 'true PTCL, truth of', *zhenshi* 'tru(ly) so', *zhenshide* 'tru(ly) so PTCL, true correctness of', *zhenxin* 'sincere heart', *zhenxinde* 'sincere heart PTCL, sincere heart of', *zhenzheng* 'sincerity', etc., in which *de* is a nominalizer.[5] Of these, *zhende* and *zhenshi* (and its variant *zhenshide*) function as DMs.

The 'tru(th)' meaning seems to have been extended to various semantic fields and notably to religious idea of truth or truthfulness. For instance, in one of the oldest dictionaries, *Shuowen Jiezi* of the 2nd century, the meaning of *zhen* is given with the following explanation, strongly suggesting that *zhen* was related to religious idea of 'true person', 'fairy', 'un-worldly man', 'hermit', etc.

(7)　(眞) 僊人變形而登天也。此眞之本義也。
　　(Zhēn)　xiān　rén　biànxíng　ér　dēngtiān　　　　yě.
　　(*zhen*)　true　person　transform　and　ascend.to.heaven　PTCL
　　Cǐ　　zhēn　zhī　běnyì　　　　　　yě.
　　this　*zhen*　of　original.meaning　PTCL
　　'(*Zhen*) True persons (= immortals) transform into heavens. This is the original meaning of ＇zhen＇.'

Despite the long history of *zhen*, the occurrence of *zhen-de*, involving *zhen* 'tru(th)', is very rare in old texts: before the Yuan dynasty (1271–1368), it occurs only five times in the CCL corpus, with only two of them clearly meaning 'true', 'genuine'. During the Ming dynasty (1368–1644) and the Qing dynasty (1644–1912), *zhen-de* occurs as a well-established syntactic construction, meaning 'true', 'genuine', 'real'. It is only in the 19th century that *zhende* begins to occur as a univerbated adverb denoting 'in reality', 'truly'. *Zhende* thus diverges into the original lexical use as a noun and the grammaticalized use as a DM. In the latter use, it is now not part of an argument (i.e., of a noun phrase) or a predicate (i.e., of a verb), and thus it is omissible. The following excerpt from the 19th century exemplifies the adverbial use of *zhende*.

(8)　Adverbial *zhende* 'truly' (19c)
　　[My younger brother aspires to study benevolence, righteousness, morality, and self-cultivation. I was very pleased when I read his letter...]
　　弟弟真的有此志向，要熟读《小学》及《五种遗规》两书。
　　dìdì　　　　　zhēnde　yǒu　cǐ　zhìxiàng,　yào　　shúdú
　　younger.brother　truly　　have　this　aspiration　need　　peruse

---

[5] *De* in Modern Chinese is a multifunctional morpheme, e.g., markers of genitive ('of'), relativization ('that/which'), nominalization, affirmativity, among others.

"xiǎoxué"	*jí*	*"wǔzhǒng yíguī"*	*liǎng*	*shū*
[book.title]	and	[book.title]	two	book

'If he **truly** has such aspirations, he should diligently study the books *Xiaoxue* and *Five Kinds of Yigui*.'

(19c. *Zeng Guofan's family letter*)

The development into an adverb, notably a sentential adverb, paves a way to the development of DMs (Traugott & Dasher 2002). This is evident in that many DMs across languages carry the dual function of adverbs and DMs, and that DMs, by virtue of not constituting a word class in their own right in traditional grammar, are typically categorized under adverbs. From the adverbial stage, *zhende* acquires diverse discursive functions.

4.2.1. Emphasis (*zhende*)

The DM use of *zhende* is first attested in the 20th century, often standing alone by gaining positional freedom. Paraphrasable as 'I'm serious/That's true!', the DM *zhende* emphasizes the content of the foregoing statement, i.e., it is an anaphoric emphasizer. This usage is illustrated in the following excerpts taken from Modern Chinese.

(9) Emphasis *zhende* (Modern Chinese)
  a. [Li Qin, let me first express my gratitude for your criticism.]
    真的！我从报上见到你的文章后，剪了下来，　　　夹于一本书中，予以保留。
    Lǐqín, ràng wǒ yě shǒuxiān duì nǐ de pīpíng biǎoshì gǎnxiè.

Zhēnde!	*Wǒ*	*cóng*	*bào-shàng*	*jiàn-dào*	*nǐ-de*	*wénzhāng*	*hòu,*
DM	I	from	newspaper-on	see-RC	you-GEN	article	after

jiǎn-le	*xiàlái,*	*jiā-yú*	*yì.běn*	*shū*	*zhōng,*	*yǔ*	*yǐ*	*bǎoliú.*
cut-PERF	down	place-in	one-CLSF	book	in	give	to	keep

    '**DM** (= **Really!**) After I saw your article in the newspaper, I cut it out and kept it in a book for safekeeping.'

(1996, *Writer digest*, Newspapers)

  b. 她马上回答：“我不会介意的，真的。”

Tā	*mǎshàng*	*huídá:*	*"Wǒ búhuì*	*jièyì-de*	***zhēnde.''***
she	immediately	answer	I will.not	mind-ptcl	DM

    'She immediately responded, "I won't mind, **DM (= really).**"'

(1995, Novel, *Hóng Píngguǒ, Love of the Red Apple* by Yu Qing.)

  c. Yang Lan: [But actually, when a person reaches their forties, they start to think about what they really want in life. In the past two years, have you experienced any changes in your mindset?]
    刘德华：　　变化了很多年，真的。需要一个自己的家。

Liúdéhuá:	*Biànhuà-le*	*hěnduō*	*nián,*	***zhēnde.***	*Xūyào*	*yí-gè*	*zìjǐ-de*	*jiā.*
	change-PERF	very.many	year	DM	need	one-CLSF	own-GEN	home

    Andy Lau: 'There have been many changes over the years, **DM (= really).** I realized the need for a family of my own.'

(21c. *Yang Lan's Conversation with Prominent Figures: Yang Lan's Interview* Record II, Informal Dialogue)

4.2.2. Preface to noteworthy information (*zhende*)

The next function, preface to noteworthy information, is similar to emphasis, in that both highlights the importance of the information in the discourse. As the name suggests, however, the function of preface to noteworthy information makes reference to what is to follow, i.e., cataphoric, in contrast with the emphasis function with the anaphoric reference. Thus, the DM in this function can be paraphrased as 'Let me tell you this', 'Listen up', 'This is important', etc. This function is exemplified in the following excerpts.

(10) Preface to noteworthy information *zhende* (Modern Chinese)
  a. [Liu Zhaohua shook the empty cup in his hand. At that moment, Ah Guang and the others were busy taking notes, and no one noticed that he needed water. I got up to pour him some water, and he thanked me.]
  然后他不忘推销他的方剂说: "真的，你不信，我在桂林的研制的保健品都是天然的中药配方，倘若你服用之后，你这个人就像是换了一个人。

  Ránhòu tā búwàng tuīxiāo tāde fāngjì shuō:
  *"**Zhēnde**, nǐ búxìn, wǒ zài guìlín-de yánzhì-de bǎojiànpǐn*
  DM   you not.believe I in [place.name]-GEN develop-ADN health.product
  dōu shì tiānrán-de zhōngyào   pèifāng, tǎngruò nǐ fúyòng zhīhòu,
  all be   natural-ADN Chinese.medicine formula if you take after
  nǐ zhège rén   jiùxiàng shì huàn-le   yí-gèrén."
  you this person just.like be change-PERF one-CLSF person
  'Then, he didn't forget to promote his herbal medicine, saying, "**Really (Let me tell you this)**, it may be hard for you to believe me, but the health products I've developed in Guilin are all made from natural traditional Chinese medicine formulas. Once you take them, you'll feel like a completely different person."'
                          (20c. *Female Journalist's Face-to-Face Conversation with Notorious Drug Lord Liu Zhaohua,* Informal Dialogue)

  b. [The conversation between Song Ailing and Kong Xiangxi did not know where it started, and gradually focused on their views on money. When it comes to money, both of them are very excited, and their faces are red with wine and blood.]
  宋蔼龄先说道: "真的，我从小就有这种感觉，钱这个东西非常奇怪，非常神秘 ……"
  *Sòng'ǎilíng xiān shuōdao: "**Zhēnde**, wǒ cóngxiǎo jiù yǒu zhèzhǒng ǎnjué,*
  [name]   first say   DM   I from.young just have this.kind feeling
  qián zhège dōngxī fēicháng qíguài, fēicháng shénmì…"
  money this thing very   strange very   mysterious
  'Song Ailing first said: "**DM (= Really)**, I have had this feeling since I was a child. Money is very strange and mysterious..."'
                          (1998, *Complete Biography of the Song Family* by Cheng Guang and Ye Si.)

In the excerpts above, the speaker uses the DM *zhende* to signal that some important information is about to be revealed. To ensure the addressee's attention, the DM is usually followed by a pause.

4.2.3. Surprise (*zhende*)

The next function is to mark the speaker's surprise, and the DM can be paraphrased as 'I'm surprised!' Surprise signaled by *zhende* is value-neutral, thus for both good and bad in nature.[6] This function is exemplified in the following excerpt:

(11)   Positive surprise *zhende* (Modern Chinese)
       (Song has been distressed because her husband, President of the Republic of China, Chiang Kai-shek, has been very ill, which made her worry that he would die soon. Now, the doctor says he will not die so soon.)
       ["Madam, please don't worry. The President's immune system can last at least six months without any issues," the doctor reassured Soong Mei-ling.]
       "真的！如果是这样，我太谢谢你了。"

       *"**Zhēnde!**   Rúguǒ shì   zhèyàng,   wǒ   tài   xièxiè   nǐ   le."*
       DM   if   be   this   I   great   thank   you   PTCL
       "**Really!** If that's the case, I am so grateful to you."'
                                                     (1998, *Complete Biography of the Song Family*
                                                      by Cheng Guang and Ye Si.)

In the excerpt, the speaker, happily surprised by the relieving remarks by the doctor, uses the DM *zhende* at the receipt of the information. This function is similar to the Korean DM *cincca* with the function of marking positive surprise in (4a).

(12)   Negative surprise *zhende* (Modern Chinese)
       "真的，无语了，这些人的脑子得坏成什么样才能搞出这么个玩意儿来啊！"

       *"**Zhēnde,**   wúyǔ   le,   zhèxiē   rén-de   nǎozi   děi   huài-chéng*
       DM   no.word   PTCL   these   person-GEN   brain   have   bad-RC
       shénme   yàng   cái-néng   gǎo-chū   zhème-gè   wányìr   lái   a!"
       what   form   then-can   bring-RC   such-CLSF   stuff   come   PTCL
       'DM (= Really), I'm speechless. How messed up do these people's brains have to be to come up with something like this!'
                                                     (2010s\201x\Internet corpus\Weibo\weibo_004)

In (12), similar to (11) above, the DM *zhende* signals surprise of the speaker, with the notable difference that the signaled surprise is negative in nature.

4.2.4. Preface to upcoming disalignment (*zhende*)

The last functional category of *zhende* is to preface an upcoming disalignment, paraphrasable as 'Well, but...', which is often used with *danshi* or *keshi* (both meaning 'but'). In other words, the DM signals that the speaker has information that does not align with the interlocutor's prior statement or expectation. In its

---

[6] As Haiping Long (p.c.) points out, *zhende* in this function tends to occur with a rising tone. As is the case with the Thai 'truth'-based DM (Khammee 2024), which utilizes the appeal contour for certain functions, tonal or suprasegmental differentiations may be associated with particular discourse functions.

discourse function and its lexical semantics, this DM is similar to the English DM *actually*, which carries the function of prefacing upcoming disalignment. One major difference is that *actually* itself marks the shift (thus, does not usually occur with *but*) but the Chinese *zhende* is normally accompanied by the adversative connective 'but'. This usage is exemplified below:

(13) Preface to upcoming disalignment *zhende* (Modern Chinese)
["Well," said Jessop, "now you understand the method of suicide I propose to you. I propose that Mrs. Betterton should go on traveling. And you should be Mrs. Betterton."]
"但是，真的，" 希拉里说，"那将很难做到。..."
*"Dànshì, **zhēnde,"**   xīlālǐ shuō, "nà   jiāng hěnnán   zuò-dào...*
 but   DM   [name] say   that   will   very.hard do-RC
"But, **DM (= actually; well, but),"** Hilary said, "that's going to be hard to do. [I mean, they're going to recognize right away that I'm not Mrs. Betterton.]"

(c. 1998, Novel *Mùdìdì Bùmíng*,
trans. of Agatha Christie's *Destination Unknown*)

In the excerpt above, the speaker uses the DM *zhende* to signal that some disalignment (i.e., the speaker's reluctance to accept the proposal) is to follow.

4.3. Chinese *zhenshi*

The Chinese expression *zhen-shi*, literally 'be truly/correctly so' (*zhen* 'true' + *shi* 'be so'), denotes truthfulness of the described entity, either a choice of expression or the content of a statement. In the historical record, the expression first occurs from the Warring States period (475–221 BCE), as a syntactic construction, not univerbated, thus having a compositional meaning of 'be truly (so)'. It continues to be used as a syntactic construction until the Northern Song dynasty (960–1127). In the Ming dynasty (1368–1644), the internal cohesion increases and the syntactic construction (*zhen-shi*) becomes a univerbated word (*zhenshi*) carrying the adverbial function. In this word, the semantic focus is on *zhen* while *shi* is peripheral (thus, departing from the predicative use ('(X) is truly so')). In this period, however, the adverbial usuage of *zhenshi* 'truly' is numerically insignificant. One such example is shown in the following excerpts:

(14) Adverbial *zhenshi* (Ming dynasty)
[Just now, the mother's worries turned into joy, and she smiled brightly, saying,]
"亏得儿子峥嵘有日，奋发有时，真是谢天谢地！"
*"Kuīdé érzi   zhēngróng-yǒu-rì,   fènfā-yǒu-shí,   **zhēnshi** xiètiānxièdì!*
fortunately son prominent-have-day strive-have-time truly thank.heaven.thank.earth
' "Fortunately, son is extraordinary, and sometimes he works hard; **truly**, thank God."'

(Ming dynasty, 17c. *Jīn Gǔ Qíguān*,
Miracles of the Past and Present by Bao Weng Lao Ren)

At a later time, in the Qing dynasty (1636–1912), the adverbial usage increases. As was the case with *zhende* in the previous section, *zhenshi* also begins to develop into a DM. With the persisting semantics of the source, its primary function is to affirm what is said (e.g., 'That is true and correct'), but in the 20th century, it has acquired diverse discourse functions with the influence of the context and conceptual extension.

4.3.1. Discontent (*zhenshi*)

The first, and most frequent function of the DM *zhenshi* is to signal negative stance of the speaker, such as discontented emotion or critical attitude, paraphrasable as 'I'm not happy with you/that...'. The discontent is generally a mild one, typically with respect to the situation, or someone indicated in the discourse, or the interlocutor, but it can be a stronger as in blaming someone for misconduct. The negative attitude signaled by *zhenshi*, however, is not as strong as challenge indicated by the Korean DM *cincca* (see 4.1.3 above). What the speaker is unhappy with may be expressed before or after the DM, as shown in the following excerpts, taken from Early Modern Chinese and Modern Chinese data:[7]

(15) Discontent *zhenshi* (Early Modern Chinese and Modern Chinese)
a. 这个人一想，再要是孩子凑多了，更不好办了。真是！

Zhège rén yì-xiǎng, zài yàoshi háizi còu duō le,
this person once-think again if child gather more PTCL
gèng bùhǎo bàn le. Zhēnshi!
further not.good handle PTCL DM
'This person thought that if there were even more children, it would be even more difficult to handle. **DM (= Indeed**!)'

(Qing dynasty, 1890, *Xiǎo Wǔ Yì, The Five Little Moralities*)

b. 秀芬急得说：＂真是！凤姐怎么还不来？＂

Xiùfēn jídé shuō: "Zhēnshi! Fèngjiě zěnme háibù lái?"
[name] anxiously say DM [name] how still.not come
'Xiu Fen anxiously exclaimed, **"DM (= Indeed**!) Why hasn't Feng Jie arrived yet?"'

(Modern Chinese, 1964, *Zhàndòu de Qīngchūn, Fighting Youth*)

c. [Anyway, it's not something honorable or prestigious,]
何必问来问去，自寻烦恼？真是！

hébì wènlái wènqù, zì-xún fánnǎo? Zhēnshi!
why ask.come ask.go self-seek trouble DM
'Why bother asking around and causing unnecessary trouble? **DM (= Indeed**!)'

(1947, *Wéi Chéng, Siege* by Qian Zhongshu)

---

[7] According to the generally accepted periodization of the Chinese language, the Qing dynasty times (1636−1912) belong to Early Modern Chinese (1271−1912), encompassing the Yuan, Ming and Qing dynasties for over six centuries. The language of the Early Modern Chinese records dating from the late Qing dynasty times, especially those written around the turn of the 20th century, are similar to that of Modern Chinese.

d. A: [Is it the Li family again?]

B: "那还不是。那个男人啊，真是！丢了事，天天喝酒，越没有钱越喝酒。"

*Nà hái búshì. Nàgè nánrén a, **zhēnshi**!*
that still not that man ah DM
Diū-le shì, tiāntiān hējiǔ, yuè méi-yǒu qián yuè hējiǔ.
lose-PERF job everyday drink.liquor more not-have money more drink.liquor
'That's right. That man, **DM (= really; how dreadful he is)**! He never goes to work but drinks every day. The less money he has, the more he drinks.'

(Modern Chinese, 20th c. *Bēi Gē, Elegy* by Zeng Zhuo)

In the examples above, *zhenshi* signals that the speaker (in (15a), the person being quoted, not the narrator) is unhappy with a hypothetical situation of having too many children to take care of; in (15b) the speaker is unhappy with the tardy person Feng Jie or with the delay being caused by him; in (15c) the speaker is annoyed by the repeated occurrence of an undesirable event or action; and in (15d) the speaker is expressing his disapproval toward the man with a blameworthy lifestyle. It is interesting to note that early instances of *zhenshi* nearly always occur as an affirmation/confirmation ('That's right') of something negative ('That's right – it's bad'), unlike its formal relative *zhende*, which does not show such propensity in evaluative polarity.

4.3.2 Sudden remembrance (*zhenshi*)

The next function of *zhenshi* is to signal sudden remembrance or realization, paraphrasable as 'Oh, I forgot', 'Oh, that reminds me!' Since people utter this DM normally spontaneously at the moment of remembrance, it is said before the content of the remembrance is presented. This is illustrated in the following excerpt[8]:

(16) Sudden remembrance *zhenshi* (Early Modern Chinese and Modern Chinese)
['Zheng Xiong said, "Brother Yang, what do you three usually do for a living?" Yang Ming replied, "We work as bodyguards outside. May I know your honorable surname?"]

马俊说："真是，我也忘了，这是我拜兄，他姓郑名雄，名号人称铁面天王。"

*Mǎjùn shuō: "**Zhēnshi**, wǒ yě wàng-le, zhè shì wǒ bàixiōng,*
[name] say DM I also forget-PERF this be I sworn.brother
tā xìng zhèng míng xióng, mínghào
he surname [name] first.name [name] name.and.alias
rénchēng tiěmiàn tiānwáng."
person.call iron.face heavenly.king
'Ma Jun said, "**Indeed (= oh, I forgot!)**, I forgot to introduce. This is my sworn brother. His surname is Zheng, and his name is Xiong. He is widely known as the Iron-faced Heavenly King."'

(Qing dynasty, *Jìgōng Quánzhuán, The Complete Biography of Jigong*)

---

[8] The exact year of publication of *Jìgōng quán chuán* is unknown, but is generally believed to be around the turn of the century, i.e., between 1875−1908, during the reign of Emperor Guangxu of the Qing dynasty.

Since sudden remembrance or realization normally involves the speaker's memory lapse, this type of affirmation ('you are right') marked by the TRUTH-DM normally involves speaker's self-reproachment about the lapse, thus, indicating the general negative-bias of *zhenshi*, as shown in the example above.

5. Discussion

In the foregoing sections, we have looked into diverse functions of TRUTH-DMs in Korean and Chinese with examples taken from historical and contemporary corpora. Now we turn to a discussion of some issues of theoretical importance, i.e., notable changes at various levels of grammar and commonalities and differences between Korean and Chinese.

5.1. Notable changes at various levels of grammar

DM development in general involves change at multiple levels of grammar. Among the most prominent changes is the semantic change. For instance, the meaning changes from the sentential meaning to the sentence-external meaning (Heine et al. 2020, Heine et al. 2021). The original 'tru(th)' meaning of TRUTH-DMs has undergone desemanticization, i.e., loss of meaning, in most usages of the TRUTH-DMs. Korean TRUTH-DM *cincca* used as a pause-filler is among the clearest instances of extensive desemanticization. In DM uses, the original content meaning of 'true' has become a speech-act unit (cf. Sweetser 1990) or of procedural function (Hansen 1997, Schourup 1999, Wilson 2011). For instance, usages with the content meaning, e.g., 'true love', 'truly happy' make a good contrast with the procedural meaning, e.g., 'what I'm saying truthfully is that...' or 'what you truly need to pay attention to is that...' As shown in these paraphrases, despite the desemanticization that has occurred to a variable extent, the functions of the TRUTH-DMs are directly or indirectly, related to 'truthfulness'.

It is generally accepted that DMs lose their ability to make propositional contributions, i.e., they do not contribute to the truth conditions (cf. Hansen 1997, Schourup 1999, Haselow 2017) and so they may be truth-conditionally optional (Jucker 2002). As illustrated in the foregoing exposition, the TRUTH-DMs are often positionally detached from the clause and thus its direct semantic link to the proposition is lost. The loss of propositional meaning is obvious in the contrast between 'a <u>true</u> story' (*cincca yayki*) and 'Look at this guy, truly!' (*i yangpani cincca*; in (3) above). This is in line with the observation that DM development may involve functional change from sentence-structuring function to metatextual function (Heine et al. 2020, 2021).

It has also been widely observed that DMs carry diverse functions (Jucker 2002, Rhee & Koo 2021). Indeed, the TRUTH-DMs in Korean and Chinese are particularly multifunctional. A brief mention of classification is in order in this context. An interesting aspect is that, as Rhee and Koo (2021: 77) note, the functions of de-contextualized DMs are hard to identify but those of the

contextualized DMs are largely unambiguous because diverse interpretive cues are available in the context, even though it is often difficult to name the function explicitly. The functions of DMs are subtle, and their classification largely depends on granularity of analyses. For instance, the discontent function and the challenge function of the Korean DM *cincca* (in 4.1.2 and 4.1.3, respectively), may be collapsed into one for their conceptual similarity of negative emotions, which, however, were separately named for their differences in directedness, i.e., self-directed discontent and other-directed challenge, as well as their differences in illocutionary force. Conversely, the discontent function in (15a) through (15c) and the reproaching function in (15d) of the Chinese DM *zhenshi* may be separated by the perceived differences in strength, i.e., mild discontent and strong reproach, which, however, were treated under a single function for greater similarity than differences.

Another notable aspect of DM development is the change in prosody. It has been observed that DMs tend to be set off prosodically or 'less integrated' (Heine et al. 2021: 48), and further that, as a result of erosion, a common concomitant of grammaticalization, DMs tend to be phonologically short (Brinton 2017, Heine et al. 2020, 2021). The TRUTH-DMs tend to have a pause before and/or after them, more rigidly so with Chinese DMs. In this context it is notable that Korean DMs occurring at the right-periphery position tend to be fully integrated with the preceding word (Koo & Rhee 2013, Kim & Sohn 2015), a tendency also observed in other languages (Mulder and Thompson 2008, among others). As for the length, the TRUTH-DMs in Korean and Chinese are indeed phonologically short, but they have not undergone phonological erosion to a meaningful level, largely due to the fact that the source constructions were already short from the beginning (note that all of them were bi-syllabic), not having enough phonological substance to be subjected to erosion.

The last point relates to positionality. As a result of grammaticalization of DMs, the forms change from positionally constrained placement to less constrained placement (Heine et al. 2021: 48). This is also the case with the TRUTH-DMs in Korean and Chinese: they are no longer in the modifier position as an adjective or adverb, or in the argument position as a noun. They typically occur at the left-periphery or right-periphery position, whereby their semantic-pragmatic scope is wider than before.

As shown above, a comparative investigation of the TRUTH-DMs in Korean and Chinese reveals a number of interesting commonalities, despite minor variations, with respect to meaning, function, prosody, and placement, among others (cf. Brinton 2017, Heine et al. 2014, 2020, 2021, Traugott 2022). However, there are differences in the two languages, to which we turn.

5.2. Korean and Chinese TRUTH-DMs compared

There are a few areas where the development of the TRUTH-DMs in Korean and Chinese diverge. Among the most prominent differences is that Chinese TRUTH-

DMs show different specialization between the two TRUTH-DMs, i.e., *zhenshi* being negatively-biased and *zhende* not being polarity-specific. Unlike *zhende*, *zhenshi* from the incipient stage frequently occurs in negative contexts, mostly marking discontent and sudden remembrance (affirmation of negative evaluation, self-reproach, etc.). In Modern Chinese, the negativity associated with *zhenshi* tends to be more stable and stronger than it was in the earlier times, i.e., it is used nearly always in negative contexts and the degree of negativity is stronger as in reproach and annoyance. This polarization strongly suggests contextual influences in functional development, similar to the notion of 'absorption' (Bybee et al. 1994). The other TRUTH-DM *zhende* does not exhibit such bias. The divergent development of Chinese TRUTH-DMs stands in sharp contrast to that of the Korean TRUTH-DM *cincca*, which emerges only a century ago, with a nominal or an adjectival use, and becomes a DM used in positive, negative, or neutral contexts.

Another difference between Korean and Chinese TRUTH-DMs lies in the different levels of semantic bleaching, as briefly indicated in 5.1 above. The Korean TRUTH-DM *cincca* can carry the function of pause-filling. Pause-fillers do not have any semantic content at a meaningful level: they are uttered just to fill the gap caused by the inability to produce unbroken utterance for various reasons, e.g., failure to find a proper word spontaneously (cf. 'delay device', Svartvik 1980), strategic display of hesitance for politeness (Rhee & Koo 2021), etc. The extreme level of semantic bleaching is easily observable with the English pause fillers *uh, um, y'know, well,* etc. (Östman 1981, Jucker 2015, Müller 2004), of which lexical meanings are either completely absent or not easily discernible. The pause-filling function is not found with the Chinese TRUTH-DMs *zhende* and *zhenshi*, which suggests that semantic bleaching is more extensive with the Korean *cincca*. The 'truth' meaning is relatively more strongly preserved in Chinese TRUTH-DMs (cf. 'persistence', Hopper 1991).

Still another difference is observed with respect to prosody. Prosody is often different depending on the function (Hancil 2013, Rhee 2020, Rhee & Koo 2021). Among prosodic features, the utterance length (i.e., speed) is particularly strongly correlated with the functions. For instance, Korean TRUTH-DM *cincca* is spoken fast when it carries the emphasis, surprise, and sudden remembrance functions, and it is spoken slowly when it carries the discontent and pause-filling functions. In the latter case, the DM occurs at the right-periphery with a trailing intonation, signaling inability to complete the utterance due to the difficulties in lexical search or due to overwhelming emotions in the contexts of irritation, frustration, etc. It is interesting to note that the TRUTH-DMs in Chinese, especially the negatively-biased *zhenshi*, do not exhibit trailing intonation pattern.

The results of comparison of the TRUTH-DM in Korean and Chinese with respect to their functions and notable features are summarized in Table 1.

Table 1. **Comparison of Korean and Chinese TRUTH-DMs**

Functions and features	Korean *cincca*	Chinese *zhende*	Chinese *zhenshi*
Emphasis	√	√	
Discontent	√		√
Challenge	√		
Surprise	√	√	
Sudden remembrance	√		√
Pause-filling	√		
Noteworthiness		√	
Disalignment		√	
Trailing intonation	√		
Positional freedom	√	√	√
Metatextuality	√	√	√
Desemanticization	more extensive	less extensive	less extensive
Evaluative bias			Negative

6. Summary and conclusion

This paper investigated the grammaticalized functions of TRUTH-DM in Korean and compared them with two Chinese TRUTH-DMs. The Korean lexeme *cincca* (in the form of *cinsca*) first occurs around the turn of the 20th century, and its grammaticalized form, the DM *cincca,* emerges in the mid-20th century, inheriting much of the 'truth' semantics from its predecessor forms. Korean *cincca* carries the functions of marking emphasis, discontent, challenge, surprise, sudden remembrance/realization, and pause-filling.

Chinese TRUTH-DMs *zhende* and *zhenshi* became univerbated as adverbs and then developed into DMs around the turn of the 20th century. The Chinese *zhende* carries the functions of marking emphasis, preface to noteworthy information, surprise, and preface to upcoming disalignment, whereas *zhenshi* carries the functions of marking discontent and sudden remembrance, the latter with a strong overtone of self-reproach. *Zhenshi* signals negative stance, revealing the influence of the context at the incipient stage. Thus, Chinese TRUTH-DMs show divergence in marking evaluative-emotional polarity, e.g., *zhenshi* is generally negatively-biased, whereas *zhende* is not.

The developments of TRUTH-DMs in Korean and Chinese share much in common, corroborating the thesis that grammatical functions largely depend on the semantics of the source lexeme (Heine et al. 1991, Heine 1997, Bybee et al. 2024). The conceptual motivation of the emergence of TRUTH-DMs in Korean and Chinese seems to be based on the conceptualization of 'truth' as being not easily accessible or hidden, thus its manifestation being extraordinary and its realization inducing surprise (cf. Rhee 2022). All DM functions are directly or indirectly related to the notion of 'truth', but semantic bleaching is extensive with the pause-filler function. Korean and Chinese TRUTH-DMs exhibit features typical of grammaticalization in terms of their meaning, function, placement, etc., but phonological erosion is not prominent.

Abbreviations

ACC: accusative; ADN: adnominalizer; CLSF: classifier; CONN: connective; DM: discourse marker; END: sentence-ender; EXCL: exclamative; GEN: genitive; HON: honorific; HORT: hortative; INTJ: interjective; NOM: nominative; PERF: perfective; POL: polite; PST: past; PTCL: particle; PURP: purposive; RC: resultative complementizer

Acknowledgments

References

Adamou, Evangelia & Yaron Matras (eds.) 2020. *The Routledge Handbook of Language Contact.* London: Routledge.

Aijmer, Karin. 1986. Why is *actually* so popular in spoken English? In Gunnel Tottie & Ingegerd Bäcklund (eds.), *English in speech and writing: A symposium,* 119–129. Uppsala: Almqvist & Wilsell International.

Bisang, Walter. 2006. Contact-induced convergence: Typology and a reality. In Keith Brown (ed.), *Encyclopedia of language & linguistics,* 2nd ed. Vol. 3, 88–101. Oxford: Elsevier.

Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: Cambridge University Press.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Čapková, Andrea. 2015. *On adverbial English-French faux amis: A contrastive view of actually and actuellement.* BA thesis, Praha, Czechia.

Chen, Ying. 2010. "Zhēnde" de xū huà, yǔyán yánjiū [Grammaticalization of *zhende*]. *Yǔyán yánjiū* 30 (4). 62–66.

Dong, Xiufang. 2005. Yǔyì yǎnbiàn de guīlǜ xìng jí yǔyì yǎnbiàn zhōng bǎoliú yì sù de xuǎnzé [The regularity of semantic evolution and the selection of preserved sememes in semantic evolution]. In Yun Lu Wang (ed.), *Hànyǔ shǐxué bào* 5, 287–293.

Fang, Qingming. 2012. Zài lùn "zhēn" yǔ "zhēn de" de yǔfǎ yìyì yǔ yǔ yòng gōngnéng, [Revisiting the grammatical meaning and pragmatic function of *zhen* and *zhende*]. *Hànyǔ Xuéxí* 5. 95–103.

Grant, Anthony P. (ed.) 2020. *The Oxford Handbook of Language Contact.* Oxford: Oxford University Press.

Guo, Xiaolin. 2015. "Zhēnshi de" fùmiàn píngjià gōngnéng tànxī. [An analysis of the function of *zhenshide* negative evaluation]. *Yǔyán Jiàoxué Yǔ Yánjiū* 1. 97–104.

Han, Zhizhou. 2016. "Zhēn de" de huàyǔ biāojì gōngnéng fēnxī [Functional analysis of the discourse marker zhende]. *Cháng'ān Xuéshù* 10. 182–190.

Hancil, Sylvie. 2013. Introduction. In Sylvie Hancil & Daniel Hirst (eds.), *Prosody and iconicity,* 1–31. Amsterdam: John Benjamins.

Hansen, Maj-Britt Mosegaard. 1997. *Alors* and *donc* in spoken French: A reanalysis. *Journal of Pragmatics* 28. 153–187.

Haselow, Alexander. 2017. *Spontaneous Spoken English. An Integrated Approach to the Emergent Grammar of Speech*. Cambridge: Cambridge University Press.

Heine, Bernd. 1997. *Cognitive Foundations of Grammar.* Oxford: Oxford University Press.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework.* Chicago: The University of Chicago Press.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers.* Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2011. The areal dimension of grammaticalization. In Heiko Narrog & Bernd Heine (eds.), *The Oxford handbook of grammaticalization,* 291–301. Oxford: Oxford University Press.

Heine, Bernd, Tania Kuteva & Gunther Kaltenböck. 2014. Discourse grammar, the dual process model, and brain lateralization: Some correlations. *Language & Cognition* 6 (1). 146–180.

Heine, Bernd, Tania Kuteva & Haiping Long. 2020. Dual process frameworks on reasoning and linguistic discourse: A comparison. In Alexander Haselow & Gunther Kaltenböck (eds.), *Grammar and cognition: Dualistic models of language structure and language processing,* 59–89. Amsterdam: John Benjamins.

Hopper, Paul J. 1991. On some principles of grammaticalization. In Elizabeth Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, 2 vols. Vol. 1, 17–35. Amsterdam: John Benjamins.

Jucker, Andreas H. 2002. Discourse markers in Early Modern English. In Richard Watts & Peter Trudgill (eds.), *Alternative histories of English*, 210–230. London: Routledge.

Jucker, Andreas H. 2015. Pragmatics of fiction: Literary uses of *uh* and *um. Journal of Pragmatics* 86. 63–67.

Khammee, Kultida (2024). From objective to subjective and to intersubjective functions: The case of the Thai 'truth'-lexeme. *Russian Journal of Linguistics* 28 (4). 942–965. https://doi.org/10.22363/2687-0088-40496

Kim, Stephanie Hyeri & Sung-Ock S. Sohn. 2015. Grammar as an emergent response to interactional needs: A study of final *kuntey* 'but' in Korean conversation. *Journal of Pragmatics* 83. 73–90. https://doi.org/10.1016/j.pragma.2015.04.011

Kim, Taeho & Seon-yeong Jeong. 2012. A corpus-based study of the truth-related words in Korean used as discourse markers. *Cross-cultural Studies* 29. 453–477.

Koo, Hyun Jung. 2009. Force dynamics as a variational factor: A case in Korean. *LACUS Forum* 33. 201–210.

Koo, Hyun Jung & Seongha Rhee. 2013. On an emerging paradigm of sentence-final particles of discontent: A grammaticalization perspective. *Language Sciences* 37. 70–89 https://doi.org/10.1016/j.langsci.2012.07.002

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World Lexicon of Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.

Li, Jijun. 2003. "Zhēn" hé "zhēn de" lùn xī —jùfǎ, yǔyì, yǔ yòng gōngnéng jí qí yǔfǎ huà guòchéng hé jīzhì fēnxī [*Analysis of zhen and zhende: Analysis of syntax, semantics, pragmatic functions and their grammaticalization process and mechanism*]. MA thesis, Shanghai Normal University, China.

Li, Xianyin. 2015. Jīyú zìrán kǒuyǔ de huàyǔ fǒudìng biāojì "zhēnshi" yánjiū [Research on negative marking of discourse *zhenshi*, based on natural spoken language]. *Yǔyán Jiàoxué Yǔ Yánjiū* 3. 59–69.

Li, Xiaojun. 2011. Biǎo fùmiàn píngjià de yǔ yòng shěnglüè [Pragmatic ellipsis for negative evaluation]. *Dāngdài Xiūcíxué* 166. 35–42.

Liu, Chenyang. 2021. Hùdòng jiāojì zhōng "zhēn de" de huàyǔ biāojì gōngnéng [The discourse marker function of *zhende* in interactive communication]. *Císhū Yánjiū* 1. 111–120.

Long, Haiping, Feng Gu & Xiaoping Xiao (translation), Bo Hong & Feng Gu (annotation). 2012. *Translations of World Lexicon of Grammaticalization.* Beijing: Beijing World Publishing Corp.

Mulder, Jean & Sandra A. Thompson. 2008. The grammaticalization of *but* as a final particle in conversation. In Laury Ritva (ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions,* 179–204. Amsterdam: John Benjamins.

Müller, Simone. 2004. 'Well you know that type of person': Functions of *well* in the speech of American and German students. *Journal of Pragmatics* 36. 1157–1182.

Östman, Jan-Ola. 1981. *You Know: A Discourse Functional Approach*. Amsterdam: John Benjamins.

Rhee, Seongha. 2020. On determinants of discourse marker functions: Grammaticalization and discourse-analytic perspectives. *Linguistic Research* 37 (2). 289–325. https://doi.org/10.17250/khisli.37.2.202006.005

Rhee, Seongha. 2021. On grammaticalization of truthfulness-based emphatic discourse markers. In Fuxiang Wu, Yonglong Yang, Haiping Long (eds.), *Grammaticalization and studies of Grammar* 10. 461–499. Beijing: Commercial Press.

Rhee, Seongha. 2022. Where to go at the end: Polylexicalization and polygrammaticalization of *kaz* 'edge' in Korean *Russian Journal of Linguistics* 26 (3). 571–595. https://doi.org/10.22363/2687-0088-30616

Rhee, Seongha & Hyun Jung Koo. 2021. On divergent paths and functions of 'background'-based discourse markers in Korean. In Alexander Haselow & Sylvie Hancil (eds.), *Studies at the Grammar-Discourse interface,* 77–100. Amsterdam: John Benjamins.

Schourup, Lawrence. 1999. Discourse markers. *Lingua* 107. 227–265.

Svartvik, Jan. 1980. 'Well' in conversation. In Sydney Greenbaum, Geoffrey Leech & Jan Svartvik (eds.), *Studies in English linguistics for Randolph Quirk,* 167–177. London: Longman.

Sweetser, Eve E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 2022. *Discourse Structuring Markers in English*. Amsterdam: John Benjamins.

Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Wang, Yuanyuan. 2010. Xiàndài hànyǔ zhōng "zhēnshi" de gōngnéng fēnxī [The functional analysis of *zhenshi* in Modern Chinese]. *Xiàndài Yǔwén* 29. 27–29.

Wilson, Deirdre. 2011. The conceptual-procedural distinction: Past, present and future. In Victoria Escandell-Vidal, Manuel Leonetti & Aoife Ahern (eds.), *Procedural meaning: Problems and perspectives,* 3–31. Bingley: Emerald Group Publishing.

Yan, Hongju. 2006. Huàyǔ biāojì de zhǔguān xìng hé yǔfǎ huà [The subjectivity and grammaticalization of discourse markers]. *Húnán Kējì Dàxué Xuébào* 9. 80–85.

Zhang, Yisheng. 2004. *Xiàndài Hànyǔ Fùcí Tànsuǒ* [*Exploration of Adverbs in Modern Chinese*]. Xuélín Publishing.

Dictionaries and Internet resources

CCL Corpus (The Center for Chinese Linguistics at Peking University Corpus). n.d., http://ccl.pku.edu.cn:8080/ccl_corpus/.

Koesacen [A Historical Dictionary]. 2007[1997]. Seoul: Kyohaksa Publishing.

Phyocwun Kwuke Taysacen. 1999. Seoul: The National Institute of Korean Language.

Wulimalsaym [Korean Dictionary]. n.d. https://opendict.korean.go.kr

Article history:

Bionotes:

Seongha RHEE is Professor of Linguistics at Mahidol University, Thailand and Professor Emeritus at Hankuk University of Foreign Studies, Korea. He received his Ph.D. in linguistics from the University of Texas, Austin in 1996. His primary research interest is to identify cognitive and discursive mechanisms that enable language change from the crosslinguistic and typological perspectives. He published *World Lexicon of Grammaticalization* (co-author, 2019, CUP); book chapters in *The Cambridge Handbook of Korean Linguistics* (2022, CUP); and research articles in *Nature*, *Journal of Pragmatics, Language Sciences, Russian Journal of Linguistics* and *Lingua,* among others.
e-mail: srhee@hufs.ac.kr
https://orcid.org/0000-0003-0312-0975

Lin ZHANG is a doctoral candidate in the Department of Korean Language and Literature at Seoul National University. Her primary research interests focus on historical development of lexical and grammatical forms in Korean, as well as comparative diachronic studies between Korean and Chinese.
e-mail: linwkd@snu.ac.kr
https://orcid.org/0009-0004-3826-536X

Сведения об авторах:

Санха РИ – профессор лингвистики Университета Махидол, Таиланд, и почетный профессор Ханкукского университета иностранных исследований, Корея. Получил степень PhD по лингвистике в Техасском университете в Остине в 1996 году. Сфера его научных интересов – выявление когнитивных и дискурсивных механизмов изменения языка в межъязыковой и типологической перспективе. Среди его публикаций – книга *World Lexicon of Grammaticalization* (2019, CUP); главы книг в *The Cambridge Handbook of Korean Linguistics* (2022, CUP), статьи в журналах *Nature*, *Journal of Pragmatics*, *Language Sciences*, *Russian Journal of Linguistics*, *Lingua* и др.
e-mail: srhee@hufs.ac.kr
https://orcid.org/0000-0003-0312-0975

Линь ЧЖАН – докторант кафедры корейского языка и литературы Сеульского национального университета. Ее основные исследовательские интересы сосредоточены на историческом развитии лексических и грамматических форм в корейском языке, а также на сравнительных диахронических исследованиях корейского и китайского языков.
e-mail: linwkd@snu.ac.kr
https://orcid.org/0009-0004-3826-536X

Research article / Научная статья

From *truth* to *truly*: The case of *shinni* 'truly' in Japanese compared to Chinese, Korean and Thai counterparts

Yuko HIGASHIIZUMI[1] iD ✉, **Reijirou SHIBASAKI[2]** iD **and Keiko TAKAHASHI[1]** iD

[1]Toyo University, Tokyo, Japan
[2]Meiji University, Tokyo, Japan
✉higashiizumi.yuko.1@gmail.com

Abstract

This study traces the development of the adverb *shinni* 'truly', a hybrid form consisting of the Sino-Japanese noun *shin* (眞/真) 'truth' and the adverbializer *-ni* of Japanese origin, in the history of Japanese. The goal of the study is to compare the developmental pathway of *shinni* with that of 'truth'-related words derived from the same Chinese word 眞/真 in Chinese, Korean, and Thai in order to seek commonalities and differences among them. In these languages, it is reported that some words with 眞/真 have developed from "TRUE" to "INTENSIVE", and have further developed a number of interactional functions as discourse markers (DMs). The data were obtained from various historical and modern corpora and database. A total of 1810 occurrences of *shinni* in written and spoken Japanese of various genres were analysed with the focus on their pragmatic function. The study confirms a commonality in the development from "TRUE" to "INTENSIVE" in *shinni*, illustrating its evolution from the noun *shin* ("TRUE") to the use of *shinni* that can be interpreted as serving to intensify the illocutionary force of what is being said (i.e., "INTENSIVE"). In contrast, this study finds that *shinni* does not bear any DM functions that some words with 眞/真 in Chinese, Korean, and Thai have developed. This study discussed the possibility that, due to various factors, the pace of change at advanced stages of grammaticalization may be more diversified than has previously been suggested.

Keywords: *adverb, discourse marker, illocutionary force, intensive, shinni, Sino-Japanese*

От *истины* к *истинно*: *shinni* в японском языке в сравнении с китайскими, корейским и тайским когнатами

Юко ХИГАСИИДЗУМИ¹ ⊠, Рейджиро ШИБАСАКИ², Кейко ТАКАХАСИ¹

¹Университет Тойо, Токио, Япония
²Университет Мэйдзи, Токио, Япония
⊠higashiizumi.yuko.1@gmail.com

Аннотация

В настоящем исследовании прослеживается развитие наречия *shinni* 'истинно' – гибридной формы, состоящей из китайско-японского существительного *shin* (眞/真) 'истина' и наречия *ni* японского происхождения – в японском языке в диахроническом аспекте. Цель исследования сравнить траекторию развития *shinni* с траекторией слов, производных от того же китайского слова 眞/真 в китайском, корейском и тайском языках, и выявить их общие черты и различия. В этих языках некоторые слова с 眞/真 развились от «ИСТИНА/ПРАВДА» до «УСИЛИТЕЛЬ» и в дальнейшем приобрели ряд функций дискурсивных маркеров (ДМ). Данные были получены из различных исторических и современных корпусов, а также баз данных. В общей сложности были проанализированы 1810 случаев использования *shinni* в письменных и устных японских текстах различных жанров с акцентом на их прагматических функциях. Исследование подтверждает общность в развитии *shinni* от «ИСТИНА» до «УСИЛИТЕЛЬ» иллюстрируя его эволюцию от существительного *shin* («ИСТИНА») до *shinni* – как усилителя иллокутивной силы высказывания. Исследование показывает, что *shinni* не несет никаких функций ДМ, которые развились у некоторых слов с 眞/真 в китайском, корейском и тайском языках. Полученные результаты позволяют сделать предположение о том, что из-за различных факторов темпы изменений, происходящих на продвинутых стадиях грамматикализации, могут не совпадать.

Ключевые слова: *наречие, дискурсивный маркер, иллокутивная сила, интенсивный,* shinni*, китайско-японская интерференция*

1. Introduction

This study examines the historical development of the Sino-Japanese noun 眞/真 *shin* 'truth, sincerity, righteousness, genuineness' into the adverb *shinni* 'truly, genuinely, really', which consists of *shin* of Chinese origin and the adverbializer *-ni* of Japanese origin. It aims to compare it with the development of discourse markers (DMs, hereafter) that originate from the same Chinese word 眞/真 in Chinese, Korean, and Thai (see Rhee & Zhang 2024, Khammee 2024).[1]

---

[1] The Chinese characters 眞 and 真 are both used in the history of Japanese. In modern standard Japanese, *shinni* is written as 真に (with the latter form 真 and the Japanese phonographic

Kuteva et al. (2019: 443) have demonstrated that some words with the meaning "TRUE" ('true', 'real') have developed an "INTENSIVE" meaning in some languages, including Chinese, English, French, and Hungarian. Similarly, in Japanese, the Sino-Japanese noun *shin* 'truth', i.e., "TRUE", has developed into an adverb *shinni* 'truly' that can be interpreted as functioning as "INTENSIVE". In Chinese, Korean, and Thai, "TRUE"-related words that share the etymon 眞/真 have also developed into "INTENSIVE" and have further developed a number of interactional functions as DMs (Rhee & Zhang 2024, Khammee 2024). In contrast, *shinni* has not developed such interactional functions. The goal of the current study is to explore the history of the adverb *shinni*, drawing on various historical and contemporary corpora, to seek commonalities and differences in the words originating from 眞/真 in Chinese, Korean, and Thai. It aims to answer the following research questions: (i) whether a "TRUE"-related word 眞/真 *shin* in Japanese has developed into an "INTENSIVE", and (ii) whether it has further developed interactional functions as DMs in a similar way to "TRUE"-related words that share the etymon 眞/真 in Chinese, Korean, and Thai.

This paper is organized as follows: Section 2 provides some background to the current study. Section 3 introduces the corpora and methodology used in the current study, and Section 4 presents the survey results. Section 5 discusses the results in relation to pragmatic functions and gives some examples of Sino-Japanese "TRUE"-related words that carry similar functions to the DMs originating from 眞/真 in Chinese, Korean, and Thai. Section 6 gives some concluding remarks.

2. Background

2.1. From TRUE to INTENSIVE in grammaticalization

Kuteva et al. (2019: 443) provide some words with the meaning "TRUE ('true', 'real')" that have evolved to serve the function of "INTENSIVE" in Chinese, English, French, Hungarian, American Sign Language, and Baka.[2] Such examples are summarized in (1) based on Kuteva et al. (2019). Note that in the case of Hungarian *igaz* 'true' and *igaz-án* 'really', Kuteva et al. (2019) juxtapose them with a comma, instead of showing the direction of change, e.g., *igaz* 'true' > *igaz-án* 'really'; we follow their description here.

---

character に). For convenience, the glossing for the noun *shin* and the adverb *shinni* are standardized to 'truth' and 'truly' respectively in what follows. Discourse markers have been discussed and described using various terms in the literature. In this study, we will use the term discourse marker as an umbrella term for expressions that can serve discourse-pragmatic functions.

[2] Ethnologue (https://www.ethnologue.com) gives two languages named Baka; one is Nilo-Saharan and the other is Niger-Congo.

(1)   TRUE ('true, real') > INTENSIVE
        Archaic Chinese *qing* 'truth' > *qing*, intensive marker, adverb
        Early Modern Chinese *zhen* 'truth' > *zhen*, intensive marker, adverb
        Old French *verai* 'tru(ly), truthful(ly)' > (borrowing) Middle English *verray* >
        Modern English *very*
        Hungarian *igaz* 'true', *igaz-án* 'really'
        American Sign Language *TRUE > INTENSIVE*
        Baka *ko* 'truly', 'really', 'completely', adverb > 'very', intensive marker

In a similar manner, the Sino-Japanese noun *shin* 'truth', i.e., "TRUE", has evolved to serve the function of "INTENSIVE", as in (2). In (2), the adverb *shinni* 'truly', a hybrid form consisting of *shin* of Chinese origin and the adverbializer *-ni* of Japanese origin, serves to intensify the illocutionary force of apology, meaning 'I am very sorry'. At the same time, *shinni* sounds rather archaic in contemporary Japanese. Note that (2) is indeed taken from a historical novel about the Zen monk named Dōgen (1200–1253). The adverb *shinni* has the stylistic effect of creating the impression of a conversation between people from long ago for modern readers.

(2)   Intensive adverb (intensifying the illocutionary force)
        あなたのお心を挫くようで真に申し訳ないのですが、…ご老師の看病を
        お願いできないでしょうか？
        anata-no-o.kokoro-o   *kujiku-yō-de*
        you-GEN-RES.heart-ACC   discourage-way-COP.CVB
        shinni   *mōshiwakenai-no-desu-ga,*   [...]
        truly   sorry-NML-COP.POL-but
        go.rōshi-no-kanbyō-o   *onegai-deki-nai-deshō-ka?*
        RES.old.master-GEN-care-ACC   ask-can-not-COP.CONJ.POL-QP
        "I am truly/very sorry to discourage you, but […] could you please take care of
        your old master?"
                    (BCCWJ, PB29_00269, 23790, Mizushima Hajime, *Dōgen,* 2002)[3]

The developmental pathway from the Sino-Japanese noun *shin* 'truth' to the adverb *shinni* 'truly' that can be interpreted as meaning 'very' is in line with "TRUE ('true', 'real')" to "INTENSIVE" suggested by Kuteva et al. (2019).

2.2. A note on the Sino-Japanese shin(ni)

The basic meaning of the Sino-Japanese noun *shin* is 'truth' (Kadokawa 1982–1999, Todo 1978, Morohashi 2018). According to Nihon Kokugo Daijiten (Nikkoku 2006), the noun *shin* is attested in *Shōmangyōgisho*, a Buddhist commentary written in the early seventh century, while the adverb *shinni* 'truly', i.e., the noun *shin* agglutinated by the adverbializer *-ni* of Japanese origin, is found in *Sasamegoto*, the theory of Renga poems written in 1463–1464. Later, *shinni* is

---

[3] Information on the examples found by the Chunagon corpus search application (Corpora (b)–(j) in Tables 1 and 2) is indicated as follows: the abbreviation for the corpus, sample ID, the beginning number in the corpus, (Author), (Title), and the year of the data. The bolding in the examples is by the authors.

used in conversation in novels written in a colloquial style, as in (3). The adverb *shinni* modifies the verb phrase *o-ide-nansu* '(you) will go' in (3a) and the adjective *ureshii* 'happy' in (3b). Note that *shinni* in (3b) can be interpreted as meaning 'very', i.e., "INTENSIVE" (Kuteva et al. 2019) (see Section 5.1 for further discussion).

(3) Adverb (modifying the predicate)

   a. じらさづとも、**しんに**おいでなんすところを、けふはあかしておきか
      せなんし
      jiras-azu-tomo, ***shinni*** *oidenansu-tokoro-o,*
      tease-not-even.if **truly** go.RES-place-ACC
      kyō-wa *akashi-te* *okikasenanshi*
      today-TOP reveal-CVB tell.RES.IMP
      "Stop teasing, reveal (to me) and let me hear where you will really go today."
      (Nikkoku, *Sharebon Keiseikaifutasujimichi*, 1798)

   b. 真に嬉しいヨ、どふぞ
      shinni *ureshii-yo,* *dōzo*
      truly be.happy-FP please
      "(I) am truly/very happy, please."
      (Nikkoku, *Ninjōbon Shunshokuumegoyomi*, 1832–33)

In contemporary Japanese dictionaries, *shinni* is described as an adverb or an adverb-like phrase that modifies the predicate (e.g., Hida & Asada 2018, Daijirin 2019, Iwanami 2019). Hida & Asada (2018: 196) comment that in contemporary Japanese, *shinni* is employed in formal writing, as shown in (4).

(4) Adverb (modifying the predicate)
     彼は**しんに**妻を愛していた。
     kare-wa ***shinni*** *tsuma-o* *aishi-te-i-ta.*
     he-TOP **truly** wife-ACC love-CVB-be-PST
     "He truly loved (his) wife."
     (Hida & Asada 2018: 196)

These dictionaries allow us to trace the historical process that led to the evolution of the Sino-Japanese lexeme *shin* from a noun meaning 'truth' to the adverb *shinni* meaning 'truly'. However, to deepen our understanding of *shinni*, we will extend our research by analyzing data retrieved from both written and spoken Japanese corpora. In the following sections, therefore, we will conduct a detailed corpus-based analysis of *shinni*.

3. Data and methodology

The data used for this study were obtained from various historical and modern corpora as well as database listed in Table 1 (see Appendix for further information). Database (a) SZ provides online access to a collection of Japanese classical literature. From this database, we manually collected data, ensuring that each

instance was annotated as *shinni*. Corpora (b)–(j) were developed by the National Institute for Japanese Language and Linguistics. From these corpora, examples that are annotated as *shinni* were collected using the Chunagon corpus concordance system. We then checked them manually and removed any irrelevant examples from the database used in the current study.

Table 1. **Database and corpora used for this study**

Database and corpora	Periods
Written Japanese	
(a) SZ=*Shimpen Nihon Koten Bungaku Zenshū*	late 7th c.–late 19th c.
(b) CHJ=*The Corpus of Historical Japanese*	8th c.–early 20th c.
(c) SHC=*Showa-Heisei Corpus of Written Japanese*	1933–2013
(d) BCCWJ=*Balanced Corpus of Contemporary Written Japanese,* excluding the sub-corpus *Diet Record* (Corpus (f) DR)	1971–2008
Spoken Japanese	
Formal speech (mostly monologue)	
(e) SSC=*Showa Speech Corpus*	1952–1969
(f) DR=the sub-corpus *Diet Record* in Corpus (d) BCCWJ	1971–2008
(g) CSJ=*The Corpus of Spontaneous Japanese*	1999–2001
Daily conversation	
(h) NUCC= *Nagoya University Conversation Corpus*	1993–2000
(i) CWPC=*Gen-Nichi-Ken Corpus of Workplace Conversation*	1993 & 1999–2000
(j) CEJC=The Corpus of Everyday Japanese Conversation	2016–2019

A total of 1810 occurrences of *shinni* in written and spoken Japanese of various genres were analysed with the focus on their pragmatic function.

4. Results

Table 2 presents a summary of the survey results. Here we will highlight two main observations and provide examples to illustrate each point.

Table 2. **The occurrence of *shinni* in the database and corpora used for this study**

Database and corpora		Periods	Freq. (pmw)	Ill. force
written	(a) SZ	late 7th c.–late 19th c.	5 (NA)	0
	(b) CHJ	Edo (1603–1867)	24 (28.9)	0
		Meiji (1868–1912)	76 (8.1)	0
		Taisho (1912–1926)	116 (23.6)	0
	(c) SHC	Showa (1926–1989)	711 (29.7)	2
		Heisei (1989–2019)[4]	92 (9.3)	0
	(d) BCCWJ (excl. DR)	1971–2008	701 (7.0)	8
		Total	1725	10

[4] We would like to thank Jiyeon Park for the comment that it is possible that the use of some complementary words, such as *honto*, was noticeably prevalent in the Heisei period. Although this may be the case, it is outside of the scope of the current study, so we will leave it for future investigation. See the brief descriptions in Table 4 in Section 5.3.

		Database and corpora	Periods	Freq. (pmw)	Ill. force
spoken	formal speech	(e) SSC	1952–1969	3 (5.7)	0
		(f) DR	1971–2008	76 (14.9)	0
		(g) CSJ	1999–2001	6 (0.8)	0
	daily conv.	(h) NUCC	1993–2000	0	0
		(i) CWPC	1993 & 1999–2000	0	0
		(j) CEJC	2016–2019	0	0
			Total	85	0

Freq.=raw frequency; pmw=per million words; Ill. force=raw frequency of the instances in which the adverb *shinni* can be interpreted as serving to intensify an illocutionary force of what is being said.

The first observation is that the adverb *shinni* is present in the database and corpora of written Japanese (Database (a) and Corpora (b)–(d)), as well as in the corpora of formal speech (Corpora (e)–(g), mainly monologues), but no instances of *shinni* are found in the corpora of daily conversation (Corpora (h)–(j)).[5]

In the corpora of written Japanese, (5) is the earliest attestation of the adverb *shinni*, found in a travel story which showcases famous historical landmarks in Edo/Tokyo for readers through two fictive characters visiting there. In (5), *shinni* modifies the nominal predicate *zokushi-nari* '(the person) is a vulgarian'.

(5) Adverb (modifying a predicate)
その心徳を得ずんば、**真に**俗子なり。

sono-kokoro	*toku-o*	*e-zunba,*	***shinni***	*zokushi-nari.*
that-heart	virtue-ACC	obtain-not.if	**truly**	vulgarian-COP

"If (one's) heart does not obtain the virtue, the person is truly a vulgarian."

(SZ: Toda Mosui, *Murasaki no Hitomoto*, 1628)

It is interesting to note here that the occurrence of *shinni* is relatively frequent in conversation in novels written in the Edo period (1603–1867), while it is not frequent in the Meiji (1868–1912) and Taisho (1912–1926) periods, as shown in Table 3. To be more specific, 17 out of 24 (70.8%) instances of *shinni* are found in conversation in novels written in the late eighteenth century to the mid-nineteenth century.

Table 3. **The occurrence of *shinni* in conversation in novels in Corpus (b) CHJ**

Corpus	Periods	Total	*shinni* in conversation in novels	
		Freq.	Freq.	%
(b) CHJ	Edo (1603–1867)	24	17	70.8
	Meiji (1868–1912)	76	11	14.5
	Taisho (1912–1926)	116	8	6.9

[5] Of the instances of *shinni* across Corpora (e)–(g), only a single example of *shinni* is found in dialogue in Corpus (e) SSC; all other examples are used in monologue.

In the corpora of spoken language, the earliest attestation of *shinni* is (6). In (6), the male speaker is talking about what he thinks is important for new employee training. In (6), *shinni* modifies the verb *shugan-to-suru* 'think of as the main point'.

(6) Adverb (modifying a predicate)
　　［略］言葉遣いが悪かったらもう何にもうまくないんだとゆうことでわたくしえーしんにそういうことを主眼としてですねえー［略］わたし考えているんです

[…]	*kotoba.zukai-ga*	*warukat-tara,*	*mō*	*nanni-mo*
	language.use-NOM	bad.PST-if	any.more	anything-PT

umaku-nai-n-da-toyū-koto-de　　　　　*watakushi*　　*ē*　　***shinni***
good-not-NML-COP-QUOT-NML-COP.CVB　I　　　　　uh　　**truly**
sō-yū-koto-o　　　　*shugan-toshite-desu-nē*　　　　　［…］
so-QUOT-thing-ACC　main.point-as-COP.POL-FP
watashi　　　　　　*kangaeteiru-n-desu*
I　　　　　　　　　think-NML-COP.POL
"[Even if the facilities are good, if the customer service has a bad attitude and] if (the customer service has) a bad language, it won't be any good at all, and I uh truly think of these things as the main point (of new employee training), […] I think."
　　　　　　　　　　　　　　　　　　(SSC: C52_06_CT, 65150, 1952)

The second observation concerns the adverb *shinni* that can be interpreted as serving to intensify the illocutionary force of what is being said, as in (2). In our written Japanese database (Database (a) and Corpora (b)–(d)), there were only 10 instances of this specific usage out of 1,725, representing approximately 0.57% (see Section 5.2 for further details). No instances were found in our spoken Japanese database (Corpora (e)–(j)).

Example (7) illustrates that the adverb *shinni* can be seen as serving to intensify the illocutionary force of regret, expressing 'I truly/deeply regret' or 'I am truly/very sorry'.

(7) Intensive adverb (intensifying the illocutionary force)
　　［略］また快方に向つたことをいつて居られたので、安心して居つたのは眞に相濟まなかつた。

[…]	*mata*	*kaihō.ni.mukat-ta-koto-o*	*it-te-orare-ta-node,*
	again	convalesce-PST-NML-ACC	say-CVB-be.RES-PST-because

anshinshi-te-ot-ta-no-wa　　　　　　***shinni***　　*aisumanakat-ta.*
relieve-CVB-be.HUM-PST-NML-TOP　　**truly**　　be.sorry-PST
"Because (he) said (he) was getting better again, (I) truly/deeply regret that (I felt) relieved (and didn't do anything for him)"".
　　　　　　　　　　　　　　　　(SHC: 70MChūkō1933_05049, 3550, 1933)

In summary, our survey results indicate that *shinni* is predominantly used in written Japanese and rarely appears in spoken Japanese; if at all, it is typically found only in formal monologues. Additionally, our data reveal that *shinni* serves to intensify the illocutionary force in only a few instances.

5. Discussion

5.1. "TRUE" to "INTENSIVE" seen in the Sino-Japanese shin(ni)

This section traces the historical development of the Sino-Japanese noun *shin* 'truth' to the adverb *shinni* 'truly' that can be interpreted as functioning as "INTENSIVE", as in (2) and (7), presenting examples found in our database (see Table 1), to confirm that the development is consistent with "TRUE" to "INTENSIVE" in grammaticalization (Kuteva et al. 2019).

In a similar way to the description in dictionaries mentioned in Section 2.2, the noun *shin* occurred earlier than the adverb *shinni*. Example (8) is the earliest attestation of the noun *shin* 'truth, sincerity, righteousness' used in a history book.

(8) Noun
性に率ひ**真**に任せ、嬌飾する所無し。

sei-ni	*shitagai*	**shin**-*ni*	*makase,*	*kyōshokusuru-tokoro*	*nashi.*
nature-PT	following	truth-PT	depending	distort.embellish-NML	not

"(Emperor Suinin) following (his) nature and depending on (his) sincerity, (he) did not distort or embellish (things)."

(SZ: *Nihonshoki*, 720)

The adverb *shinni* was attested in later centuries in our database. The earliest example is (5), in which *shinni* modifies the nominal predicate. Example (9) is the second-earliest example of *shinni* in our database. It modifies the nominal adjective *iya* 'unwilling'.

(9) Adverb (modifying the predicate)
真にいやと思ふ客人が来ても

shinni	*iya-to*	*omou-kyakujin-ga*	*ki-te-mo*
truly	unwilling-QUOT	feel-guest-NOM	come-CVB-even.if

"Even if a guest (you) think (you are) truly/very unwilling (to entertain) comes (to you), [you must welcome (him).]"

(SZ: Umeborikogaku, *Sharebon Keiseikaifutasujimichi*, 1798)

It is worth noting here that when *shinni* modifies a nominal adjective (e.g., *iya* 'unwilling' in (9)), or an adjective (e.g., *ureshii* 'happy' in (3b) and *kanashi* 'sad' in (10)), it can be interpreted as meaning 'very'.

(10) Adverb (modifying the predicate)
そんな無理な事ばかり。おつせへすと**しんに**哀しくなりいす

sonna	*murina-koto-bakari.*	*ossēsu-to*	**shinni**	*kanashiku-nariisu*
such	impossible-thing-just	say.RES-if	**truly**	sad.ADV-become.POL

"If (you) say just such impossible things, (I) become truly/very sad".
(CHJ: 52-Share1822_01062, 187720, Hanasanjin, *Sharebon Satokagami*, 1822)

It is likely that this ambiguity in interpretation may have given rise to the "INTENSIVE" *shinni* that can be regarded as serving to intensify the illocutionary force, as in (2) and (7), because the "INTENSIVE" *shinni* was found later in the early twentieth century as far as our database is concerned. Therefore, the use of *shinni*

to modify a nominal adjective or an adjective can be seen as a "bridging context" (Heine 2002, Narrog & Heine 2021: 58–61) or a "critical context" (Diewald 2002).

In conclusion, although, as reported in Section 4, the "INTENSIVE" *shinni* occurred infrequently, it is suggested that the developmental pathway of the noun *shin* 'truth' via the adverb *shinni* 'truly' to the "INTENSIVE" *shinni* 'very' conforms to "TRUE" to "INTENSIVE" (Kuteva et al. 2019).

5.2. Pragmatic functions of shinni

In this section, we will focus on the pragmatic functions of *shinni*. According to Rhee & Zhang (2024) and Khammee (2024), "TRUE"-related words in Chinese, Korean, and Thai that originated from 眞/真 have also developed into "INTENSIVE" and have further developed a number of functions as DMs. More specifically, the Korean *cinnca* serves to mark emphasis, frustration/annoyance, challenge, surprise, pause-filling, and sudden remembrance, while these functions are marked by *zhende* and *zhenshi* in Chinese; *zhende* is used to mark emphasis, and as a preface to noteworthy information, surprise, or upcoming disalignment, and *zhenshi* marks discontent, sudden remembrance, reproach, and annoyance (Rhee & Zhang 2024). Khammee (2024) identifies the following DM functions in the Thai *zin*: agreement response token, sudden realization/remembrance, surprise, confirmation/agreement solicitation, perspective shift, elaboration, emphasis, and disagreement response token.

In contrast, it appears that the Japanese *shinni* does not have DM functions similar to those of the Chinese *zhende* and *zhenshi*, the Korean *cinnca*, and the Thai *zin*, except for marking emphasis (but see Note 13 on the usage of *zhenshi*). In other words, as noted in Section 4, the "INTENSIVE" *shinni* carries the pragmatic function of intensifying the illocutionary force, although we only identified 10 examples of this specific function in our database.[6] For example, in (2) *shinni* can be interpreted as serving to intensify the illocutionary force of apology, meaning 'I am truly/very sorry', and in (7) it can be seen as serving to intensify the illocutionary force of regret, expressing 'I truly/deeply regret' or 'I am truly/very sorry'.

Let us look at some other examples of the "INTENSIVE" *shinni*. In (11a), *shinni* can be regarded as serving to intensify the illocutionary force of wish, meaning 'I truly/very much pray', while in (11b) it can be interpreted as intensifying the illocutionary force of gratitude, meaning 'I am truly/very honored'.

---

[6] Out of the 10 examples of *shinni* in this specific usage, 3 occur in essays written by authors who were born in the late nineteenth century (e.g., (7)), 1 in an essay written by an author born in the early twentieth century (11b), 1 in a historical novel written in 2002 (2), 1 in a story of a mythical creature written in 2003 (12a), 1 in a judicial decision text cited in a newspaper in 2002 (11a), and 3 in blogs written in 2008 (e.g., (12b)).

(11) Intensive adverb (intensifying the illocutionary force)

　a.　生活共同体として、あらたに復興することを**真に**祈念する

　　seikatsu.kyōdōtai-toshite,　aratani　　　fukkōsuru-koto-o
　　life.community-as　　　　　newly　　　restore-NML-ACC

　　shinni　　　　*kinensuru*
　　truly　　　　　pray

　　"As a community of life, (I) truly/very much pray (you would) newly restore it".
　　　　　　　　　　(BCCWJ: PN2d_00010, 13070, Sangyō Keizai Shimbunsha,
　　　　　　　　　　　　　　　　　　　　　　　　Sankei Shimbun, 2002)

　b.　**真に**光栄でございます。
　　shinni　　　　*kōē-degozaimasu.*
　　truly　　　　　honor-COP.POL
　　"(I) am truly/very honored."
　　　　　　　　　　　　(BCCWJ: PB30_00030, 74380, Umehara Takeshi,
　　　　　　　　　　　　　　　　Umehara Takeshi Chosakushū, 2003)

In (12), *shinni* can be seen as serving to intensify the illocutionary force of affirmation, meaning 'I truly/very much agree'. Example (12a) is taken from a story in which a Yōkai, a supernatural being from Japanese folklore and mythology, is the main character. In (12a), the speaker in the story expresses strong agreement with what was said in the preceding discourse. In (12b), the writer of the blog extends his/her strong agreement with the opinion s/he received from a reader.

(12) Intensive adverb (intensifying the illocutionary force)

　a.　**真に**その通り。
　　shinni　　　*sono-tōri.*
　　truly　　　so-just.like
　　(Thanks to those creatures with magical and psychic power who processes us, we, the fox family, are able to have strong power of transformation.) "(I) truly/very much agree."
　　　　　　　　　　(BCCWJ: PB39_00749, 72330, Kyōgoku Natsuhiko,
　　　　　　　　　　　　　　Tōfukozō Sugoroku Dōchū Furidashi, 2003)

　b.　**真に**その通りですよね。
　　shinni　　　*sono-tōri-desu-yo-ne.*
　　truly　　　so-just.like-COP.POL-FP-FP
　　(I have received the opinion that it is not right to ask people to vote for you while also telling them not to require them to register as fans of your blog.) "(I) truly/very much agree."
　　　　　　　　　　(BCCWJ: OY14_14139, 7550, *Yahoo! Blog*, 2008)

The remaining three examples of the "INTENSIVE" *shinni* can be counted as serving to intensify the illocutionary force of apology in a similar way to (2).

This specific pragmatic function of "INTENSIVE" *shinni* 'very', as well as the meaning of the adverb *shinni* 'truly', may be characterized as an increase in discourse orientation in semantic change and grammaticalization (Narrog and Heine 2021: 92–116). Narrog and Heine (2021: 92) point out that "in semantic change in general and in grammaticalization in particular, meanings tend to become increasingly oriented toward the speech act participants, that is, speaker and hearer, and toward organizing speech or discourse itself", as summarized in (13).

(13) I   increasing orientation toward the speaker's perspective (increased speaker orientation)

II   increasing orientation toward the speech situation including the hearer (increased hearer orientation)

III increasing orientation toward speech (text) and discourse itself

(Narrog & Heine 2021: 93)

Tendency I and Tendency II may be seen in the overall development of the adverb *shinni* and the "INTENSIVE" *shinni* from the Sino-noun *shin*, while Tendency III cannot be identified in this development.[7] We may see Tendency I (increased speaker orientation) in the later development of *shinni* than *shin*. In other words, the adverb *shinni* acquired the speaker-oriented meaning in that its use is dependent on the speaker's/writer's subjective judgment. Tendency II (increased hearer orientation) may be observable in the "INTENSIVE" *shinni* that serves to intensify the illocutionary force of apology that is directed to the hearer by the speaker (in conversation in novels as far as our database is concerned). On the other hand, the use of *shinni* as an intensifier seems to remain mostly within the range of the speaker's side, because the hyperbolic or expressive functions of intensifiers attest to the effect of the speaker's involvement in the immediate discourse or interaction (Athanasiadou 2007), and do not necessarily exert influence on the hearer's side, especially in written discourse.

5.3. Why has shinni *not become a discourse marker?*

Why, then, has *shinni* not become a DM like the corresponding forms in Chinese, Korean, and Thai? One reasonable explanation is that *shinni* has been specialized in written discourse, as clearly seen in Table 2. It is true that one of the spoken corpora, Corpus (f) DR (Diet Record), includes a relatively large number of examples; however, parliamentary statements during the Diet session appear to reflect formal tones close to written language.

---

[7] According to Narrog and Heine (2021: 99–100), Tendency III is "a distinctive tendency that cannot be subsumed under speaker- or hearer orientation", for example, the development of complex clause constructions out of independent speech acts, such as concessive conditionals out of imperatives. The development of the adverb *shinni* out of the Sino-noun *shin* does not involve Tendency III because *shinni* does not "serve to connect two or more propositions in discourse and indicate semantic relationships between them" (Narrog and Heine 2021: 100).

In addition, it is worth noting that each reading of a given Chinese character has its own stylistic role in Japanese discourse (Shibasaki 2008, see also Widdowson 1975: Chap 2, Farb 1993); the same can be applied to 真. For example, 真に can be read in at least the following three ways: *shin-ni*, *makoto-ni*, and *ma-ni* in the collocational phrase *ma-ni ukeru* [truth-as take 'take (something) seriously']. The first two are described as written language, while the last is described as everyday language (Ohno & Hamanishi 2012 s.v. *shin-ni*). It is well known that the spoken and written languages have developed in different ways (Miller 2011: Chap 11), and interestingly, both *shinni* and *ma-ni* are used in one text in the first half of the nineteenth century, as follows. (3b) is repeated here as (14), with the addition of the Chinese character 真 for *shinni*; *ma-ni* in (15) is recorded as the earliest example according to Nikkoku.

(14) Adverb (modifying the predicate)
　　　真に嬉しいヨ、どふぞ

　　　shin* (真) *ni　　*ureshii-yo,*　　　*dōzo*
　　　truly　　　　　be.happy-FP　　　please
　　　"(I) am truly/very happy, please."
　　　　　　　　　　　　　　　(Nikkoku, *Ninjōbon Shunshokuumegoyomi*, 1832–33)

(15) Collocation
　　　嬉しがらせを**真に**うけて、今日まで尽した心の操

　　　ureshi-gar-ase-o　　　　　***ma* (真)-*ni***　　*uket-te,*　　　*kyō-made*
　　　be.happy-V.AFF-CAUS-ACC　　**truth-as**　　　**take**-CVB　　today-until
　　　tsukushi-ta　　　*kokoro-no-misao*
　　　serve-PST　　　　heart-GEN-chastity
　　　"I have (ever) taken seriously (my partner, Tōbei's) kindness and love, and preserved (my) chastity up to now."
　　　　　　　　　　　　　　　(Nikkoku, *Ninjōbon Shunshokuumegoyomi*, 1832–33)

While the Chinese character 真 is used in one and the same text, the two instances differ from each other from both semantic and stylistic points of view. In (14), *shinni* modifies the following predicate *ureshii* 'be happy' and can be interpreted as meaning 'very', i.e., "INTENSIVE" (Kuteva et al. 2019, see Section 5.1 for relevant discussions). In (15), on the other hand, *ma-ni* collocates with *ukeru* 'to take' and this collocation means 'take (something) seriously, accept (something) as true': *ma-ni* does not serve as an intensifier. In other words, the collocational phrase retains its truth-related meaning rather than following the shift to an abstract meaning 'very' of intensification or to a DM.[8] Stylistically, the author of this

---

[8] The collocational phrase *shin-ni semaru* (truth-PT close.in) 'approach the heart (of the matter)' also retains the original nominal function 'truth' of *shin*. Note that *shin-ni shikari* (truth-PT as.it.is) 'that is correct' quoted in *Genkai* (Otsuki 1889 s.v. *shin-ni*) might have been another collocation in the late nineteenth century. Further, this specific expression was likely to be a clausal pragmatic marker 'certainly' to prompt language users to ask themselves something, especially in written discourse as

human-interest story, Tamenaga Shunsui, might have used *shinni* and *ma-ni* for different purposes, which would have become a springboard for a phraseology of written vs. everyday language uses in later stages. A BCCWJ search provides 107 examples of *mani ukeru*, 25 of which are used in online question–answer interactions and blogs in *Yahoo!* as below. This survey proves that *shinni* is skewed toward written language, as in Table 2, while *ma-ni ukeru* is used in a wider variety of discourse genres.

(16) Collocation
しかしあの大久保は何者ですか？あんなのの言うことなんて**真に受ける**必要なし。

shikashi	*ano-Ōkubo-wa*	*nani-mono-desu-ka?*
but	that-PN-TOP	what-person-COP.POL-QP

anna-no-no-iu-koto-nante	***ma-ni***	***ukeru***	*hitsuyōnashi.*
like.that-fellow-GEN-say-NML-EMPH	**truth-as**	take	unnecessary

"But who on earth is *Ōkubo*? (You) don't have to take seriously what he says."
(BCCWJ, OC06_05208, 1040, *Yahoo!Chiebukuro*, 2005)

(17) Collocation
あなたの思うことすべてを**真に受け**ますから、あなたの為になることを望んで下さい。

anata-no-omou-koto-subete-o	***ma-ni***	***uke-masu-kara,***
you-GEN-think-NML-all-ACC	**truth-as**	take-POL-because

anata-no-tame.ni.naru-koto-o	*nozon-de-kudasai.*
you-GEN-benefit.bring-NML-ACC	wish-CVB-please.do

"because (your subconsciousness) takes all you think about at face value, (please wish yourself) what turns out to be good for you."
(BCCWJ, OY14_08390, 2530, *Yahoo! Blog*, 2008)

What kinds of expressions have been used in everyday language instead of *shinni*? As explained above, *ma-ni* could be one of the candidates but it appears to be restricted mostly to the collocation *ma-ni ukeru* 'take (something) seriously'. Table 4 summarizes some typical adverbs meaning 'truly' based on Ohno & Hamanishi (2012) and Nikkoku. This table tells us that some truth-related words are used in a complementary style and that *shinni* assumes a role in written discourse.

---

below, although it sounds old-fashioned now. We will develop this point in our future study. Note that # means a sentence boundary, while the last comma is equivalent to a period.

(i) 見ろ、見ろ、 あれ 溝道を出たぞ# 眞に然り、

miro,	*miro,*	*are*		*kōdō-o*	*de-ta-zo.*	*shinni*	*shikari,*
look.IMP	look.IMP	that (=the French army)		trench-ACC	go.out-PST-FP	truly	be.correct

"Look out, the French army went out of the trench. No doubt (they are closing in)."
(CHJ: 60MTaiyo1901_14016, 311710, 1901)

Table 4. **Some 'truth'-related words in Japanese** (based on Ohno & Hamanishi 2012)

Form	Genre	Earliest use as adverb
masa-ni 'literally, very'	written	828 onward
jitsu-ni 'indeed, in faith'[9]	everyday	1231-53 onward
shin-ni 'truly'	written	1463–64 onward
masashiku 'truly'	written	1657 onward (according to CHJ)
hontō(-ni) (or *honto(ni)*) 'really'	everyday	1787 onward
ma-ni ukeru 'take (something) seriously'	everyday	1832–33 onward
makoto-ni 'truly'[10]	written	1887 onward
jissai 'in fact'	everyday	1896 onward

In addition, one can see a process of change: once a form begins to be used to some extent in everyday language, it is likely to extend its function to a DM, as illustrated below. In contemporary Japanese, *hontō(-ni)/honto(ni)* 'really, truly' is one of the most commonly used 'truth'-related words that can serve DM functions of expressing the speaker's surprise, assurance, agreement, etc., akin to the Chinese *zhende* and *zhenshi*, the Korean *cinnca*, and the Thai *zin*.[11] For example, in (18), *hontō* marks the speaker's surprise, knowing that their uncle said to "sister-in-law" that he felt relieved when his wife was hospitalized because he had been caring for her before her hospitalization.

(18) Wife: えー そんなことゆったの
                     ē   sonna-koto   yutta-no
                     INT   such-thing   said-FP
                     "Oh, (did he) say such things?"

      Sister-in-law: ゆったよ
                     yutta-yo
                     said-FP
                     "(He) said."

      Wife: 本当
                     hontō
                     really
                     "Really"

      Sister-in-law: ん いや
                     n   iya
                     uh   no
                     "Uh well"

                                              (CEJC: T013_009, 17980, 2017)

---

[9] The nominal part *jitsu* 'truth, faith' can be used as part of the collocational phrase *jitsu-o ieba* (truth-ACC say-if) 'to tell (you) the truth' (Hida & Asada 2018: 486).

[10] According to Nikkoku, *makoto-ni* can be written in several ways such as *makoto* (真)-*ni*, *makoto* (実)-*ni*, *makoto* (誠)-*ni*, *makoto* (洵)-*ni*, and *makoto* (亶)-*ni*. All of these forms mean 'truth'; the last two forms are archaic now. As we mentioned in Note 8, collocational phrases with these adverbial expressions deserve further investigation.

[11] Moriyama (2022) analyses the functions of *hontō(-ni)/honto(ni)* and *maij(-de)* as interjections (i.e., DMs in the current study) in contemporary Japanese based on questionnaires.

The historical development of *hontō* is from the noun *hontō* 'realness, truth', as in (19a), to the adverb *hontōni* 'really, truly', as in (19b), and then to the intensifier 'very', as in (19c) (Narumi 2015: 196).[12] This developmental pathway illustrates the shift from "TRUE" to "INTENSIVE" (Kuteva et al. 2019).

(19)   a.   さては狐ではねへ。**ほんとう**の北八か
             satewa   *kitsune-de-wa-nee.*   ***hontō*-*no-Kitahachi-ka***
             then   fox-COP.CVB-TOP-not   real-GEN-PN-QP
             "Then, (you are) not a fox. (Are you) real Kitahachi?"
                 (Jippenshaikku, *Tōkaidōchūhizakurige*, vol. 4, jō, 1802 (Narumi 2015: 196))

        b.   此のごろは**ほんとうに**、呂律が廻って來たぜ
             konogoro-wa   ***hontōni*,**   *roretsu.ga*   *mawat-te-ki-ta-ze*
             recently-TOP   really   clarity-NOM   flow-CVB-become-PST-FP
             "Recently, (you) have become able to speak clearly."
             (Shikiteisanba, *Ukiyoburo*, vol. 4, maki no ge, 1809–1813 (Narumi 2015: 196))

        c.   奥さんが、あの時は**ほんとに**呆れたと云つて、気が附いて僕にあやまる。
             okusan-ga,   *ano-toki-wa*   ***hontoni***   *akire-ta-to-it-te,*
             woman-NOM   that-time-TOP   really   astonish-PST-QUOT-say-CVB
             ki-ga   *tsui-te*   *boku-ni*   *ayamaru.*
             attention-NOM   notice-CVB   I-to   apologize
             "The woman, saying that (she) was really astonished at that time, realized and apologized to me."
                                         (Mori Ogai, *Uwita Sekusuarisu*, 1909 (Narumi 2015: 196))

To sum up, a variety of words indicating 'truly' in Japanese have been in competition with each other but have gradually come to be distributed in a complementary manner. As for *shinni*, it has been skewed toward written language over time, while some other words such as *jissai* 'in truth', *jitsu-ni* 'indeed', and *hontō(-ni)* (or *honto(ni)*) 'really' have found their way into everyday language. Taking into consideration the fact that the latter three words are used as DMs while *shinni* is not, albeit in the majority of cases, what matters is whether a given word is used as an adverb or as an intensifier on a routine basis in spoken or everyday language. This is not a hard-and-fast rule but a diachronic process that is observed in many cases of the words in focus. One may thus get a polite and formal feeling if one hears the use of *shinni*, *masani*, *masashiku*, and *makotoni*, in conversation at all. This will be a plausible reason why *shinni* has not fully reached the DM stage.

---

[12] Narumi (2015: 201) notes that the etymology and origin of *hontō* is unknown.

5.4. Theoretical implications for cyclicity phenomena

Finally, let us touch on one theoretical issue in the development of DMs across languages. Hansen (2018a, b) proposes that words derived from the same source or carrying the same meaning tend to follow similar historical pathways into DMs, i.e., cyclicity. This proposal provides an explanation with respect to Romance languages and sounds intuitively convincing. That said, there is still some room for a reconsideration on the development of DMs in East Asian languages and Thai. As discussed in Section 5.2, the Chinese *zhende* and *zhenshi*, the Korean *cinnca*, and the Thai *zin*, all of these are historically related to the Japanese *shinni* and are used as DMs (Khammee 2024, Rhee & Zhang 2024[13], see papers in Higashiizumi and Shibasaki (in preparation) for comparison). In this respect, cyclicity works well other than for *shinni*, except for cases of some significantly different functions of particular DMs (Note 13). Does this mean that *shinni* is an exception to the idea of cyclicity? If so, why does it differ from these cognate-like expressions in the other languages?

One reason is that in contemporary Japanese, *shinni* is specialized in written language, while some of the relevant expressions are prevalent in everyday language. The other reason is that expressions used adverbially and repeatedly in everyday language have a tendency to become DMs. As summarized in Table 2, *shinni* shows a strong tendency toward use in written and formal contexts and in fact has not yet reached the DM stage (Section 5.3).

Here we are not intending to argue against Hansen's (2018a, b) hypothesis, but we are focusing on the degree of this specific grammaticalization pathway from "TRUE" to "INTENSIVE" through an analysis of some limited sets of cognate-like examples in an East Asian context. Obviously, Japanese was in close contact with Chinese over a lengthy period, and *shin* 'truth' was borrowed from Chinese as a noun in the early seventh century; it was also used as a nominal adjective, i.e., nominal *shin* plus a copula, which is categorized under the heading of noun (Nikkoku). In the fifteenth century, the adverbializer *-ni* was added to *shin* to serve as the adverb *shinni* 'truly', and *shinni* has been used since then up to the present (Tables 2 and 3). In a nutshell, *shin* has a long history of intensive contact with Chinese, which we believe played a foundational role in the rise of the intensifier meaning 'truly' as a piece of evidence for Kuteva et al. (2019: 443) and Hansen (2018a, b). However, unlike its counterparts in Chinese, Korean, and Thai, *shinni* has not (yet) advanced to the stage of DM, or it is very slow (or conservative) in undergoing semantic-pragmatic extensions to a DM, presumably due to the reasons mentioned above. If *shinni* comes to fulfill a DM function in the future, it would back up the assumption of cyclicity. Only time will tell.

---

[13] According to Haiping Long, the Chinese *zhenshi* differs significantly from *zhende* in terms of its pragmatic function, presumably due to the unknown origin of *shi*. We are grateful to him for this invaluable comment.

6. Conclusion

We have traced the developmental pathway of the Sino-Japanese noun *shin* 'truth' via the adverb *shinni* 'truly' to *shinni* that can be interpreted as serving to intensify the illocutionary force of what is being said. This development is in line with the general pathway of change labeled "TRUE" to "INTENSIVE" (Kuteva et al. 2019: 443). It is reported that 'truth'-related words with the same etymon 眞/真 'truth' in Chinese, Korean, and Thai have developed their own expressions from "TRUE" to "INTENSIVE". Thus, it appears that the extension of the etymon 真/眞 from "TRUE" to "INTENSIVE" is one of the commonalities among these four Asian languages as well. On the other hand, in Chinese, Korean, and Thai, these 'truth'-related words have developed further into DMs, while the Sino-Japanese adverb *shinni* has not taken on any DM function, as of the present: only some examples of *shinni* can barely (but not fully) be seen as serving to intensify an illocutionary force. The extent to which each expression has fully acquired its own DM function differs at least between Japanese and the other three languages (Chinese, Korean, and Thai). The idea of cyclicity (Hansen 2018a, b) would predict that *shinni* would become a DM sometime in the future, but we have to continue our research to test the hypothesis from an East Asian perspective. This study raises the possibility that, due to various factors, the pace of change at advanced stages of grammaticalization may be more diversified than has previously been suggested.

Abbreviations

ACC: accusative; ADV: adverbial; CAUS: causative; CONJ: conjectural; COP: copula; CVB: converb; EMPH: emphatic; FP: final particle; GEN: genitive; HUM: humble; IMP: imperative; INT: interjection; NML: nominalizer; NOM: nominative; PN: personal name; POL: polite; PST: past; PT: particle; QP: question particle; QUOT: quotative; RES: respective; TOP: topic; V.AFF: verb affix.

References

Athanasiadou, Angeliki. 2007. On the subjectivity of intensifiers. *Language Sciences* 29 (4). 554–565.

Diewald, Gabriele. 2002. A model for relevant types of contexts in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections of grammaticalization*, 103–120. Amsterdam & Philadelphia: John Benjamins.

Farb, Peter. 1993. *Word Play: What Happens When People Talk*. New York: Vintage Books.

Fujimura, Itsuko, Shuji Chiba & Mieko Ohso. 2012. Lexical and grammatical features of spoken and written Japanese in contrast: Exploring a lexical profiling approach to comparing spoken and written corpora. *Seventh GSCP International Conference: Speech and Corpora*. 393–398.

Gendai, Nihongo Kenkyūkai. 2011. *Gappon Josei no Kotoba·Dansei no Kotaba: Shokuba-hen (A Bound set of Female Language and Male Language: At workplace)*. Tokyo: Hituzi Syobo.

Hansen, Maj-Britt Mosegaard. 2018a. Cyclic phenomena in the evolution of pragmatic markers: Examples from Romance. In Salvador Pons Bordería & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers: New issues in the study of language change*, 51–77. Leiden: Brill.

Hansen, Maj-Britt Mosegaard. 2018b. The role of inferencing in semantic/pragmatic cyclicity: The case of Lain *nunc* and French *or/maintenant*. *Open Linguistics* 4. 127–146.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections of grammaticalization*, 83–101. Amsterdam & Philadelphia: John Benjamins.

Higashiizumi, Yuko & Reijirou Shibasaki (eds.). in preparation. *The Emergence of Pragmatic Markers from Chinese Compounds in Chinese, Japanese and Korean: Perspectives from East Asian Languages and Beyond*. Leiden: Brill.

Khammee, Kultida. 2024. From objective to subjective, and to intersubjective functions: The case of the Thai 'truth'-lexeme. *Russian Journal of Linguistics* 28 (4). 942–965. https://doi.org/10.22363/2687-0088-40496

Koiso, Hanae, Haruka Amatani, Yasuharu Den, Yuriko Iseki, Yuichi Ishimoto, Wakako Kashino, Yoshiko Kawabata, Ken'ya Nishikawa, Yayoi Tanaka, Yuka Watanabe & Yasuyuki Usuda. 2022. Design and evaluation of the Corpus of Everyday Japanese Conversation. *LREC [Language Resources and Evaluation]*. 5587–5594.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World Lexicon of Grammaticalization*. 2nd edn. Cambridge & New York: Cambridge University Press.

Maruyama, Takehiko. 2020. On the possibility of a diachronic speech corpus of Japanese. In Andrej Bekeš & Irena Srdanović (eds.), *Japanese language from empirical perspective: Corpus-based studies and studies on discourse*, 219–234. Ljubljana: Znanstvena založba FF.

Miller, Jim. 2011. *A Critical Introduction to Syntax.* New York: Bloomsbury Publishing.

Moriyama, Takuro. 2022. Kandōshi to shite no *maji* to *honto* (*Maji* and *honto* as interjections). In Kenji Tomosada (ed.), *Kandōshi Kenkyū no Tenkai (Expansion of research concerning Interjections)*, 43–61. Tokyo: Hituzi Syobo.

Narrog, Heiko & Bernd Heine. 2021. *Grammaticalization*. Oxford: Oxford University Press.

Narumi, Shin'ichi. 2015. *Studies on the Japanization of Chinese Loanwords: Focusing on Adverbialization*. Tokyo: Hituzi Syobo. (In Japanese).

Ogiso, Toshinobu, Asuko Kondo, Yuta Takahashi, Makiro Takanaka & Yoko Mabuchi (eds.). 2023. *Showa-Heisei Corpus of Written Japanese*. National Institute for Japanese Language and Linguistics.

Ogiso, Toshinobu, Asuko Kondo, Yuta Takahashi & Yoko Mabuchi (eds.). 2024. Design, construction and publication of the *Showa-Heisei Corpus of Written Japanese*. *Information Processing Society of Japan* 65 (2). 278–291.

Rhee, Seongha & Lin Zhang. 2024. The way of truth: The case of the Korean discourse marker *cincca* in comparison with Chinese discourse markers *zhenshi* and *zhende*. *Russian Journal of Linguistics* 28 (4). 818–842. https://doi.org/10.22363/2687-0088-40500

Shibasaki, Reijirou. 2008. Two levels of blending with homophonic compounds in Japanese. *Cultura, Lenguaje y Representación* 5. 221–239. [Special issue: Metaphor and discourse.]

Widdowson, Henry G. 1975. *Stylistics and the Teaching of Literature*. London: Longman.

Dictionaries

Daijirin = *Daijirin*. 4th edn. 2019. Tokyo: Sanseido.

Hida, Yoshifumi & Hideko Asada. 2018. *Gendai Fukushi Yōhō Jiten* (A Usage Guide to Adverbs in Contemporary Japanese). new edn. Tokyo: Tokyodo Shoten.

Iwanami = *Iwanami Kokugo Jiten* (Iwanami Dictionary of the Japanese Language. 8th edn. 2019. Tokyo: Iwanami Shoten.

Kadokawa = *Kadokawa Kogo Daijiten* (Kadokawa Large Dictionary of Classical Japanese), JapanKnowledge Lib, 1982–1999. Tokyo: Kadokawa. https://japanknowledge.com/library/.

Morohashi, Tetsuji. 2018. *Dai Kanwa Jiten* (Large Dictionary of Chinese Characters), digital version. Tokyo: Taishukan Shoten.

Nikkoku = *Nihon Kokugo Daijiten* (The Dictionary of the Japanese Language), JapanKnowledge Lib. 2006. Tokyo: Shogakukan. https://japanknowledge.com/library/.

Ohno, Susumu & Masato Hamanishi (eds.). 2012. *Kadokawa Ruigo Shin Jiten* (A New Kadokawa Thesaurus). https://japanknowledge.com/library/.

Otsuki, Fumihiko. 1889. *Nihonjisho Genkai* (*A Japanese Dictionary Genkai*). Privately published. Reprinted, Tokyo: Chikuma-shobo, 2004.

Todo, Akiyasu (ed.). 1978. *Gakken Kanwa Daijiten* (Gakken Dictionary of Chinese Characters). Tokyo: Gakken.

Appendix: Corpora used for this study

	Abbreviation, Database and corpora, URL, date accessed	Period	Chunagon version	Data version	Total number of words in the corpus
(a)	SZ=*Shimpen Nihon Koten Bungaku Zenshū* (The Complete Works of Classical Japanese Literature. new edn.). https://japanknowledge.com/library/ (accessed 1 July 2024)	late 7th c. – late 19th c.	N/A	N/A	Unknown
(b)	CHJ=*The Corpus of Historical Japanese*. http://clrd.ninjal.ac.jp/chj/ (accessed 1 July 2024)	8th c. – early 20th.c	2.7.1	2023.03	18,550,049
(c)	SHC=*Showa-Heisei Corpus of Written Japanese*. (Ogiso et al. (eds.) 2023; Ogiso et al. (eds.) 2024). https://clrd.ninjal.ac.jp/shc (accessed 1 July 2024)	1933–2013	2.7.2	2023.05	33,404,844

	Abbreviation, Database and corpora, URL, date accessed	Period	Chunagon version	Data version	Total number of words in the corpus
(d)	BCCWJ=*Balanced Corpus of Contemporary Written Japanese,* excluding the sub-corpus *Diet Record* (Corpus (f) DR). https://clrd.ninjal.ac.jp/bccwj/index.html (accessed 1 July 2024)	1971–2008	2.6.0	2021.03	99,808,991
(e)	SSC=*Showa Speech Corpus* (Maruyama 2020). https://www2.ninjal.ac.jp/conversation/showaCorpus/ (accessed 1 July 2024)	1952–1969	2.7.1	2022.02	528,589
(f)	DR=the sub-corpus *Diet Record* in Corpus (d) BCCWJ. https://clrd.ninjal.ac.jp/bccwj/index.html (accessed 1 July 2024)	1999–2001	2.6.0	2021.03	5102469
(g)	CSJ=*Corpus of Spontaneous Japanese.* https://clrd.ninjal.ac.jp/csj/index.html (accessed 1 July 2024)	1971–2008	2.7.1	2018.01	7,576,046
(h)	CEJC=*The Corpus of Everyday Japanese Conversation* (Koiso et al. 2022). https://www2.ninjal.ac.jp/conversation/cejc.html (accessed 1 July 2024)	2016–2019	2.7.2	2023.03	2,419,171
(i)	NUCC=*Nagoya University Conversation Corpus* (Fujimura et al. 2012). https://www2.ninjal.ac.jp/conversation/nuc.html (accessed 1 July 2024)	1993–2000	2.7.2	2020.1	1,135,329
(j)	CWPC=*Gen-Nichi-Ken Corpus of Workplace Conversation* (Gendai Nihongo Kenkyūkai (ed.). 2011). https://www2.ninjal.ac.jp/conversation/shokuba.html (accessed 1 July 2024)	1993 & 1999–2000	2.7.2	2018.03	186,906

Bionotes:

Yuko HIGASHIIZUMI is a Visiting Research Fellow of the Institute of Human Sciences at Toyo University, Tokyo, Japan. She received her PhD in English linguistics from Dokkyo University, Japan, in 2004. Her primary research interest is Japanese linguistics and historical pragmatics. She published research articles in *Journal of Historical Pragmatics* and *East Asian Pragmatics*, among others.
e-mail: higashiizumi.yuko.1@gmail.com
https://orcid.org/0009-0005-0086-572X

Reijirou SHIBASAKI is Professor of English at the School of Interdisciplinary Mathematical Sciences, Meiji University, Tokyo, Japan. He received his PhD in Linguistics from University of California at Santa Barbara. His research focuses on historical linguistics and discourse analysis, especially in English and Japanese. His recent research

articles are included in, inter alia, *Studies at the Grammar-Discourse Interface* (2021, John Benjamins), *Different Slants of Grammaticalization* (2023, John Benjamins), *Journal of Japanese Linguistics* (2023, De Gruyter Mouton), and *Gengo Kenkyu* (2024, Journal of the Linguistic Society of Japan).
e-mail: reijiro@meiji.ac.jp
https://orcid.org/0009-0006-2268-1486

Keiko TAKAHASHI is a Visiting Research Fellow of the Institute of Human Sciences at Toyo University, Tokyo, Japan. She received her degree of M.A. (Language and Information Sciences) from Tokyo University, Japan, in 2000. Her primary research interest is Japanese linguistics and historical sociopragmatics. She published research articles in the *Japanese Journal of Language in Society*, *Journal of Japanese Language Teaching*, *Academic Japanese Journal*, among others.
e-mail: ktakahashi@toyo.jp
https://orcid.org/0009-0007-8370-0043

Сведения об авторах:

Юко ХИГАШИИДЗУМИ – приглашенный научный сотрудник Института гуманитарных наук в Университете Тойо, Токио, Япония. Она получила степень PhD по английской лингвистике в Университете Доккё, Япония, в 2004 году. Основные научные интересы – японская лингвистика и историческая прагматика. Автор научных статей, опубликованных в *Journal of Historical Pragmatics, East Asian Pragmatics* и др.
e-mail: higashiizumi.yuko.1@gmail.com
https://orcid.org/0009-0005-0086-572X

Рейджиро СИБАСАКИ – профессор английского языка, преподает в Школе междисциплинарных математических наук в Университете Мэйдзи, Токио, Япония. Получил степень PhD по лингвистике в Калифорнийском университете в Санта-Барбаре, США. Его исследования сосредоточены на исторической лингвистике и анализе дискурса, особенно на английском и японском языках. Его последние статьи опубликованы в *Studies at the Grammar-Discourse Interface* (2021, John Benjamins), *Different Slants of Grammaticalization* (2023, John Benjamins), *Journal of Japanese Linguistics* (2023, De Gruyter Mouton) и *Gengo Kenkyu* (2024, Журнал лингвистического общества Японии).
e-mail: reijiro@meiji.ac.jp
https://orcid.org/0009-0006-2268-1486

Кейко ТАКАХАСИ – приглашенный научный сотрудник Института гуманитарных наук в Университете Тойо, Токио, Япония. Она получила степень магистра (языковые и информационные науки) в Токийском университете, в 2000 г. Основные области исследования – японская лингвистика и историческая социопрагматика. Автор статей, опубликованных в журналах *Japanese Journal of Language in Society*, *Journal of Japanese Language Teaching*, *Academic Japanese Journal* и др.
e-mail: ktakahashi@toyo.jp
https://orcid.org/0009-0007-8370-0043

The evolution of pragmatic marker *zenzen* in Japanese: From objectivity to intersubjectivity

Jiyeon PARK 🔟 ✉

Matsuyama University, Ehime, Japan
✉j.park@g.matsuyamau.ac.jp

Abstract

This paper investigates the semantic-pragmatic functions of the Japanese adverb *zenzen* (全然) 'completely, entirely, not at all' highlighting its semantic and functional development from its initial borrowing from the Chinese etymon to its contemporary use. The goal of the study is to investigate how *zenzen* is used in spoken discourse and what functions it has from the perspective of cooptation, subjectification, and intersubjectification. The paper traces the historical trajectory of *zenzen* from the early Meiji Period to the present, analyzing quantitative data from six corpora. A total of 2,154 examples were analysed. The results showed that in contemporary Japanese, *zenzen* serves to accentuate a state of perfection and reassurance, reflecting the speaker's attitude toward the interlocutor. The paper argues that *zenzen* has evolved as a pragmatic marker, indicating the speaker's epistemic stance and viewpoint. The paper attempts to explain how *zenzen* has transitioned from a lexical item with objective meanings to a pragmatic marker with (inter)subjective functions. The findings of the paper indicate that, *zenzen* functions as an adjectival noun combined with copula -*da/-desu* 'to be', creating a new unique construction [*zenzen-da/desu*]. These constructions play a role in expressing (inter)subjective meanings. The findings of the paper will prove useful in expanding our understanding of how diachronic language changes occur from the perspectives of cooptation, subjectification, and intersubjectification.

Keywords: *pragmatic marker, cooption, subjectification, intersubjectification, spoken discourse, Japanese*

Эволюция прагматического маркера *zenzen* в японском языке: от объективности к интерсубъективности

Джиён ПАРК ⓘ ✉

Университет Мацуяма, Эхимэ, Япония
✉j.park@g.matsuyamau.ac.jp

Аннотация

В данной статье исследуются семантико-прагматические функции японского наречия *zenzen* (全然) «полностью, целиком, совсем нет», прослеживается его семантическое и функциональное развитие от первоначального заимствования из китайского этимона до современного использования. Целью исследования является изучение того, как *zenzen* используется в устной речи, выявление его функций с точки зрения кооптации, субъективации и интерсубъективации. На основе количественного анализа данных из шести корпусов в статье прослеживается историческая траектория *zenzen,* начиная с раннего периода Мэйдзи и до наших дней. Всего было проанализировано 2154 употребления. Результаты показали, что в современном японском языке *zenzen* служит для подчеркивания совершенства и уверенности, что отражает отношение говорящего к собеседнику. В статье показано, что *zenzen* развился как прагматический маркер, указывающий на эпистемическую позицию и точку зрения говорящего. Делается попытка объяснить, как *zenzen* перешел из лексической единицы с объективными значениями в прагматический маркер с (меж)субъективными функциями. Результаты исследования показывают, что *zenzen* может функционировать как адъективное существительное в сочетании со связкой *-da/-desu* «быть», создавая новую уникальную конструкцию [*zenzen-da/desu*]. Эти конструкции играют определенную роль в выражении (меж)субъективных значений. Результаты исследования расширяют понимание того, как происходят диахронические изменения языка с точки зрения кооптации, субъективации и интерсубъективации.

Ключевые слова: *прагматический маркер, кооптация, субъективация, интерсубъективация, устный дискурс, японский язык*

1. Introduction

Within the Japanese language lexicon, a significant portion comprises Sino-Japanese words of Chinese origin, accounting for 40 to 50 percent (Mizutani et al. 2017: 60–61), making the research of these words, such as *zenzen*（全然）, crucial for understanding diachronic language change and development. The Japanese adverb *zenzen* stands out, having gained widespread use in both written and spoken registers. *Zenzen*, literally 'completely thus', is used in affirmative and negative sentences, representing 'completely, totally, entirely' or 'not at all', respectively. The word was initially borrowed from the Chinese noun *quánrán* (全然), meaning 'completely, totally, entirely' in the early 1800s from *báihuà xiǎoshuō* 'vernacular

fiction' which was written in colloquial texts. Initially, the word *zenzen* was often accompanied by *furigana*[1], indicating a reading of *mattaku*, meaning 'completely' (*Nihon Kokugo Daijiten* 'Unabridged Japanese Dictionary' 1988). *Zenzen* gained popularity in the early 1900s (Sano 2012) and has since undergone significant changes in its meaning and usage over time. In contemporary Japanese, *zenzen* is usually used with verbs or adjectives that include plain negative forms such as *-nai* 'not'(e.g., *Kare-wa* **zenzen** *benkyoo shi-nai* "He does not study *at all*") or co-occur with words that connote negativity.

To date, extensive research has been conducted on the meaning of *zenzen*, particularly lexical cooccurrence and historical semantic changes in Japanese linguistics. However, there has been insufficient focus on works evaluating the role of the semantic-pragmatic functions of *zenzen*. Indeed, earlier works have tended to be based on written texts such as literary works, and far too little attention has been paid to the use of *zenzen* in conversational contexts. For instance, in contemporary Japanese, *zenzen* is frequently found in the form of holophrases and is used as a response in spoken discourse (e.g., A: *Kore-wa oishii?* 'Is this delicious?' B: *Zenzen* '[No], *not at all*'). Furthermore, since the 2000s, a unique construction combined with copula *-da/desu* 'to be' has emerged (e.g., A: *Renraku-ga okure-te gomen-nasai* 'I am sorry for the delay in contacting you' B: *Zenzen-desu* '*It is totally fine to me. Do not concern about it*').

This paper aims to investigate the use of *zenzen* in spoken discourse, specifically through the lenses of cooptation, subjectification, and intersubjectification (Heine et al. 2021, Traugott 2003). It focuses on the pragmatic functions and semantic features of *zenzen*, utilizing a range of historical and contemporary corpora for analysis. This paper argues that *zenzen* has evolved from an objective lexical item to a pragmatic marker (PM, hereafter) to represent a state of perfection and reassurance, reflecting the speaker's attitude toward the interlocutor.

The paper is organized as follows: Section 2 summarizes previous works on *zenzen*, Section 3 explains the methodology, and Section 4 reports the results from quantitative analysis based on corpora data. Section 5 addresses the historical changes of *zenzen* in its semantic and syntactic shifts. Section 6 discusses the semantic and pragmatic functions, focusing on the adjectival noun construction [*zenzen-da/desu*] in spoken discourse, and indicates the development of *zenzen* as a PM. Finally, Section 7 concludes the paper.

---

[1] Furigana (振り仮名) in Japanese, is a reading aid that helps to indicate the pronunciation of Chinese characters. It consists of smaller kana, either hiragana or katakana. In horizontal writing, furigana is typically placed above the kanji, while in vertical writing, it is placed to the right of the kanji.

2. Background

2.1. Previous studies

As mentioned in Section 1, it is known that Japanese *zenzen* was initially borrowed from the Chinese noun *quánrán* (全然), meaning 'completely, totally, entirely' in the early 1800s from *báihuà xiǎoshuō* 'vernacular fiction' (*Nihon Kokugo Daijiten* 'Unabridged Japanese Dictionary' 1988). In the originated source, the Chinese *quánrán* mostly co-occurred with negative expressions[2]. When the word was first adopted into the Japanese language, it was often accompanied by furigana indicating readings such as *suppari* 'completely' (Hashimoto 2010), *marude* 'entirely, wholly', *mattaku* 'completely, thoroughly, not at all' and *sukkari* 'completely' (Niino 2011:117). Those readings continued through the early and middle Meiji Period[3] and were often found in some literary works. The reading of *zenzen* became more widespread starting in the late 1890s (*Nihon Kokugo Daijiten* 'Unabridged Japanese Dictionary' 1988). Initially, *zenzen* was used in sentences with no apparent restrictions on its collocation within sentences. Over time, however, the meaning of *zenzen* and its usage with collocational elements in a sentence has changed.

Numerous studies in Japanese linguistics have investigated the diachronic changes in the meaning of *zenzen*. Notably, Niino (2011) categorizes sentences containing *zenzen* into six types based on their collocational elements. These types are divided mainly into two groups: [*zenzen* + affirmative[4]] and [*zenzen* + negative] groups. The form [*zenzen* + verb/adjective-*nai* 'not'] is categorized as the [*zenzen* + negative] group. Other forms are categorized as [*zenzen* + affirmative] groups and further subdivided into five groups: (i) Sentences including Sino-Japanese nouns combined with prefixes or suffixes indicating negativity, such as *mu-* 'none' or *fu-* 'not' (e.g., *Zenzen **fu**-sansei-de aru* '[It is] totally *disapproval*', Niino 2011: 118); (ii) Sentences with words representing a contrast or difference between more than two things (e.g., *Zenzen **kotonaru*** '[It is] completely *different*', Niino 2011: 119); (iii) Sentences with words indicating negative meanings (e.g., *Zenzen **uchikowasu*** '[It] totally *demolishes*', Niino 2011: 120); (iv) Sentences with words indicating the speaker's negative evaluation (e.g., *Watashi-ga zenzen **yoochi**-na atama-de*, I have a totally *childish* mind', Niino 2011: 120), which should be distinct from (iii). Niino (2011: 121) notes that although the words classified under (iii) represent negative meanings, they do not involve the speaker's evaluation and instead denote a continued negative situation; (v) Sentences including words without negative

meanings and represent more neutrality (e.g., *Zenzen **dooitsu**-de aru* '[It is] totally *identical*', Niino 2011: 122).

Extensive research has been conducted on classifying the meanings of *zenzen* based on the words it is often collocated with. However, there is much less knowledge about how *zenzen* is used in the context rather than just within a single phrase or sentence. Furthermore, in spoken discourse, functioning as holophrases (i.e., *Zenzen*) and adjectival noun (i.e., *Zenzen-desu*), typically used as a response, are not fully understood[5].

2.2. Some controversial issues on the use of zenzen

There have been several controversial arguments on the usage of *zenzen*. Firstly, it has been believed that the original meaning of *zenzen* is 'not at all', which means *zenzen* should be used to represent negative statements. Thus, its use of affirmative expressions has been criticized as erroneous. However, research in Japanese linguistics has rediscovered that affirmative and negative expressions coexisted when *zenzen* was initially borrowed in Japanese (Matsui 1977, Koike 2001, Sano 2012). This argument will also be proven true from my observation in Sections 4 and 5 of the current paper.

Secondly, using *zenzen* as an intensifier of the speaker's stance (e.g., *Kore-wa **zenzen** oishii* 'It is *extremely* delicious') has been criticized as a vulgar expression (Endo 1994, Endo & Yabe 1995)[6]. This usage was already observed in the late 1940s in a spoken register. Niino (2011, 2020) points out that this issue might be due to a lack of detailed examination of everyday language. In addition, Niino (2011: 134) argues that almost all these kinds of usage could be interpreted as 'completely, entirely', and should not be treated as adverbs of degree.

This paper fundamentally aligns with Niino's perspective, which argues that 'completely, entirely' is considered emphatic. In the current paper, however, its intensified meaning is not the same as the meaning of *zenzen* in affirmative sentences. Instead, the paper suggests that the meanings of 'completely, entirely' remain and add intensified meanings by the speaker's stance toward an event or hearer. In other words, this paper argues that *zenzen* retains its meaning of 'completely' and has evolved to represent a (inter)subjective stance-marking and functions as a PM through internal language change. Furthermore, the current paper will focus on examples where *zenzen* is used as a phrasal (i.e., an adjectival noun) as a response in spoken discourse and argues that some of them play a role in indicating (inter)subjective functions.

---

[5] Sano (2012) observes using *zenzen* in a spoken register using Corpus of Spontaneous Japanese (CSJ) offering by NINJAL (https://clrd.ninjal.ac.jp/csj/). However, the corpus data is limited to speeches at conferences and mock lectures, resulting in a more monologue-style rather than dialogue-style conversation.

[6] According to a recently updated version (April 2024 updated) of dictionaries, this meaning (i.e., intensifier) is defined as "slang" (*Dejitaru Daijirin* 'Digital Daijirin', Tokyo: Shōgakukan).

2.3. Pragmatic marker

In this paper, the term "pragmatic marker (PM)" is used to describe the function of *zenzen*. There is little consensus regarding the definitional and terminology issues of this marker. So far, many scholars have proposed definitions and tried to explain their functions (Fraser 1990, 2009, Östman 1981, Redeker 1991, Schiffrin 1987, and so on). In their works, several terms have been used to indicate the marker. For example, the term pragmatic marker was used by Brinton (1996). The other well-known terms are discourse marker (DM) (Jucker & Ziv 1998, Schiffrin 1987), discourse particle (Aijmer 2002, Hansen 1998), and pragmatic particle (Östman 1995). Fraser (2009) claims that discourse markers are a sub-type of pragmatic markers and classified pragmatic markers into four types: basic pragmatic markers (BPMs, e.g., *I promise*, *Please*, *My complaint*, Fraser 2009: 295), commentary pragmatic markers (CPMs, e.g., *Fortunately*, *Bluntly speaking*, *Certainly*, *Reportedly*, *Sir*, Fraser 2009: 296), discourse markers (DMs, e.g., *On the contrary*, *Anyway*, *So*, Fraser 2009: 296), and discourse structure markers (DSMs, e.g., *In summary*, *Now*, Fraser 2009: 297).

In the current paper, however, the terms PM and DM are treated as independent and separate: PM has "expressive functions", both subjective (e.g., expressing evaluation, emphasis, focusing on the speaker) and interpersonal (e.g., evoking the hearer's attention, expressing common knowledge, denoting negative or positive politeness, Brinton 1996: 36-40), while DM is a typically signal a relation between the discourse segment which hosts them and the prior discourse segment, perhaps produced by another speaker (Fraser 2009: 296, Traugott 2018: 27, e.g., Mark, a good guy. *On the contrary*, he's a jerk, Fraser 2009: 296).

The term PM is employed in the paper to help us understand the semantic-pragmatic function of *zenzen*. As it will be discussed in the paper, *zenzen* serves "expressive functions" but does not have a function of "signaling some relationship between clauses, utterances, units of talk, or discourse segments" (Heine et al. 2021: 9).

2.4. Subjectification and Intersubjectification

The present paper employs the concepts of subjectification and intersubjectification to help understand how *zenzen* has developed into a PM, particularly in the construction [*zenzen-da/desu*] in Sections 5 and 6. Subjectification is the process through which linguistic elements gradually come to encode the speaker's internal states and viewpoints over time. This concept, originally introduced by Benveniste (1971), Lyons (1982), and Langacker (1990), underscores the role of linguistic elements in expressing the speaker's attitudes and beliefs. Traugott (1989) defines subjectification as the process of evolving meanings to reflect the speaker's subjective beliefs or attitudes toward what is being communicated.

Intersubjectification builds on the concept of subjectification by highlighting the speaker's focus on the interlocutor's perspectives and social interaction needs.

According to Traugott (2010), intersubjectification involves the evolution of functions to express the speaker's attention to the interlocutor's self in both epistemic and social senses. Intersubjectification develops from subjectification, using its mechanisms to meet the communicative needs of the interlocutor. Traugott (2010) emphasizes the inherent connection between intersubjectification and subjectification, indicating that the former cannot occur without some degree of the latter. In short, while subjectification focuses more on the speaker, intersubjectification shifts this focus to the interlocutor.

Zenzen describes a speaker's perspective toward the interlocutor, and further enhancing interaction with the interlocutor could be considered suggestive of cooptation (Heinei et al. 2022). In the next section, the concept of cooptation is introduced.

2.5. Cooptation

Cooptation is defined as a cognitive-communicative operation which enables speakersto switch their perspective from the level of reasoning anchored in the meaning of sentences to a meta-level of reasoning immediately anchored in the situation of discourse (Heinei et al. 2021: 67). Grammatical effects typically associated with cooptation listed as (1).

(1) Common effects of cooptation (Heine et al. 2021: 68)
  a Meaning: From meaning as part of the sentence to meaning outside the sentence.
  b Function: From sentence-structuring function to metatextual function.
  c Syntax: From syntactic constituent of the sentence to syntactically unattached status.
  d Prosody: From prosodically integrated to unintegrated or less integrated status.
  e Semantic-pragmatic scope: From more restricted to wider scope.
  f Placement: From positionally constrained to less constrained placement.

In the original meanings, *zenzen* expresses the meanings 'completely, absolutely' or 'not at all', which can be classified as objective meanings. However, over time, *zenzen* further involves a speaker's positive attitude towards an interlocutor in discourse. In other words, *zenzen* uses a wider perspective by commenting on the hearer, and its function is determined by the situation of discourse. Moreover, in contemporary Japanese, *zenzen* functions as an adjectival noun combined with copula *-da/-desu* 'to be' and syntactically unattached status in a sentence, creating a new unique construction [*zenzen-da/desu*] in spoken discourse (we will discuss this construction in detail in Section 5 and 6).

3. Methodology

This paper uses qualitative and quantitative analyses to trace the development of *zenzen* using data extracted from six corpora, as listed in (2)[7]. Additionally, some

---

[7] Refer to the detailed information about each corpus in the References.

widely used Japanese dictionaries aid in this analysis. All corpora are available in the online search engine Chuunagon[8] and open to the public.

(2) Corpora Sources
   a *The Corpus of Historical Japanese* (CHJ, from Nara to Taisho Period): The online search engine Chuunagon, contains approximately 17.6 million words.
   b *Showa Speech Corpus* (SSC, from 1950 to 1970): The online search engine Chuunagon, contains approximately 0.53 million words.
   c *Balanced Corpus of Contemporary Written Japanese* (BCCWJ, 1971 to 2008): The online search engine Chuunagon, contains approximately 104.9 million words.
   d *Showa-Heisei Corpus of Written Japanese* (SHC, 1933 to 2013): The online search engine Chuunagon, contains approximately 33.40 million words.
   e *Nagoya University Conversation Corpus* (NUCC, 2001 to 2003): The online search engine Chuunagon, a total of 129 conversations, contains approximately 100 hours.
   f *Corpus of Everyday Japanese Conversation* (CEJC, 2016 to 2020): The online search engine Chuunagon, a total of 577 conversations, contains approximately 225 hours and 2.4 million words.

It has been demonstrated that PMs are usually used in spoken discourse (Östman 1982, Fraser 1990). For an expression to be considered a PM, markers have the interactional and argumentative function in the discourse (Aijmer & Simon-Vandenbergen 2011: 225). As seen in (2), the current paper includes three conversation corpora to understand the semantic-pragmatic function of *zenzen* in spoken discourse. This extensive dataset allows for quantitative and qualitative analysis of *zenzen*'s development as a PM.

4. The corpora survey

Using six written and spoken corpora mentioned in Section 3, 2,154 example sentences were extracted, including *zenzen*. Figure 1 and Table 1 illustrate the progress of the development of *zenzen*. Figure 1 shows the changing trend every ten years from the 1870s to the present. In the Figure, I divided *zenzen*'s meanings into three groups: affirmative, negative, and (inter)subjective. The negative group contains three different sentence types: (i) Co-occurring with verbal phrases (e.g., *koto-ni-suru* 'make different') or verbs (e.g., *chigau* 'to differ') that describe differences between two more things; (ii) Co-occurring with Sino-Japanese nouns combined with prefixes or suffixes indicating negativity (e.g., ***mu-imi*** 'meaning*less*'); (iii) Co-occurring with verbs or adjectives that include plain negative forms such as *-nai* 'not' (e.g., *kawara-**nai*** '*no* change'). The (inter)subjective group includes: (i) To intensify the speaker's perspective of something, i.e., subjectivity meaning; (ii) The speaker considers the addressee's

stances, reflecting on how the conversation might affect the addressee's feelings or attitudes, i.e., intersubjectivity meaning.
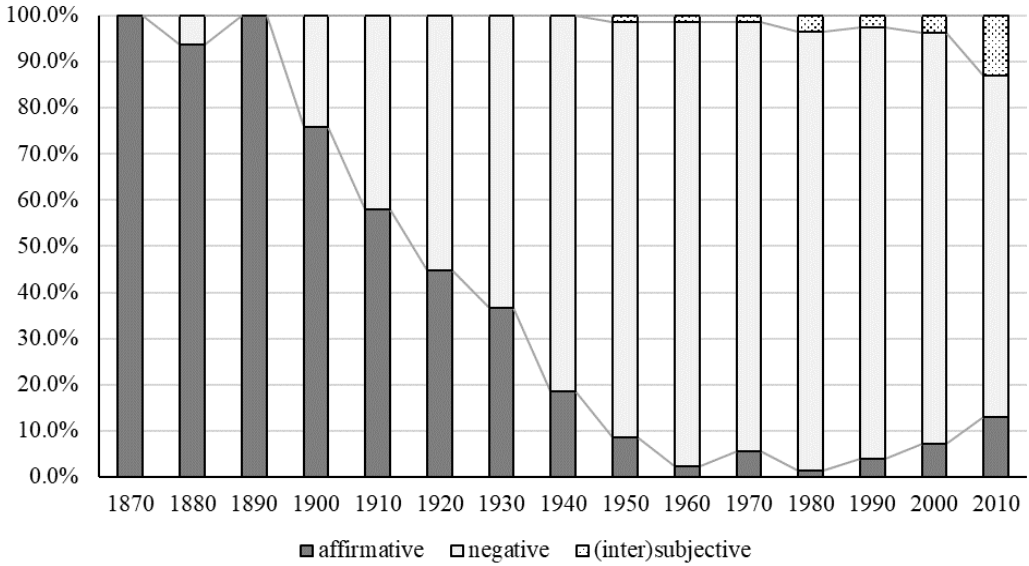
Figure 1. The semantic development of *zenzen*

Table 1. **The sematic distribution of *zenzen***

	Affirmative	**Negative**	**(Inter)Subjective**
1870	100% (6/6)	0.0% (0/6)	0.0% (0/6)
1880	93.8% (15/16)	6.3% (1/16)	0.0% (0/16)
1890	100% (30/30)	0.0% (0/30)	0.0% (0/30)
1900	75.9% (85/117)	24.1% (32/117)	0.0% (0/117)
1910	58.0% (47/81)	42.0% (34/81)	0.0% (0/81)
1920	44.9% (35/78)	55.1% (43/78)	0.0% (0/78)
1930	36.6% (26/71)	63.4% (45/71)	0.0% (0/71)
1940	18.5% (27/146)	81.5% (119/146)	0.0% (0/146)
1950	8.5% (17/200)	90.0% (180/200)	1.5% (3/200)
1960	2.3% (2/87)	96.2% (5/87)	1.5% (0/87)
1970	5.6% (4/71)	93.0% (66/71)	1.4% (1/71)
1980	1.4% (2/144)	95.1% (137/144)	3/5% (5/144)
1990	3.9% (12/310)	93.5% (291/310)	2.6% (7/310)
2000	7.2% (25/347)	89.1% (309/347)	3.7% (13/347)
2010	13.1% (59/450)	74.0% (333/450)	12.9% (58/450)

Figure 1 shows how *zenzen* has evolved over time. This result aligns with those of previous studies reviewed in Section 2. This longitudinal and quantitative analysis of the current paper provides a comprehensive assessment of the usage of

zenzen, as one of the few attempts to investigate the development using large-scale written and spoken data (cf., Okazaki 2008).

Notably, subjective and intersubjective meanings are more frequently observed in spoken than written registers. For instance, when comparing data from spoken and written sources using corpora in the 2000s, there was a higher frequency of subjective and intersubjective meanings in spoken discourse. NUCC and CEJC ((2e) and (2f)) were used to analyze the spoken data, while BCCWJ (2c)) was used for analyzing the written data. The results show that in the spoken sources, subjective and intersubjective meanings accounted for 12.4% (65/524), whereas in the written sources, they were observed less frequently at 3.2% (8/247). This result corroborates the idea that PMs are specialized in their characteristic of *orality* (Brinton 1996: 33).

Figure 2 shows how the rate of *zenzen* accompanying negative forms has changed over time. The negative forms frequently followed by *zenzen* are *-nai* 'not', and *-zu* 'without doing', as mentioned in Section 2.
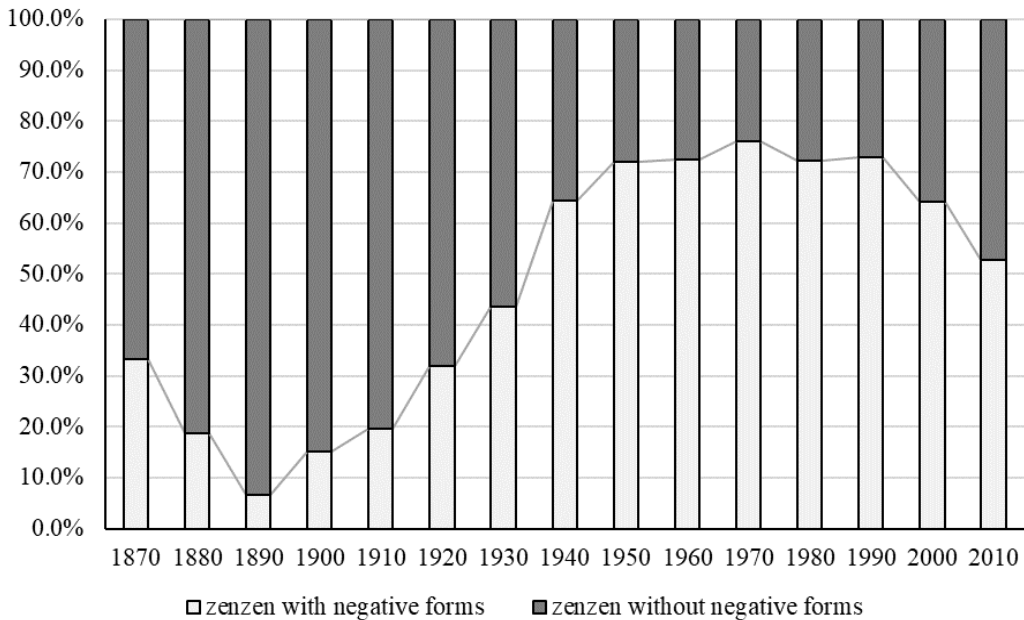
Figure 2. The rate of accompanying negative forms

In Figure 2, we can find a sharp increase in verbs and adjectives that include negative forms from the 1920s to a peak in the 1970s. The data in Figure 2 correlates with the shift in semantic development of the negative meaning of *zenzen* shown in Figure 1, supporting the idea that *zenzen* has experienced substantial semantic change.

5. Historical changes of *zenzen*

This section will provide an overview of the historical change in using *zenzen* from the Meiji Period to the present based on findings from earlier works and the data I collected from corpora.

5.1. From the Meiji to the early Shōwa Period (1868 to 1926)

It is known that *zenzen* was absorbed widely by the Japanese around the early 1900s (Niino 2011). In my data, even a few examples were observed, including *zenzen* in the 1870s to the 1880s, as in (3) to (4). In example (3), *zenzen* co-occurred with the verb describing negativity (i.e., *chigoo* 'to differ: two or more things do not correspond or match, or the information is incorrect'), and in (4), it is used in the sentence describing naturality (i.e., *nihon-no zokkoku-ni-shite* 'be the vassal state of Japan'). *Zenzen*, in both, expresses 'completely, totally, entirely'.

(3) … 政をすると教を施すとは全然仕法の違ふた。
    sei-o   *suru-to kyoo-o*   *hodokosu-to-wa* **zenzen**   *shihoo-no*
    politics-ACC   do-and   education-ACC   perform-with-TOP   completely   way-GEN
    chigoo-ta.
    differ-PST
    "The way of doing politics is *completely* different from performing in education."
              (1874, *Hyakuichishinron*, CHJ, 60C 口語 1874_06101, 30150)

(4) 沖繩の主權を確定し、沖繩は、全然日本の屬國にして、…
    Okinawa-no   *shuken-o*   *kakutei-shi okinawa-wa* **zenzen**   *nihon-no*
    Okinawa-GEN   sovereignty-ACC   confirm-do Okinawa-TOP   completely   Japan-GEN
    zokkoku-ni-shite, …
    a.vassal.state-in-be
    "Okinawa's sovereignty is confirmed, and Okinawa [will] *completely* be the vassal state of Japan."

           (1888, *Kokuminnotomo*, CHJ, 60M 国民 1888_14016, 2430)

In the initial stage, *zenzen* is used in sentences to indicate both affirmative and negative statements (Matsui 1977, Niino 2011, 2020). Notably, using affirmative statements was much more frequently found in my data from 1870 to 1910 in Figure 1. Syntactically, *zenzen* functions as an adverb and is usually positioned in the middle of the sentence.

From the early 1900s, there was a notable increase in the frequency of sentences conveying negative connotations, and this tendency has become increasingly prominent. It has become the main use of the word to date. Before the early 1900s, *zenzen* did not appear frequently combined with plain negative forms such as *-nai* 'not', *-zu* 'without doing', unlike its use in contemporary Japanese, as seen in Figure 2. Instead of combining plain negative forms, *zenzen* was much more used with general verbs or verbal phrases that convey negative meanings, such as *chigau* 'to differ', *koto-ni-suru* 'make different'. My data also found that *zenzen* often co-occurred with Sino-Japanese nouns that combined with prefixes or suffixes

indicating negativity. These examples were first observed in the late 1890s, such as *mu-* 'none', *fu-* 'not', and *-botsu* 'without' (Wakatabe 1991). Example (5) is the earliest example in my data.

(5) 左樣すると私は全然無関係だ。
     *sayoo-suru-to   watashi-wa **zenzen**   **mu**-kankei-da.*
     like.that-do-COND I-TOP   not.at.all none-connection-COP
     "If [it] likes to do that, I have *no* connection [with it] *at all*."
                                        (1909, *Taiyoo*, CHJ, 60M 太陽 1909_02026, 179090)

In the early stages of *zenzen* adoption by Japanese, there seemed to be no strict restrictions on its meaning and collocational relationships. However, since the early 1900s, its meaning and collocational relationships have undergone drastic changes within the language (i.e., "internally motivated" change, Heine & Kuteva 2005).

5.2. From the Shōwa to the early Heisei Period (1927 to 1989)

Since the early 1900s, the number of examples combined with elements that indicate the negative meaning of *zenzen* has noticeably increased, as seen in Figures 1 and 2, and since the middle 1900s, the negative meaning overwhelmingly increased, taking over the mainstream in the use of *zenzen* (see also the survey by Sano 2012). In addition, collocation elements with *zenzen* also have changed: *zenzen* frequently collocates with verbs or adjectives, including plain negative forms such as *-nai* 'not', as in (6) and (7).

(6) あなたと結婚する意志はもう全然ないんですから。
     *anata-to kekkon-suru ishi-wa   moo   **zenzen** nai-n-desu-kara.*
     you-with marriage-do intention-TOP anymore at.all   NEG-NOM-COP.POL-as
     "As I have no intention of marrying you anymore *at all*."
                                (1933, *Daiichi no Shussan*, SHC, 70B 女の 1933_10913, 2160)

(7) 勿論、政治的なことは全然關與しないといふ前提のもとに …
     *muron   seiji-teki-na   koto-wa **zenzen** kanyo-shi-nai-toiu*
     of.course politics-like-ADJ thing-TOP at.all   involvement-do-NEG-QUOT
     zentei-no-moto-ni …
     precondition-GEN-basis-in
     "Of course, on the precondition that [we] have no involvement in political matters *at all*."
                                (1944, *Chuookooron*, SHC, 70M 中公 1941_02011, 94390)

In my data, cases where *zenzen* collocates with negative words are particularly from the early 1950s to the late 1990s, occupying up to 90%, as seen in Figure 1. Furthermore, from around the 1940s, we can attest to examples that intensify the

speaker's epistemic stance, as in (8)[9] and (9). In (8), the speaker discusses an art exhibition hosted by the organization and expresses an opinion that the exhibition was subpar. In this sentence, *zenzen* functions emphasize the speaker's judgment that the exhibition was *absolutely* not good.

(8) 一水會の若い人は全然駄目だ。
    *issuikai[10]-no   wakai hito-wa   **zenzen**   dame-da.*
    issuikai-GEN   young people-TOP   at.all   no.good-COP
    "Young people at Issuikai are no good *at all*."
                      (1941, *Chuookooron*, SHC, 70M 中公 1941_01032, 27990)

In (9), the speaker thinks and says that the hearer probably dislikes a war *absolutely*. In this sentence, *zenzen* functions to add force to the emotional state, indicating that he/she *absolutely* hates a war.

(9) 戦争なんか全然嫌だと思うでしょう。
    *sensoo nanka   **zenzen**   iya-da-to   omou-deshoo.*
    war   things.like   absolutely   dislike-COP-QUOT   think-MOD
    "Do you think [you] *absolutely* dislike a war, [right]?"
                    (1952, *N-ke Zatsudan*, SSC, C52_03_CT, 23640)

The widespread use of *zenzen* combined with elements with negative meanings has led to the idea that using *zenzen* with positive or neutral meanings in a sentence is grammatically incorrect. As mentioned in Section 2.2, this idea is still partially accepted in contemporary Japanese.

In this stage, the use of *zenzen* underwent a significant change, shifting towards a more negative connotation. Additionally, examples that included morphological elements (e.g., *-nai* 'not') to express negativity became prevalent. Moreover, since around the 1940s, it has appeared that *zenzen* emphasized the speaker's stance at events. However, it is important to note that the meaning of 'completely, entirely, totally' has not disappeared; instead, in addition to these meanings, it has acquired a meaning of the subjective stance of a speaker[11].

5.3. From the Heisei Period to the present (1990 to the present)

Although the change was slight, negative expressions with *zenzen* began to decrease gradually, and examples of combining with affirmative ones increased, as seen in Figure 1. However, it should be noted that they are not the same as the

---

[9] The word *dame* does not always convey the speaker's negative stance; the word has various meanings, such as 'useless', 'cannot', and 'must not' (e.g., *Kore-ni sawat-tara dame-desu* 'You *should not* touch this').

[10] It is the name of a Japanese art organization founded in 1936.

[11] Niino (1997: 278) points out that the meaning of *zenzen* has not changed when used in both affirmative and negative sentences. It essentially means 'completely, entirely, totally'. However, using *zenzen* in a sentence affirms something at 100%, not just 50% or 60%. This unique usage of *zenzen* has led to an emphasis on the meaning of the words and phrases modified by it. Consequently, some *zenzen* in sentences could be interpreted as meaning 'very, extremely'.

examples we have looked at in 5.1. The examples seen in the very early stage had more objective meaning; however, in many of the examples found in this Period, *zenzen* functions as an intensifier of the speaker's subjective stance and expressing 'completely, entirely, totally' in spoken discourse.

In the following example (10), the speaker talks about the plant he bought at a store, which looks better than the one he did not buy. In (10), *zenzen* intensifies the niceness of the appearance.

(10) こっちのほうが全然見栄えがいいやと思って。
　　　kotchi-no hoo-ga ***zenzen*** *mibae-ga ii-ya-to omot-te.*
　　　this-GEN one-NOM absolutely look-NOM nice-SFP-QUOT think-CVB
　　　"I thought this one looks *absolutely* nicer."

<div align="right">(2016, Zatsudan, CEJC, T007_005a, 31580)</div>

Morphosyntactically, *zenzen* is typically positioned in the medial of the sentence and modifies the following verbs or adjectives, as we have seen above. However, in spoken discourse, it has also emerged that stand-alone forms like holophrases were used as a response, as in (11). This use of *zenzen* is considered syntactically unattached status, one of the common effects of cooption. In (11), speakers A and B talk about the progress of speaker B's studies, and speaker B intensifies the lack of progress in studies. It uses an elliptic form (i.e., stand-alone form) without repeating the previous proposition, resulting in a compelling focus of negation (see also Lee 2018).

(11) A: ちょっと読んでたんだよ。
　　　chotto yon-de ta-n-da-yo.
　　　little read-CVB PST-NMZ-COP-SF
　　　"[I] read the book a little."
　　B: あ、そうなんだ。
　　　a, soo-na-n-da.
　　　oh that-COP-NMZ-COP
　　　"Oh, I see."
　　A: うん。
　　　un.
　　　yeah
　　　"Yeah."
　　B: はかどった？
　　　hakadot-ta?
　　　make.a.progress-PST
　　　 "Did you make any progress on your studies?"
　　A: 全然。
　　　zenzen.
　　　not.at.all
　　　"[I could not make any progress in my studies] *at all*."

<div align="right">(2001, NUCC, data003, 118160)</div>

It is important to note that in this context, *zenzen* is used to respond in interactive conversation using stand-alone form, always to convey negativity. It is attributed to the fact that *zenzen* has predominantly been associated with negativity for nearly a century, except for the initial stage when it was borrowed from Chinese. Niino (2011: 154) examines the use of holophrases in *zenzen* and argues that the latter part of the negative expression can be dropped. Niino (2011) explains that *zenzen* is usually used with negative expressions, making it unnecessary to explicitly state the negative part in a sentence, with the other part of *zenzen* being omitted as a result (also see Koike 2001). Therefore, the latter part of the sentence might be redundant.

In spoken discourse, *zenzen* is often found at the end of the sentence, as in (12) to (13). Its position at the end of the sentence suggests that the word's (inter)subjectivity may be related to its position in the discourse context (Beeching & Detges 2014: 11).

The following example (12) is interesting, as we can see and identify that *zenzen* has different meanings and functions in the second and third sentences, respectively.

(12)　A: 言葉が知りたいってゆうことになったら、
　　　　kotoba-ga　　shiri-tai-tte　　　yuu-koto-ninat-tara,
　　　　language-NOM　know-want-QUOT　say-thing-become-COND
　　　　"If you want to know the language."
　　　　あ、なんでもいいよ、全然。
　　　　*a,　nan-demo　ii-yo,　　**zenzen**.*
　　　　INJ what-even　fine-SFP　totally
　　　　"Oh, anything is fine, *totally*."
　　　　絶対これってゆうものは全然ないので。
　　　　*zettai　　　kore-tte　　yuu-mono-wa **zenzen**　　nai-node.*
　　　　absolutely　this-QUOT　say-thing-TOP　not.at.all　not-as
　　　　"As [there is] absolutely nothing like this *at all*."
　　　　　　　　　　(2019, *Jugyoo and Ressun*, CEJC, W003_001, 31000)

In (12), in the case of the speaker speaking to the audience, the speaker is open to any choice, emphasizing flexibility or indifference. In the second sentence of (12), *zenzen* intensifies the speaker's subjective attitude. From the perspective of the interactional aspect, *zenzen* functions to acknowledge the addressee's potential concern, and the sentence-ending position of *zenzen* may strengthen the speaker's reassurance to the addressee that any option is entirely acceptable. On the other hand, in the third sentence, *zenzen* represents objective meaning rather than the speaker's epistemic stance that emphasizes that there are no such things as good questions *at all*.

In (13), speaker B discusses the name chosen by the hearer and describes it as outdated. The word *zenzen*, positioned at the end of the sentence, emphasizes the speaker's attitude that the name of the selection sounds extremely old-fashioned and undesirable. Therefore, the speaker is requesting the name change.

(13)   A: ··· セレクションみたいなやつ作って。
          serekushon-mitai-na   yatsu tsukut-te.
          selection-like.to-ADJ   thing make-CVB
          "[I] made things like as a selection."

       B: うん。
          un.
          ah
          "Ah."

       A: 駿介セレクション。
          Sunsuke serekushon.
          PSN   selection
          "[The name of the selection is] *Sunsuke* selection."

       B: や、それはださいじゃん、全然。
          ya, sore-wa dasai-jan,   zenzen.
          INJ   that-TOP outdated-surely   absolutely
          "Oh, that is *absolutely* outdated."
          もう俺にはなんも関係ないぽい名前付けて。
          moo   ore-ni-wa nanmo   kankei-nai poi namae-tsuke-te.
          anymore me-to-TOP anything relation-not ish   name-make-CVB
          "To make a name that does not seem to have anything to do with me anymore."
                                    (2018, *Shokuji and Kyuusoku*, CEJC, K011_015, 196140)

In the second utterance of speaker B, *zenzen* not only emphasizes the outdatedness of the name but also serves a function expected to convince the addressee to change it.

Since the 2000s, it has been observed that in an adjectival noun combined with copula, the construction [*zenzen-da/desu*] emerges and functions as phrasal.

(14)   A: 見て、中が、全然まだ、中がちょっと。
          mi-te,   naka-ga,   zenzen mada, naka-ga   chotto.
          look-CVB inside-NOM   not.at.all yet   inside-NOM bit
          "Look. Inside, [it is] still *not at all* [cooked] *at all*, [it is not cooked] a bit inside."

       B: あ、あれ。
          a,   are.
          INJ   oh
          "Ah, oh."

       A: 全然だよ。
          zenzen-da-yo.
          not.at.all-COP-SFP
          "[It is] *not* [cooked] *at all*."

       B: うそ。
          uso.
          joke
          "[You must be] joking."

A: ほら、全然だね。
hora, **zenzen-da-ne.**
look   not.at.all-COP-SFP
"Look. [It is] *not* [cooked] *at all*."
B: じゃ、もっ回揚げるわ。
ja,   mokkai   ageru-wa.
well again   fry-SFP
"Well, I will fry it again."

(2018, *Kaji Zatsudan*, CEJC, K013_009, 86550)

In (14), speaker A comments that the inside of fried food is not cooked. In the following discourse, speaker A uttered twice the construction [*zenzen-da*] that contains a negative statement: 'The food is not cooked at all'. On the other hand, in the first sentence of A, the part *zenzen mada* seems partially omitted in the latter part of the sentence, such as a verb (e.g., *deki-te i-nai* 'not done'). Thus, this sentence should not be treated like speaker A's second and third utterances: the fixed construction [*zenzen-da*], which functions as an independent sentence per se.

In addition, the construction [*zenzen-desu*], which conveys intersubjective meanings, is often found in spoken discourse, as shown in (15). Speaker A is a customer at speaker B's shop and visits the shop on a holiday. Speaker A apologizes for visiting on a holiday. However, in the second utterance of speaker B, speaker B reassures speaker A that it is completely acceptable and attempts to alleviate the customer's guilt and concern.

(15)   A: 今日お休みですか。
        kyoo   oyasumi-desu-ka.
        today holiday-COP.POL-Q
        "Is it a holiday today?"
       B: 休みです。はい。
        yasumi-desu.   hai.
        holiday-COP.POL   yes
        "[Today is] a holiday. It is."
       A: すみませんね。そんな貴重なお休みに。
        sumimasen-ne.   sonna-kichoo-na   oyasumi-ni.
        I.am.sorry-SFP   such-precious-ADJ holiday-on
        "I am sorry. [As visiting your shop] on [your] such precious holiday."
       B: あっ、いえ、全然です。
        atsu, ie, **zenzen-desu.**
        INJ   no perfectly-COP.POL
        "Oh, no, [I am] *perfectly (fine)* [with it]."
                                (2017, *Yoodan and Zatsudan*, CEJC, T014_012, 1110)

In speaker B's second utterance, the construction [*zenzen-desu*] could be replaced with the expression *Daijoobu-desu* (fine-COP.POL), which means 'It is okay'. Interestingly, since the 2000s, *zenzen* has often been used with *Daijoobu-desu* and functions as an intensifier in spoken language. In my corpora data, I found

six examples of ***Zenzen*** *daijobu-desu* '[It is] *perfectly* fine'. It is interesting to note that *zenzen* is used as an emphatic adverb, forming the construction itself.

To sum up, when *zenzen* was initially borrowed from the Chinese, it described the objective meaning of 'completely, entirely, totally' or 'not at all' and could occur within both affirmative and negative sentences. In my corpora survey, it was used much more often with affirmative sentences when it was first introduced in Japanese. Since the early 1900s, there has been a sharp increase in the use of the word *zenzen* in negative sentences to convey the meaning of 'not at all'. Consequently, using *zenzen* to express negativity became standard in the 1900s. On the other hand, in spoken language, there has been a gradual increase in using *zenzen* within affirmative sentences since the middle of the 1900s, and a few of them represent the speaker's subjective meanings. From the 2000s onwards, significant changes have been observed in the morphosyntactic and semantic use of *zenzen*, particularly in spoken discourse. *Zenzen* represents subjective meanings to intensify the speaker's attitude and viewpoint. All these expressions have coexisted in contemporary Japanese.

In Section 6, we will discuss the morphosyntactic and semantic-pragmatic change, focusing on the construction [*zenzen-da/desu*].

6. Discussion

In this section, we will discuss how the PM function of *zenzen* emerged from the perspective of cooptation (Kaltenböck, Heine & Kuteva 2011, Heine et al. 2021), subjectification (Lyons 1982, Traugott & Dasher 2001, Traugott 2010), and intersubjectification (Traugot 2003) focusing on the construction [*zenzen-da/desu*] and its morphosyntactic and semantic-pragmatic change.

In the investigation into the common effects of cooption, as seen in Section 2.5., *zenzen* has undergone cooption progress in its semantic-pragmatic manipulation and syntactic features. In contemporary Japanese, *zenzen* typically functions as an adverbial, often standing alone (i.e., as a holophrase), especially as a response to the hearer in spoken discourse, as shown in (16). Example (16) below is a repetition of example (11) above. In (16), speaker A asked about the progress of speaker B's studies. Speaker B commented that he has not made remarkable progress *at all*.

(16) A: ちょっと読んでたんだよ。
      chotto   yon-de   ta-n-da-yo.
      little   read-CVB PST-NMZ-COP-SF
      "[I] read the book a little."
   B: あ、そうなんだ。
      a,   soo-na-n-da.
      oh   that-COP-NMZ-COP
      "Oh, I see."

A: うん。
   un.
   yeah
   "Yeah."
B: はかどった？
   hakadot-ta?
   make.a.progress-PST
   "Did you make any progress on your studies?"
A: 全然。
   zenzen.
   not.at.all
   "[I could not make any progress in my studies] *at all*."

(2001, NUCC, data003, 118160)

This use of *zenzen* is syntactically unattached, which is one of the common effects of cooptation. Interestingly, *zenzen* stands alone as a response in interactive conversation, and it is always used to express negativity. This use of *zenzen* has often been founded in spoken discourse since the late 1990s. A possible explanation for this could be linked to the economy principle of language. As shown in Figure 1 and Table 1, *zenzen* has remained predominantly associated with negativity for nearly a century except for its initial stage when borrowed from Chinese, fostering a strong collocational relationship between *zenzen* and negative meanings (Niino 2011). In a longitudinal study of early Modern English texts, Vicentini (2003: 55) concludes that due to the human tendency to reduce physical and mental efforts, the economy principle has maintained a balance between characteristics that ensure efficient and direct communication on the one hand and the natural need for least effort on the other. This finding could help us understand those examples of *zenzen* (16).

On the other hand, since the 2010s[12], a new unique construction [*zenzen-da/desu*] emerged and functioned as a phrasal (i.e., an adjectival noun), a fixed expression representing the subjective stance and intersubjective markings well. This shift from adverbial modifying verbs or adjectives to an adjectival noun involved syntactic and functional reanalysis. It was found that largely three meanings have the construction [*zenzen-da/desu*] in my data: (i) Negation ((17)); (ii) Completeness or perfection ((18)); and (iii) Affirmation and reassurance to address ((19)). The first and second ones express subjectivity, while the third one expresses intersubjectivity. It is worth noting that although the three meanings demonstrate different functions regarding (inter)subjectivity, they all share a common semantic feature of denying the contextual background proposition (Arimitsu 2008).

---

[12] To add reliability, I searched blog postings through a major online search engine (https://www.yahoo.co.jp/, accessed 7 July 2024); this usage has more frequently appeared since 2020.

In (17), speaker B explains how it takes time to reach the destination (i.e., Hakata), emphasizes the distance to the destination as undemanding, and denies any difficulties in his/her journey.

(17) A: 博多行きってあんの。
        hakata yuki-tte   a-n-no.
        hakata go-QUOT   exist-NMZ-Q
        "Is there [a train] going to Hakata?"

      B: ある。
        aru.
        exist.NPST
        "There is."

      A: あ、そう。
        a,   sou.
        INJ I.see
        "Oh, I see."

      B: うん。四時間。
        un,   yo-jikan.
        yes four-hour
        "Yes, it takes four hours."

      A: え、大したことないじゃん。
        e,   taishita-koto nai-jan.
        INJ big.deal   NEG-isn't.it
        "Oh, it is not a big deal. Isn't it?"

      B: うん、大したことないよ。全然大したことない。
        un, taishita-koto nai-yo.   zenzen   taishita-koto nai.
        yes big.deal   NEG-SFP   not.at.all big.deal   NEG
        "Yes, it is not a big deal. it is not a big deal at all."

        …
        四時間、四時間半か、四時間。全然だ。
        *yo-jikan,   yo-jikan-han-ka,   yo-jikan.   **zenzen**-da.*
        four-hour   four-hour-half-Q   four-hour   not.at.all-COP
        "[It takes] four hours, four hours and a half, four hours.
        [It is not a big deal] *at all*."

<div align="right">(2016, Shokuji, CEJC, T014_001b, 12470)</div>

In (17), speaker B uses *zenzen* twice in its third utterance, in the first and final sentences. Both instances of *zenzen* convey negative meanings. However, these two uses differ syntactically: the first is used with a plain negative form -*nai* 'not', which is the most widely used; the second is regarded as a newly emerged construction [*zenzen-da*]. The construction of that simpler structure can be fully carried out to express the same meaning the first complete sentence conveys, assisting the previous utterance and contextual background.

The construction [*zenzen-da/desu*] also emphatically represents a speaker's belief in the completeness and perfection of a situation. In (18), speakers A and B

talk about a cold, and speaker B says if he/she catches a cold, he/she will only have a runny nose while his/her eyes are completely fine, and there will be no problem.

(18) A: だから鼻ってそうですね。風邪引いた時のその鼻水ぐらいで。
      dakara hana-tte   soo-desu-ne.   kaze hii-ta   toki-no   sono
      so   nose-QUOT that-COP.POL-SFP cold catch-PST   when-GEN   that
      hanamizu gurai-de.
      snot   just-COP
      "So, the nose is like that. It is just that runny nose when you catch a cold."
      B: あっ、そのぐらいですね。はーい。全然ですね、目は。
      *atsu, sono gurai-desu-ne.   haai. **zenzen**-desu-ne,   me-wa.*
      INJ   that just.like-COP.POL-SFP yes   completely-COP.POL-SFP eye-TOP
      "Yeah, [my nose is] just like that. Okay, my eye is *completely [fine]*."
                                    (2017, *Shokuji & Tsukiai*, CEJC, T014_012, 57110)

The construction [*zenzen-desu*] in (18) conveys the meaning of the safety and troubledness of his/her eye condition. This example is interpreted as speaker B expressing the speaker's subjective belief and emphasizing what is being said.

Furthermore, the construction [*zenzen-da/desu*] indicates intersubjective function. In this usage, it was found that many constructions in my data include -*desu*[13], which is a form to express more formality and politeness than a form -*da*. This intriguing finding may reflect the differences in characteristics between subjectification and intersubjectification. Subjectification occurs when meanings become more centered on the speaker, while intersubjectification occurs when meanings become more centered on the addressee (Traugott 2003). Thus, using a more polite form -*desu* might be appropriate to indicate intersubjective function.

In my data, intersubjectification is prominent in developing the interlocutor's affirmations and reassurances. In (19), [*zenzen-desu*] indicates reassuring and mitigating concerns or apologies by the interlocutor.

(19) A: 一時間経った。
      ichi-jikan tat-ta.
      one-hour   pass-PST
      "[It] passed one hour."
      B: いいえ、全然ですよ。全然なの。
      *iie, **zenzen**-desu-yo.   **zenzen**-na-no.*
      no   absolutely-COP.POL-SFP absolutely-COP-SFP
      "No, it is *absolutely [fine]*. It is *absolutely [fine]*."
      そんな気にせず普通に話してだいじょぶだから。
      sonna ki-ni se-zu   futsuuni hanashi-te daijoobu-da-kara.
      such concerned.about-NEG   normally talk-CVB   fine-COP-as
      "Do not worry [about it]; [it is] as fine to talk normally."
                          (2016, *Shigoto & Mi-no Mawari-no Yooji*, CEJC, T011_007, 111180)

---

[13] The construction [*zenzen-da + yo/ne*] as in (14) would indicate intersubjective function as well. Speaker A uses this construction to point out that the food is not cooking well, implying an urging for speaker B to cook (i.e., to fry) the food again.

In the above (19), [*zenzen-desu*] conveys the meaning of reassuring the interlocutor, indicating no worries, and showcasing its use in providing positive affirmation. The example (19) supports evidence from previous findings: intersubjective stance is based on alignment with the interlocutor (Du Bois 2007, Kaltenböck et al. 2011).

The examples reported in the paper above indicate that the construction [*zenzen-da/desu*] deployed by speakers to capture what they characterize as their interlocutor stances or attitudes can be used in subjective and intersubjective functions. The contemporary use of the construction [*zenzen-da/desu*] in various contexts represents a significant departure from historical changes. This transformation can be attributed to a broader cooptation process (Heine et al. 2021), where the construction indicates negativity, completeness, and reassuring the interlocutor in spoken discourse.

Based on the grammatical effects by Heine et. al (2021), as seen in 2.5, we can conclude that the conditions of the parameters support the construction [*zenzen-da/desu*]. Examples (17) to (19) above show that the meaning of the construction is non-restrictive, varies with context has a wider scope (i.e., negativity, completeness, and offering reassurance), and depends on the contextual situation. This construction is favored in spoken discourse that involves high interactions and assumes the presence of interlocutors. Moreover, *zenzen* is initially used as an adverb and modifies verbs or adjectives, but it is found in a specific construction that combines with copula, leading to syntactical independence from the sentence.

The development from the Meiji Period to the present can be diagramed in Figure 3. In Stage 1, when it was initially adopted in the Japanese language, there were no significant differences in the frequency between the meanings 'completely' and 'not at all', while in Stage 2, the meaning 'not at all (the bold one)' was much more predominant.

Stage I	Stage II	Stage III	Stage IV
Adverbial	Adverbial	Adverbial, Stand-alone form	Adjectival noun (Phrasal)
'completely, not at all' >	'completely, **not at all**' >	intensifier, 'not at all' >	(inter)subjective marker

GRAMMATICALIZATION
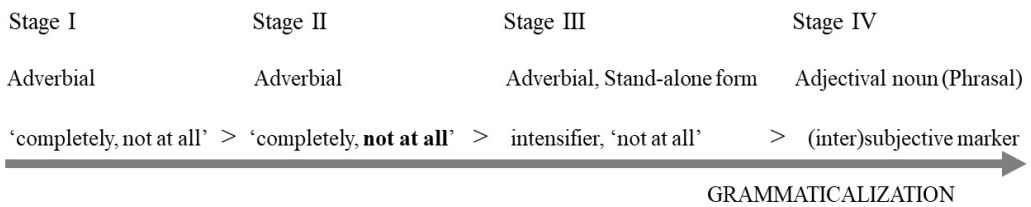
Figure 3. The development of *zenzen*

It is important to note that the semantic-pragmatic functions accumulate and co-occur. Therefore, acquiring subjective and intersubjective functions does not entail relinquishing objective meanings; 'completely, not at all'. It is a matter of embracing both, not exclusively one or the other.

7. Conclusion

This paper investigated the evolution of the Japanese PM *zenzen*（全然）from its historical origins to its contemporary usage, revealing its transition from an adverb with objective meanings to a PM with (inter)subjective functions in spoken discourse from the perspective of cooptation and (inter)subjectification. The paper traced its development from the Meiji Period to the present and highlighted significant semantic-pragmatic functions and morphosyntactic shifts.

Taken together, the findings of the current paper add to the growing body of study on historical language change in terms of cooption, subjectification, and intersubjectification.

Abbreviations

ACC = accusative case; ADJ = adjectivizer; COND = conditional mood; COP = copula; CVB=converb; GEN = genitive case; INJ = interjection; MOD = modal; NEG = negation; NMZ = nominalizer; NOM = nominative case; NPST = non-past tense; PASS = passive voice; POL = polite; PSN = personal name; PST = past tense; Q = question particle; QUOT = quotative; SFP = sentence-final particle; TOP = topic marker

References

Aijmer, Karin. 2002. *English Discourse Particles: Evidence from a Corpus*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/scl.10

Aijmer, Karin & Anne-Marie Simon-Vandenbergen. 2011. Pragmatic markers. In Jan Zienkowski, Jan-Ola Östman & Jef Verschueren (eds.), *Discursive pragmatics*, 223–247. Amsterdam/Philadelphia: John Benjamins Publishing. https://doi.org/10.1075/hoph.8

Arimitsu, Nami. 2008. Nichi, Eigo-no taihi hyoogen-ni mirareru hi-meiji-teki hitei-sei-to ryoo•shitsu•taido-ni kansuru henka-no mekanizumu [Non-explicit negativity in Japanese-English contrastive expressions and mechanisms of change regarding quantity, quality, and attitude]. In Kazuhiro Komada & Tetsuharu Koyama (eds.), *Kotoba-to Ninchi-no Mekanizumu: Yamanashi, Masa-aki kyooju Kanreki Kinen Rombun-shuu [Linguistic and Cognitive Mechanism: Festschrift for Professor Yamanashi, Masa-aki on The Occasion of His Sixtieth Birthday]*. 247–269. Tokyo: Shōgakukan.

Beeching, Kate & Ulrich Detges. 2014. *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Vol. 12. Leiden: Brill. https://doi.org/10.1163/9789004274822

Benveniste, Emile. 1971. Subjectivity in language. *Problems in General Linguistics* 1. 223–230.

Brinton, Laurel J. 1996. Attitudes toward increasing segmentalization: Complex and phrasal verbs in English. *Journal of English Linguistics* 24 (3). 186–205. https://doi.org/10.1177/00754242960240030

Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49 (2). 365–390. https://doi.org/10.1515/LING.2011.011

Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139–182. Amsterdam/Philadelphia: John Benjamins Publishing. https://doi.org/10.1075/pbns.164.07du

Endo, Orie. 1994. Shiyoo goshu-to, atarashii kotoba-no yoohoo [A use of word class and usage of new words]. *Kotoba* 15. 114–134.

Endo, Orie & Hiroko Yabe. 1995. Hanashi kotoba-ni tokuchoo-teki-na go-no atarashii yoohoo-to sedai-sa: 'Sugoi', 'toka', 'zenzen', 'kekkoo' ni-tsuite [A new usage of words characteristic of spoken language and generational differences: 'Sugoi', 'toka', 'zenzen', 'kekkoo']. *Kotoba* 16. 114–127.

Fei, Gao. 2012. *Nihongo-to Chuugokugo-no Hitei-to Koou-suru Fukushi-no Hikaku Koosatsu [A comparative study of negation and corresponded adverbs in Japanese and Chinese]*. Master Thesis. Aichi University of Education.

Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14 (3). 383–398. https://doi.org/10.1016/0378-2166(90)90096-V

Fraser, Bruce. 2009. An account of discourse markers. *International Review of Pragmatics* 1 (2). 293–320. https://doi.org/10.1163/187730909X12538045489818

Hansen, Maj-Britt Mosegaard. 1998. *The Function of Discourse Particles: A Study with Special Reference to Spoken Standard French. Vol. 53*. Amsterdam/Philadelphia: John Benjamins Publishing. https://doi.org/10.1075/pbns.53

Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511614132

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. Chapter 2. Concepts of analysis. *The Rise of Discourse Markers*. 56–90. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108982856

Jucker, Andreas H. & Yael Ziv. 1998. *Discourse Markers: Descriptions and Theory*. Amsterdam/Philadelphia: John Benjamins Publishing. https://doi.org/10.1075/pbns.57

Kaltenböck, Gunther, Bernd Heine & Tania Kuteva. 2011. On thetical grammar. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 35 (4). 852–897. Amsterdam/Philadelphia: John Benjamins Publishing. https://doi.org/10.1075/sl.35.4.03kal

Koike, Seiji. 2001. "Zenzen" sai-sai kou [A re-note "zenzen"]. *Udai Kokugo Ronkyuu* 12. 1–11.

Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1 (1). 5–38. https://doi.org/10.1515/cogl.1990.1.1.5

Lyons, John. 1982. Deixis and subjectivity: Loquor, ergo sum? In Robert J. Jarvella & Wolfgang Klein (eds.), *Speech, place, and action: Studies in deixis and related topics*. New York: Wiley.

Matsui, Shigekazu. 1977. Kindai koogo-bun-ni okeru teido fukushi-no shoochoo: Teido-no hadashisa-o arawasu baai [The rise and fall of adverbs of degree in Modern colloquial texts: The case of degree of extremeness]. *Kokugo-gaku to Kokugo-shi: Matsumura Akira Kyooju Kanreki Kinen [Japanese Linguistics and History of the Japanese Language in Commemoration of the 60th Anniversary of Professor Akira Matsumura]*. 737–758. Tokyo: Meiji Shoten.

Mizutani, Shizuo, Toshio Ishiwata & Tastuo Miyajima. 2017. *Chapter 2. Go-no ryoo-teki-na koozoo [The quantitative structure of words]. Gendai Zasshi Kyujusshu-no Yoogo Yooji,*

Dai Sambun Satsu: Bunseki [Vocabulary and Chinese Characters in Ninety Magazines of Today, Vol. III: Analysis of The Results]. Tokyo: Shueishobo Co., Ltd.

Niino, Naoya. 2011. *Chapter 2. "Zenzen" + kotei-o meguru kenkyuu [A study of "zenzen" + affirmative], Gendai Nihongo-ni okeru Shinkoo-chuu-no Henka-no Kenkyuu: Goyoo, Kizukanai Henka-o Chuushin-ni [A Study of Ongoing Changes in Modern Japanese: Focusing on Misuse and Unnoticed Changes].* Tokyo: Hituzi Shoboo.

Niino, Naoya. 2020. *Kingendai Nihongo-no Goyoo-to Gengo Kihan Ishiki-no Kenkyuu* [*A Study of Misuse and Norm Consciousness in Modern Japanese*]. Tokyo: Hituzi Shoboo.

Okazaki, Koichi. 2008. "Zenzen" kou [A note "zenzen"]. *Shinwa Kokubun* 43. 1–21.

Östman, Jan-Ola. 1981. A functional approach to English tags. *Studia Anglica Posnaniensia* 13. 3–16.

Östman, Jan-Ola. 1995. Pragmatic particles twenty years after. *Anglicana Turkuensia* 14. 95–108.

Redeker, Gisela. 1991. Linguistic markers of discourse structure. *Linguistics* 29. 1139–1172. https://doi.org/10.1515/ling.1991.29.6.1139

Sano, Shin-ichiro. 2012. Nihongo hanashi kotoba koopaasu-o mochiita zenzen-no henka-no shoosai-ka [Anatomy of the change of zenzen using the corpus of spontaneous Japanese]. *Dai Ik-kai Koopaasu Nihongo-gaku Waakushoppu Yokooshuu*. 32–42.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511611841

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65 (1). 31–55. https://doi.org/10.2307/414841

Traugott, Elizabeth Closs. 2003. From subjectification to intersubjectification. In Raymond Hickey (ed.), *Motives for language change*, 124–139. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486937.009

Traugott, Elizabeth Closs. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Kristin Davidse, Lieven Vandelanotte, Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization*, 29–71. Berlin: Mouton De Gruyter. https://doi.org/10.1515/9783110226102.1.29

Traugott, Elizabeth Closs. 2018. Modeling language change with constructional networks. In Bordería, Salvador Pons & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers*, 17–50. Leiden: Brill. https://doi.org/10.1163/9789004375420_003

Traugot, Elizabeth Closs & Richard B. Dasher. 2001. *Regularity in Semantic Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486500

Vicentini, Alessandra. 2003. The economy principle in language. *Notes and Observations from Early Modern English Grammars. Mots, Palabras, Words* 3. 37–57.

Wakatabe, Akira. 1991. "Zenzen" no goshi-teki kenkyuu: Meiji kara gendai made [A study of the evolution of the usage "zenzen": From Meiji to Modern Period]. *Kaishaku* 37 (11). 24–29.

Corpora

NINJAL (2024). *The Balanced Corpus of Contemporary Written Japanese* (BCCWJ). https://clrd.ninjal.ac.jp/bccwj/index.html (accessed 22 August 2024).

NINJAL (2024). *Corpus of Everyday Japanese Conversation* (CEJC). https://www2.ninjal.ac.jp/conversation/cejc.html (accessed 22 August 2024).

NINJAL (2024). *Showa Speech Corpus* (SSC). https://www2.ninjal.ac.jp/conversation/showaCorpus/ (accessed 22 August 2024).

NINJAL (2024). *Nagoya University Conversation Corpus* (NUCC). https://mmsrv.ninjal.ac.jp/nucc/ (accessed 22 August 2024).

NINJAL (2024). *The Corpus of Historical Japanese* (CHJ). https://clrd.ninjal.ac.jp/chj/ (accessed 22 August 2024).
NINJAL (2024). *Showa-Heisei Corpus of written Japanese* (SHC). https://clrd.ninjal.ac.jp/shc/ (accessed 22 August 2024).

Dictionaries

Dejitaru Daijirin [Digital Daijirin]. Tokyo: Shōgakukan.
Nihon Kokugo Daijiten [Unabridged Japanese Dictionary]. 1988. Tokyo: Shōgakukan.

Bionote:
Jiyeon PARK is a Lecturer of the Korean language at the Department of Humanities, Matsuyama University, Japan. She received her PhD in linguistics from Nagoya University, Aichi, Japan in 2019. Her main research interest is to investigate how Korean and Japanese ideophones are used in everyday language and which language factors make those to be used.
e-mail: j.park@g.matsuyama-u.ac.jp
http://orcid.org/0000-0003-1267-0573

Сведения об авторе:
Джиён ПАРК – преподаватель корейского языка кафедры гуманитарных наук Университета Мацуяма, Япония. Она получила степень PhD по лингвистике в Университете Нагоя, Айти, Япония, в 2019 г. Ее основной научный интерес – изучение языковых причин и особенностей употребления корейских и японских идеофонов в повседневном языке.
e-mail: j.park@g.matsuyama-u.ac.jp
http://orcid.org/0000-0003-1267-0573

Research article / Научная статья

From a noun to a discourse marker: The case of *seysang* 'world' in Korean

Hyun Sook LEE ⍟ ✉

Jangan University, Hwaseong-si, Republic of Korea
✉morningdew1020@jangan.ac.kr

Abstract
This study delves into the semantic evolution of the Korean lexeme *seysang*. Despite its rich functional development, *seysang* has received limited attention in grammaticalization research. This study addresses this gap by exploring its extended meanings and functional shifts in Korean. It aims to trace the transformation of *seysang* from a historical noun to its role as a contemporary discourse marker (DM) which conveys a range of pragmatic meanings, marking the speaker's emotions and stances. Using historical and contemporary corpora, including data spanning from the 15th century to modern media, this paper qualitatively analyzes the grammaticalization process of *seysang* while quantitatively examining the discourse functions of *seysangey*. The analysis particularly focuses on *seysangey*'s positional flexibility within Left Periphery (LP), Right Periphery (RP), and stand-alone positions, and its preference in positive, negative, and neutral contexts. This study further investigates whether *seysangey*'s LP and RP functions align with previous research, which associates LP with subjectification and RP with intersubjectification. The analyses indicate that *seysangey* exhibits both subjectification and intersubjectification across LP, RP, and stand-alone positions. Findings reveal that *seysang* has expanded to signify broader social spaces, from birth-to-death spans and societal environments beyond enclosed communities to symbolic meanings of people's hearts and the earthly realm. It also functions adverbially as 'very' (degree modifier) and 'at all' (negative polarity item). In Contemporary Korean, *seysangey* operates as a flexible DM marking subjectivity, intersubjectivity, and speaker stance, providing insight into the complexities of language evolution and external influences shaping the Korean lexicon and grammar.
Keywords: *grammaticalization, subjectification, intersubjectification, stance marking, seysang, seysangey*

От существительного к дискурсивному маркеру:
seysang 'мир' в корейском языке

Хён Сук ЛИ 🆔 ✉

Чанганский университет, Хвасон-Си, Республика Корея
✉morningdew1020@jangan.ac.kr

Аннотация
Данная работа посвящена исследованию семантической эволюции корейской лексемы *seysang*, которая, несмотря на развитие новых функций, получила ограниченное внимание в исследованиях по грамматикализации. Настоящее исследование устраняет этот пробел и изучает расширенные значения существительного *seysang* и функциональные сдвиги в корейском языке. Оно прослеживает трансформацию *seysang* от существительного в дискурсивный маркер (ДМ), который передает ряд прагматических значений, указывающих на эмоции и позицию говорящего. Используя исторические и современные корпусы, включающие данные с 15 века до наших дней, эта статья качественно анализирует процесс грамматикализации *seysang* и одновременно количественно исследует дискурсивные функции *seysangey*. Особое внимание уделяется позиционной гибкости лексемы *seysangey* в пределах левой периферии (ЛП), правой периферии (ПП) и отдельных позиций, а также ее предпочтению в положительных, отрицательных и нейтральных контекстах. В исследовании также ставится вопрос о том, совпадают ли выявленные функции ЛП и ПП с результатами предыдущих исследований, которые связывают ЛП с субъективацией, а ПП с интерсубъективацией. Анализ свидетельствует о том, что *seysangey* демонстрирует как субъективацию, так и интерсубъективацию в ЛП, ПП и отдельных позиций. Результаты показывают, что лексема *seysangey* расширила свое значение и обозначает более широкое социальное пространство, включая и эмоциональное. Она также функционирует в качестве наречий «очень» (модификатор степени) и «совсем» (отрицательная полярность). В современном корейском языке *seysangey* употребляется как гибкий ДМ, указывающий на субъективность, интерсубъективность и позицию говорящего и демонстрирующий сложность эволюции языка и влияние внешних факторов на корейский лексикон и грамматику.
Ключевые слова: *грамматикализация, субъективация, интерсубъективация, маркировка позиции, seysang, seysangey*

1. Introduction

Korea has historically engaged in extensive interactions with China due to geopolitical and cultural dynamics. This enduring relationship is evident in the significant proportion of Korean vocabulary that originates from Chinese (Narrog & Rhee 2013, Rhee 2020, 2021). Norman (1988) provides a comprehensive overview of the Chinese language, including its history, structure, and dialects, highlighting the role of written texts in the spread of Chinese vocabulary to neighboring languages such as Korean. According to Rhee (2021), based on

findings by Sohn (1999: 87) and Kim (2002), approximately 60% of Korean vocabulary is of Chinese origin (Sohn 1999: 87), and in the authoritative dictionary *Phyocwun Kwuke Taysacen* (1992) by the National Institute of the Korean Language, 57.3% of its headwords are of Chinese origin (W.Y Lee 2002). Rhee (2021) further states that quantitative studies reveal a much lower percentage of Sino-Korean words used in daily life. For instance, Kim (2005) shows that only 19% of the top 100 high-frequency words are Sino-Korean. Similarly, the proportion of discourse markers derived from Sino-Korean words is very low, at 14.4% (27 out of about 188), though the exact number of discourse markers is not definitive. It is well recognized that some Chinese-origin words serve as discourse markers in various Asian languages (e.g., Rhee et al. 2021).

Sohn (1999) delves into the influence of Chinese on Korean, particularly through the borrowing of Chinese vocabulary in written form and addresses the differences between Sino-Korean vocabulary and native Korean words, highlighting how these influences have shaped modern Korean. The extensive borrowing of Chinese vocabulary into Korean primarily occurred through written texts, which differs from typical cases of lexical borrowing through colloquial registers (Norman 1988, Sohn 1999, Narrog et al. 2018, Irwin & Zisk 2019). Narrog et al. (2018) discuss the process of grammaticalization in various Asian languages, including Korean, and highlight instances where written Chinese influenced the grammatical structures of Korean, leading to the development of new grammatical elements. They emphasize how written Chinese texts facilitated the grammaticalization process, resulting in the adoption of new adverbials, numeral classifiers, and deverbal postpositions in Korean (Narrog et al. 2018). According to Rhee (2020) and Eom & Rhee (2021), the development of a discourse marker (DM) from a Sino-Korean phrasal expression is unique (see Shibasaki (2021), and Higashiizumi & Takahashi (2021) for similar cases in Japanese). Irwin and Zisk (2019) explore the borrowing of Chinese vocabulary into Korean and Japanese, focusing on how written Chinese texts served as a primary medium for these borrowings and examining the subsequent grammaticalization of some of these borrowed terms.

Recent studies (Rhee et al. 2021, Higashiizumi & Shibasaki in preparation), and this special issue volume focus on Chinese-origin words used in Japanese, Korean, Vietnamese, Thai and Chinese, exploring theoretically significant issues related to language contact. One notable example is the Korean lexeme *seysang* (世上) meaning 'the world,' which is of Chinese origin. A historical survey shows that the first appearance of *seysang* is attested in the 15th century, meaning 'the world people live in.' While *seysang* primarily translates to 'the world,' its semantic journey through the centuries has seen it adopt various meanings and functions, including its use as a discourse marker (DM) in the formation *seysangey* ('in the world'), which combines *seysang* with the locative particle *-ey*, in contemporary Korean. Despite these intriguing transformations, including shifts from concrete to

abstract meanings and grammatical functions, the development of *seysang* has not received sufficient scholarly attention.

This study addresses this research gap by examining the grammaticalization process of *seysang*, from a noun meaning 'the world' to the discourse marker *seysangey*. Using historical and contemporary corpora, this study explores how *seysang* has expanded to convey a range of pragmatic meanings, marking speaker emotions and stances. The goal is to trace these transformations and understand their implications within the broader framework of Korean language evolution. It aims to answer the following research questions:

1. How has *seysang* evolved from a concrete noun to a discourse marker in contemporary Korean?
2. What are the specific pragmatic meanings and speaker stances conveyed by *seysangey* in discourse?
3. How does the grammaticalization of *seysang* reflect broader patterns in Korean language change?

This paper is organized as follows: Section 2 presents the theoretical framework, Section 3 describes the data and methods, Section 4 provides data analysis and answers to the research questions, Section 5 discusses the findings, and Section 6 concludes the study.

2. Theoretical framework

Grammaticalization refers to "a process that transforms lexemes into grammatical elements and further increases the grammatical nature of these elements" (Kuryłowicz 1965, Lehmann 2015[1995]). This process typically involves a transition from concrete to abstract meanings and from independent lexical items to dependent grammatical elements. For instance, a noun or verb may evolve into a preposition, conjunction, or auxiliary verb over time. Heine and Kuteva (2002) describe grammaticalization as a unidirectional process in which content words (like nouns and verbs) gradually become function words (like prepositions and conjunctions), often resulting in the reduction of phonetic substance and an increase in frequency of use. The evolution of *seysang* serves as a clear example of this linguistic phenomenon, illustrating how a term denoting 'the world' has transitioned into various grammatical forms.

Previous studies have examined the asymmetry in the peripheral functions of linguistic forms, suggesting that left-periphery (LP) functions are generally linked to subjective roles, while right-periphery (RP) functions are associated with intersubjective roles (Adamson 2000, Onodera 2007, Degand 2014, Traugott 2014, Beeching & Detges 2014, Onodera & Traugott 2016). This study applies these hypotheses to analyze *seysangey* in Present-Day Korean.

3. Data and methods

The research combines a historical review of linguistic data from classical and Modern Korean texts with an analysis of contemporary usage in media and

literature. Two corpora are utilized as data sources. Historical data comes from a 15-million-word historical corpus, part of the larger 21st Century Sejong Corpus developed by the Korean Ministry of Culture and Tourism, spanning from the 15th century to the early 20th century. Contemporary Korean data is drawn from a 24-million-word Drama & Movies Corpus, developed by Min Li of Tsinghua University, comprising 7,454 scripts from films, TV dramas, and sitcoms created between 1992 and 2015. The search engine UNICONC, used to analyze both corpora, was created by Jinho Park of Seoul National University[1].

The methodology involves several key steps:

- Data collection: Gathering relevant texts from the historical and contemporary corpora to track the usage of *seysang* and its grammatical forms.
- Data analysis: Using qualitative methods to examine the contexts and semantic shifts of *seysang* in historical texts, and quantitative methods to analyze frequency and distribution in contemporary usage.
- Comparative analysis: Comparing the findings from historical and contemporary data to identify patterns of grammaticalization and the factors influencing these changes.
- Theoretical integration: Integrating the results with existing theories of grammaticalization to draw broader conclusions about the linguistic processes involved.

This study follows the periodization proposed by K. M Lee (2006[1961]) in *Kwukesakaysel* (An Outline of the History of the Korean):

- Old Korean (OK; before 918):
  Use of early Chinese character borrowing systems in inscriptions and texts from the Three Kingdoms and Unified Silla periods[2].
- Early Middle Korean (EMK; 918~1445):
  Linguistic features observed in documents and records from the Goryeo Dynasty, including the development of Chinese character borrowing systems.
- Late Middle Korean (LMK; 1446~1600):

---

[1] I extend my heartfelt gratitude to the developers of the corpora and the search program for their generosity in allowing their use for academic research.

[2] Korea's major historical periods are as follows:
- Gojoseon: circa 2333 BCE–108 BCE
- Three Kingdoms Period:
  Goguryeo: 37 BCE–668 CE
  Baekje: 18 BCE–660 CE
  Silla: 57 BCE–935 CE (Unified Silla from 668 CE)
- Unified Silla: 668 CE–935 CE
- Goryeo Dynasty: 918 CE–1392 CE, and
- Joseon Dynasty: 1392 CE–1897 CE

Written and grammatical systems found in texts after the creation (1443) and promulgation (1446) of Hangul, the Korean writing system.

- Early Modern Korean (EMoK; 1601~1893):
  Standardized Korean language usage and foreign word incorporation evident in late Joseon Dynasty literature
- Modern Korean (MoK; 1894~present):
  Changes in vocabulary and grammar observed in documents from the post-Enlightenment period to present-day newspapers and magazines.
- Present-Day Korean (PDK; 2000~present; PDK is a component of MoK):
  Emergence of new vocabulary and expressions in the internet and digital media since the early 21st century.

Romanization adheres to the Extended Yale System (Rhee 1996). Consistent with Korean historical linguistics conventions, Chinese characters in historical data are represented in capital letters using Modern Korean pronunciation. For simplicity in typography, old Korean characters are shown in a simplified form.

The study's methodology is both qualitative and quantitative. This study also examines the discourse marker *seysangey* in Present-Day Korean, focusing on its use in left-periphery (LP), right-periphery (RP), and stand-alone positions. Drawing from the Min Li Drama & Movies Corpus, we analyze contexts (positive, negative, and neutral) in which *seysangey* is used, exploring its pragmatic functions and distribution across different sentence positions. The qualitative descriptive approach is rooted in the theory of grammaticalization, which elucidates how lexical items transform into grammatical elements, providing a framework for understanding the historical development of grammatical structures. The quantitative aspect involves the analysis of frequency and usage patterns.

4. Analysis and results

As mentioned in Section 1, the noun *seysang* (世上) is composed of two Chinese lexemes. *Sey* (世) means 'human being,' and *sang* (上) means 'top.' The primary meaning of *seysang* is 'the world or the society people live in.' Initially used to denote 'the physical world,' *seysang* appeared in various philosophical and historical texts where it connoted a sense of space and existence. Over the centuries, its usage expanded to incorporate more abstract notions, reflecting changes in society and philosophy. In the following subsections, the examples are excerpted from the Sejong Historical Corpus and contemporary corpus, Min Li's Drama & Movies Corpus to illustrate the evolution of *seysang* from a noun to a discourse marker.

4.1. Seysang in Late Middle Korean

When *seysang* appeared in Korean in the 15th century, it was used as a full-fledged noun, with examples marked by possessive, locative, and topic

markers. In the 16th century, examples with accusative markers were also found. *Twusienhay* is a book that translates the works of Du Fu (杜甫), a poet from the Tang Dynasty in China (618~907). This translation was commissioned by royal decree during the Joseon Dynasty (1392–1897). The first edition was published as a printed book during the reign of King Seongjong (1481), and the reprinted edition was published during the reign of King Injo (1632). Examples (1) & (2) are taken from this book. *Seysang*'s semantic function was predominantly to denote 'the world people live in,' as shown in (1):

(1) **世上**앳   길히   비록   해   어즈러우나
    SEYSANG-ays kil-hi   pilok   hay   eculewu-na
    the world-GEN   path-NOM   nevertheless very chaotic-CONN

   내의   사롬도   또한   가지   잇나니라
   nay-uy salom-to   stohAn kAzi   isnAnila
   I-Gen   living-also   also   end-NOM   exit-DEC
   'Although the way of the world is very chaotic, my life also has an end.'

<div align="right">(1481 Twusienhay 10:3b)</div>

In example (1), the word *seysang* as a noun is employed with a genitive marker and modifies the noun *kil* 'way.' It denotes 'the world we live in,' which is the primary meaning of *seysang*.

(2) **世上**애   그듸가티   가난하니   잇디   아니하니라
    SEYSANG-ay kutuy-kAthi   kananhAn-i is-ti   ani.hA-nila
    the.world-LOC   you-like   poor.person exist-COMP not.do.-DEC
    'In the world, there is no poor person like you.'

<div align="right">(1481 Twusienhay 16:27b)</div>

In example (2), *seysang* was used with a locative *ay* (*-ey* in Modern Korean), indicating 'in the world we live in' as well.

4.2 Seysang in Early Modern Korean to Modern Korean

The transition from Early Modern Korean to Modern Korean marked significant changes in the usage of *seysang*. It began to appear as *seysangey*, a discourse marker used for emphasis and emotional expression in everyday conversation. This transformation is indicative of the lexeme's grammaticalization process.

4.2.1. Nominal form with nominal function

In Early Modern Korean, *seysang* primarily functioned as a noun, denoting the world or society people live in. Over time, its usage expanded to encompass various abstract and metaphorical meanings, reflecting broader societal and philosophical changes.

(3) 어진   덕을   닷가   한가한   대   나아갓다가
    ecin   tel–ul   taska   hAnkahAn tAi   naakas-taka
    righteous virtue-ACC cultivate leisurely time move.forward-CONN

    일천   해만의   **셰샹**을   바려   가 신션의
    ilchyen   hayman-uy   **syeysyang**-ul pAli-e   ka sinsyen-uy
    one.thousand year.just-GEN the world-ACC abandon-CONN go immortal-LOC

    올라   뎌   흰구롬을   타 하날의   니라면
    olla   tye huykwulom-ul tha hanAl-uy   nil-Amyem
    ascend that white.cloud-ACC ride sky-LOC reach-COND

    '[…] then, having cultivated virtuous deeds and during a tranquil time, if one wishes
    to abandon the world of a thousand years and become an immortal riding on white
    clouds, […]'

    (1832 *Sipkwusalyakenhay*: text 26)

The sentence in example (3) is an excerpted example from *Sipkwusalyakenhay*. This book provides Korean annotations and commentary on the first volume of *Sipkwusalyakthongko* compiled by the scholar Yeojin from the Ming Dynasty (1368–1644). The book includes translated text for each chapter, but there is no preface or colophon, so the translator is unknown. Several versions of the book are available, but the edition published in 1832 by the Gyeongsang Provincial Office (Daegu) is noteworthy for its relevance to dialectology according to *Dialectology Dictionary* (2003). Here in (3), *seysang* is used to denote the span of a person's life, specifically referencing the total duration of their earthly existence, which is contrasted with the pursuit of immortality. This shows *seysang* as referring to the entire human life cycle.

Next, the example (4) is taken from *Cywunyenchemlyeykwangik*, which is a spiritual and moral cultivation book published in 1865 by Berneux, the fourth Apostolic Vicar of the Joseon Diocese. It was created to aid Korean Catholic believers in their spiritual and moral development.

(4) 나 뵈시고   명하샤   쥬교와   교종끠
    na po-isi-ko   myeonghA-si-ya   cywukyo-wa   kyocong-kkuy
    I see-HON-CONN   command-HON-CONN   bishop-and   monks-together

    품하야   온   **셰샹**   사람을   위하야   이 쳠례와
    phwumhA-ya   on   **syeysyang** salAm-ul   wihAya   i   chyemlyey-wa
    hold-CONN   whole the.world   people-ACC for.the.sake.of this feast-and

    'Seeing me and commanding, the bishop and the monks joined together to hold this
    ceremony for all the people in the world.'

    (1865 *Cywunyenchemlyeykwangik*: text 74)

In example (4), *syesyang* "셰샹" is an orthographic variant of *seysang* "세상." In this context, *seysang* refers to the general secular society, as perceived from the perspective of the religious community, rather than the entire world where all

people live. Therefore, *seysang* in this sentence corresponds to the outside society. This usage shows how *seysang* can denote a broader societal context.

Examples (5) through (8) are taken from *Sinhakwelpo*, which is a theological magazine first published in December 1900 by missionary G. H. Jones, while stationed in Jemulpo. The magazine was published until 1904, then temporarily ceased before resuming publication in 1907, continuing until the fall of 1909.

(5)   거룩하신   텬당보좌를   떠나사   더러온   **세상**에
      kelukhA-si-n   thyentangpocwa-ul   stena-sa   teleo-n   **seysang**-ey
      holy-HON-ADN heaven.throne-ACC   leave-HON   unclean-ADN the.world-LOC
      오서서
      o.se-se
      come.HON-CONN
      '[Jesus] left the holy heaven throne, came to the unclean earthly world, and'

(1902 *Sinhakwelpo* 2:573)

In example (5), the word *seysang* refers to 'the earthly world' in contrast to 'the heavenly world.' Here, *seysang* is used to describe the "unclean earthly world" as opposed to the "holy heaven throne" that Jesus left. This usage highlights the dichotomy between the sacred and the profane, emphasizing the contrast between the heavenly and earthly realms.

(6)   복음을   전파하엿스면   이거시   참   우리의   깃분
      pokum-ul   cenphaha-yessu-myen ikes-i   cham   wuli-uy   kispwu-n
      gospel-ACC   evangelize-PST-if   it-NOM   really   we-GEN   happy-ADN
      세상이   될지라
      seysang-i   toy-l-cila
      the.world-NOM   become-FUT-DEC
      'If the gospel is evangelized, it will be really the happy world.'

(1903 *Sinhakwelpo* 3:476)

Seysang in example (6) is part of the phrase "깃분 세상," which literally translates to "happy world." In this context, *seysang* is not merely referring to the physical world but is associated with the emotional or mental state of the people. The phrase emphasizes the inner joy or happiness brought by the gospel. Thus, *seysang* here reflects the emotional transformation of the people's hearts and minds, marking a shift from a reference to the physical world to symbolizing an emotional state. When used in this sense, *seysang* is often paired with adjectives that describe human emotions, such as "야박한" (cold-hearted) or "따뜻한" (warm-hearted), further highlighting its role in conveying emotional and mental states.

(7)   가령   우리   교회로   말하야도   그리스도교   **세상**에
      kalyeng   wuli   kyohoy-lo   malha-yado   kulisutokyo   **seysang**-ey
      for.example we   church-INST   say-CONN   Christianity   the.world-LOC

행한지	일천오백년에	만일	덕국의	마틴루터가
hayngha-n-ci	ilchenopayk.nyen-ey	manil	tekkwuk-uy	mathuynlwuthe-ka
do-ADN- NOMZ	1500.years-at	if	German-GEN	Martin Luther-NOM

안낫더면
an.nas.temyen
not.come-COND
'For example, even when we talk about our church, in the 1,500 years of the Christian world, if Martin Luther from Germany hadn't appeared …'

(1903 *Sinhakwelpo* 3:399)

Seysang in example (7) refers to the Christian world, illustrating the context or environment where individuals or groups can be active. This usage shows how *seysang* can denote the societal or cultural sphere in which significant events or movements occur.

(8)	상고에	하나님끠서	**세상**을	지으시고
	sangko-ey	hana-nim-skuyse	**seysang**-ul	ci-usi-ko
	ancient.epoch-at	the.God-HON-NOM	the.world-ACC	create-HON-CONN

'At a very ancient epoch, the God created the earth and …'

(1903 *Sinhakwelpo* 3:483)

In example (8), *seysang* indicates the earth inhabited by life. This demonstrates that *seysang*, while primarily meaning 'the world inhabited by people,' can also extend to mean the physical land or the earth.

4.2.2. Nominal form with adverbial function

In transitioning from Early Modern Korean to Modern Korean, *seysang* underwent significant changes not only in meaning but also in grammatical function. While examples (3) to (8) illustrate *seysang* functioning as a noun, its role evolved further.

The next example is excerpted from *Kwenikcwungsilki*, published in 1926. It explores the distinction between the two types of individuals, depicting stories that illustrate what constitutes loyalty versus betrayal. The below example (9) is an intriguing instance where *seysang* demonstrates two different grammatical functions within a single sentence.

(9)	**세상**	몹슬	귀신들은	부귀영화로	사는	사람
	Seysang	mopsul	kwisin-tul-un	pwukwiyenghwa-lo	sa-nun	salam
	the.world	wicked	ghost-PL-TOP	wealth.and.prosperity-INST	live-AND	person

세상에	만컨마는	불상한	유리객을	이지경	속이나냐?
seysang-ey	man-khenmanum	pwulsangha-n	yulikayk-ul	i-cikyeng	soki-nanya
the.world-LOC	many-although	poor-ADN	drifter-ACC	this-NOMZ	trick-Q

'How do the very wicked ghosts trick a poor drifter this much although there are many people in the world who are in wealth and prosperity?'

(1926 *Kwenikcwungsilki* 1)

Degree modifiers are expressions used in grammar to intensify or compare the degree or intensity of adjectives or adverbs. They highlight the strength or extremity of a quality or action, often through comparison or emphasizing words like *very*, *extremely*, or *beyond compare*. In example (9), *seysang* in the first line acts as an adverbial degree modifier before the adjective 몹슬 'wicked', emphasizing the degree of wickedness by comparing it to the extremes of the world. On the other hand, the second occurrence of *seysang* underlined is used with the locative particle *-ey*, meaning 'the world where people live,' functioning as a noun.

The next example demonstrates another adverbial use of *seysang*:

(10) 사은품에   따라   백화점   매출이   왔다갔다하니
      saunphwum-ey ttala   paykhwacem   maychwul-i wasstakasstaha-ni
      freebie-to   according department store sales- NOM came.and.went-CONN

      세상   이거   되겠어?
      seysang   ike   toy-keyss-e
      the.world   it[NOM]   become-FUT-Q
      'Department store sales go up and down depending on the freebies. Is this going to work at all?'

                                                                              (2007. Drama. *Que sera sera*)

In example (10), *seysang* in a noun form functions as an adverb and reinforces the negative meaning of 'at all,' modifying the sentence, *Is this going to work?*. *Seysang* in this usage tends to become a negative polarity item. This evolution of *seysang* from a noun to an adverb illustrates its dynamic nature in the Korean language.

4.2.3. Seysangey as a discourse marker

Examples (9) and (10) showcase *seysang* in its nominal form but functioning adverbially, marking an important shift in its usage. Moreover, from example (11) onwards, we observe *seysangey*, a compound of *seysang* and the locative particle *-ey* (literally meaning 'in the world'), developing into a discourse marker (DM) with varied functions. This progression highlights the dynamic nature of *seysang* as it transitioned from a noun to a DM in Modern Korean, ultimately becoming a versatile component of Present-Day Korean.

Heine (2002) argues that there are four stages to how a linguistic expression acquires a new grammatical meaning. In the bridging context stage, there is a specific context giving rise to an inference in favor of a new meaning. The example in (11) shows a bridging stage.

(11) 내   보니   **世上**에   더   **媒人**   되엿난   이
      na-i   poni   **SEYSANG**-ey tye MAYIN   toyyes-nAn   i
      I-NOM see-CONN the.world-LOC that matchmaker become-ADN person

男家에	가	곳	소겨	니라되
NAMKA-ey	ka	kos	soki-e	nilAtoy
man.house-LOC	go	soon	cheat-CONN	say-CONN

'I see that in the world the one who became a matchmaker went to the man's house and soon cheatingly lied...'

<div align="right">(1721 Olywuncenpi: 156)</div>

The example in (11) is sourced from *Olywuncenpi*, published in 1721 during the Joseon Dynasty. This work translates Gu Jun (丘濬)'s *Olywuncenpiki* (*五輩全備記*), a comprehensive guide covering various aspects of governance, military strategy, and social organization. Gu Jun was a prominent scholar and official during the Ming Dynasty (1368–1644). Intriguingly, *seysangey* in (11) can be syntactically interpreted in two different ways. In one interpretation, *seysangey* can be understood as 'I know there exists a certain type of person in the world,' functioning as an adverbial phrase. Alternatively, it can carry a negative connotation, meaning 'on earth,' which reflects the speaker's viewpoint on the event. This dual interpretation in example (11) aligns with the concept of a 'bridging context' as described by Heine (2002), where a lexical item with dual interpretations can develop into a discourse marker, as observed in examples (12) and (13).

Examples (12) and (13) are excerpted from *Kwiuyseng*, which is a representative work of new fiction genre "sinsosel" by Injik Lee in 1908. It exposes the helplessness of the declining noble class, while simultaneously depicting the oppressed class resisting the exploitation and extortion by the ruling class. The term "sinsosel" literally means 'new novel,' which is a type of literary fiction that flourished in Korea from the late 19th century to the early 20th century. It marks the transition from classical novels to modern novels. It is also sometimes referred to as "Enlightenment Period fiction." The word *seysang* with a locative marker *-ey* in the formation of *seysangey* 'in the world' can be found in the historical corpus as in (12) and (13), with a comma.

(12) | **세상에**, | 그런 | 흉악한 | 년이 | 잇슬 | 쥴 | 누가 아라 |
|---|---|---|---|---|---|---|
| **seysangey** | kulen | hyungakha-n | nyen-i | iss-ul | cyul | nwukA al-a |
| DM | such | brutal-ADN | bitch-NOM | exit-ADN | NOMZ | who know-Q |

'Oh my gosh, who knows that there exists such a brutal bitch?'

<div align="right">(1906 Kwiuyseng)</div>

Seysangey in (12) means 'on earth' in a rhetorical interrogative from a negative viewpoint. This usage reflects the speaker's disbelief and disdain.

(13)   A: He dared to have taken our money.

B: | 응, | **셰상에**, | 참 | 우슌 | 놈, | 다 | 보갯구 |
|---|---|---|---|---|---|---|
| ung, | **syeysangey** | cham | wutyu-n | nom | ta | po.kAis-kwu |
| yes, | DM | truly | ridiculous-ADN | jerk | all | see-INTJ |

'Right, oh my gosh, how ridiculous jerk!'

<div align="right">(1906 Kwiuyseng)</div>

In (13), *syeysangey* "세상에" is an orthographic variant of *seysangey* "세상에." *Syeysangey*, meaning 'on earth' carries speaker B's negative underestimation of the guy they talk about. This usage emphasizes the speaker's negative judgment and emotional reaction. Even, stand-alone *seysangey* as a discourse marker can be found. The following example demonstrates this usage:

(14)　　A: (in a calm voice) ... I was fired!
　　　　B: what?!
　　　　A: I said I got fired. The manager said to me that I'd better work as a sale representative or canvasser than an office worker. The ones that are needed to run around outside a lot…

　　　　B: 세상에…
　　　　Seysangey
　　　　 DM
　　　　'Oh my gosh…'

<div align="right">(2000. Movie Phullantasuuy kay)</div>

In example (14), speaker A tells B that A was fired, which is surprising news to B. The use of *seysangey* by speaker B signals the speaker's stance, primarily conveying a feeling of unexpectedness. *Seysangey* in (14) functions as a DM and stands alone. As a DM, *seysangey* is independent of the sentence where it occurs. The prominent feature of discourse markers is their positional flexibility, among other characteristics. The function of *seysangey* here is to convey an emotional reaction, highlighting the speaker's surprise and engaging the listener in the conversational exchange. This example illustrates how *seysangey* operates as a DM in spoken Korean, particularly in informal settings, to convey complex emotional responses succinctly.

4.3. Seysangey as a discourse marker in Present-Day Korean

In Present-Day Korean (PDK), *seysangey* functions as a discourse marker (DM) with notable flexibility in sentence structure, appearing in left-periphery (LP), right-periphery (RP), and stand-alone positions. This analysis investigates whether *seysangey*'s usage aligns with previous hypotheses, which LP is associated with subjectivity and RP with intersubjectivity. We explore the usage patterns and contextual variations of *seysangey* across these positions, examining how it expresses subjective and intersubjective functions through various emotional stances. In the following subsections, examples are presented to illustrate *seysangey*'s pragmatic roles and the contexts in which it is preferred.

4.3.1. Left-periphery analysis

Among the tokens of *seysangey* in LP as in [seysangey, STATEMENT], 289 were used for analysis.

(15)

Context	Positive	Negative	Neutral	Total
Tokens	34 (11.8%)	205 (70.9%)	50 (17.3%)	289 (100%)

As indicated in (15), *seysangey* is the most frequently used in negative contexts, accounting for 70.9% of the cases. This high percentage suggests that speakers often employ *seysangey* to convey surprise or disbelief in response to negative information.

The next example illustrates a neutral context where *seysangey* is used. In this situation, Jongnam, a foster daughter, has been praising her mother and is now asking Jaeman how long he has known her mother:

(16)　　[*Seysangey* in LP in the neutral context]
　　　　Jongnam: By the way, since when did you know my mom?
　　　　Jaeman: ...well... I think it's been for about 40 years.

　　　　Jongnam: 세상에, 그럼 한 동네 친구셨어요?
　　　　　　　　seysangey, kulem han tongney chinkwu-si-ess-eyo?
　　　　　　　　DM so same neighborhoold friend-HON-PST-END.Q
　　　　　　　　'Oh my gosh, so you were a neighborhood friend?'
　　　　　　　　　　　　　　(2005. Drama *Pyelnan yeca pyelnan namca* Ep. 30)

In (16), *seysangey* is used by Jongnam to express surprise upon learning that Jaeman has known Jongnam's mother for about 40 years and was a neighborhood friend. This example demonstrates the use of *seysangey* in a neutral context, where the primary function is to react to unexpected information without attaching a positive or negative sentiment. It highlights the speaker's astonishment and curiosity, adding an element of engagement to the conversation.

The next example is from a situation where Hyera and Taejin are having dinner at a restaurant when Taejin handed a shopping bag to Hyera:

(17)　　[*Seysangey* in LP in the positive context]
　　　　Hyera: [...] then, I will take it happily. Can I open it?
　　　　Taejin: Sure.
　　　　Hyera: (She opens it, and it's a sky blue scarf.)

　　　　　　세상에, 너무 예뻐요. 저 하늘색 좋아하는데!
　　　　　　seysangey, nemwu yeypp-eyo. ce hanulsay cohaha-nuntey!
　　　　　　DM, so beautiful-DEC. I sky.blue like-END.INTJ
　　　　　　'Oh my gosh, it's so beautiful. I like the sky blue color!'
　　　　Taejin: Is it? (smiling) That's a relief.
　　　　　　　　　　　　　　(2006. Drama *Nay Insaynguy supheysyel* Ep.7)

Example (17) reveals that *seysangey* carries speaker Hyera's positive attitude towards the unexpected gift. Upon receiving and opening a gift, Hyera's reaction includes both her delight and astonishment at the beauty of the scarf. This shows how *seysangey* can be used to convey positive emotions and enhance the expressive quality of speech, emphasizing the speaker's emotional response.

Example (18) is excerpted from a scene on the morning of Taeja's wedding, where his mother, Chansun, finds him reeking of alcohol:

(18)   [*Seysangey* in LP in the negative context]
         (The group looks together and sees Taeja, completely disheveled, being carried on Sapal's back. Chilku follows, running with Taeja's shoes in hand. Everyone is stunned. Sapal arrives and puts Taeja down, who, drunk, collapses onto the ground.)
         Jongchil: Oppa.
         Chansun: Taeja, what's wrong with you?
         Sapal: It's because he's still sober.

         Chansun: 세상에,   이   술냄새.

                  seysangey, i   swulnaymsay.
                  DM   this alcohol.smell
                  'Oh my gosh, this smell of alcohol.'

                  What have you been doing all night with the guy going to get married?
                  Don't you know it's his wedding today?
         Chilku: He said he was having a bachelor party the night before the wedding...
                                                  (2006. Drama *Somunnan chilkongcwu* 01–80)

Seysangey in (18) indicates not only Chansun's surprise but also her negative attitude towards Taeja's condition. Considering Taeja is about to get married, the smell of alcohol elicits a strong reaction from Chansun, who uses *seysangey* to express her disapproval and shock. This example underscores the marker's function in conveying negative evaluations and critical judgments.

Overall, the analysis shows that *seysangey* in the left-periphery position can convey both subjective and intersubjective functions. For instance, in Example (17), Hyera's use of *seysangey* not only reflects her personal delight and surprise but also prompts an interaction with Taejin, demonstrating its role in engaging the listener. This suggests that *seysangey*'s functions in the left periphery are not strictly limited to the speaker's subjective experience, as it can also foster a shared understanding or emotional connection between the speaker and listener.

4.3.2. Right-periphery analysis

Next, we analyzed 40 instances of *seysangey* in RP positions, which occur at the end of a speaker's turn as in [STATEMENT seysangey]. These instances were selected to investigate the contextual preferences for RP usage of *seysangey*.

(19)

Context	Positive	Negative	Neutral	Total
Tokens	3 (7.5%)	31 (77.5%)	6 (15.0%)	40 (100%)

As (19) indicates, *seysangey* in RP is also preferred in negative contexts by 77.5% with 31 tokens among 40. This trend aligns with its usage in the LP,

suggesting that *seysangey* consistently functions to express negative sentiments or critical judgments regardless of its position within the sentence.

The given situation in (20) is that two speakers happen to meet each other by chance in a pharmacy:

(20)   [*Seysangey* in RP in the neutral context]
Junghwan: (entering) Excuse me, do you have any other ointment? This one isn't working at all.
—Manager Kang, who was drinking a tonic nearby, stares at Junghwan—
Kang: Excuse me, but are you...?
Junghwan: (turning around) Oh, Manager Kang!
Kang: (simultaneously) Oh, Assistant Han! Wow, it's been a long time. (shakes hands warmly)
Junghwan: How long has it been? I think the last time we met was at my wedding.
Kang: When did you come back from New York?
Junghwan: It's been a while. (giving Kang a business card) My office is nearby.

Kang: Is it? 세상에... (also giving a business card to Junghwan)...
> **Seysangey**
> DM
> 'Oh my gosh'

(2002. Drama *Kechimepsnun salang* Ep.3)

Kang's reaction with *seysangey* expresses that the information on Junghwan's office location is new to him. Here in (20), *seysangey* is used to convey surprise in a neutral context, without any strong positive or negative connotations; the DM here highlights the newness of the information to the speaker. This example demonstrates the marker's flexibility in expressing astonishment in various contexts.

The following example is excerpted from a scene after Byeonghee mentioned that she received a proposal from Dr. Bae, who is considered the top eligible bachelor:

(21)   [*Seysangey* in RP in the positive context]
(After Byeonghee, a daughter, mentioned she was proposed to.)
Swunam: [...] My heart is pounding.
Junhee: (laughs) Mom, why is your heart pounding?
Swunam: Exactly, how ridiculous.
Byeonghee: Mom hasn't experienced that because she got married in an arranged marriage.
Swunam: (even clapping hands) Oh, right, yeah. I'm jealous, Byunghee.

How wonderful must that be 세상에.
> **seysangey**
> DM
> 'Oh my gosh'

(2006. *Yewuya mwehani* Ep.10)

Seysangey used in RP as in (21) indicates Swunam's joy along with admiration for her daughter's situation. Swunam's reaction includes both her surprise and happiness, which she expresses with *seysangey*. This highlights the marker's role in conveying positive emotional responses and enhancing the expressiveness of the dialogue.

The next example (22) captures a reunion of old classmates at a golf clubhouse, where they are seated around a table:

(22)   [*Seysangey* in RP in the negative context]
          (As expected, Jinhee is pretending to be graceful, and Jangmi looks at Jinhee with a disapproving expression.)
          Jinhee: Nami hasn't aged at all. I recognized her right away. Jangmi, (looking at the double eyelid surgery) ...you're the same as before.
          Jangmi: (suddenly) I can't recognize you. Hey, you got your whole face done except your mouth. 세상에.
                    Seysangey
                    DM
                    'oh my gosh'
          Jinhee: What do you mean got it done? I just had a little work done on my nose because of rhinitis. (Jangmi staring intently) And a bit on my eyes because my eyebrows were poking.

                                                              (2011. Movie *Sunny*)

In (22), *seysangey* expresses Jangmi's negative judgment on Jinhee's plastic surgery. The marker emphasizes Jangmi's shock and disapproval, demonstrating how *seysangey* can convey negative emotional responses and enhance the speaker's critical stance in the conversation.

The analysis of *seysangey* in the right-periphery position similarly shows that it performs both subjective and intersubjective functions. For instance, in Example (21), Sunam's use of *seysangey* at the end of her statement conveys her personal emotional response while also inviting her daughter to share in that emotion. This indicates that *seysangey*'s discourse role is versatile and not strictly bound by its position within the sentence.

4.3.3. Stand-alone analysis

As previously stated, in our analysis, 'stand-alone' instances are defined as cases where *seysangey* is the only utterance made by a speaker during their turn in a conversation, with no additional phrases or sentences. The cases of stand-alone *seysangey* selected for analysis are 127 when it is a speaker's turn to speak.

(23)

Context	Positive	Negative	Neutral	Total
Tokens	4 (3.1%)	85 (67.0%)	38 (30.0%)	127 (100%)

Stand-alone *seysangey* in (23) is also most frequently used in negative contexts by 67% with 85 cases among 127 cases. This distribution underscores that *seysangey* as a stand-alone DM is also predominantly employed to express negative emotions or reactions.

The following dialogue is quoted from the moment when Taeyeong and Miryeong are showing gifts to their grandmother:

(24) [*Seysangey* in stand-alone position in the neutral context]
      (Taeyang and Miryeong are explaining to their grandmother what they brought as gifts for a newborn baby.)
      Miryeong: Grandma, look at this hat, a hat...
      Grandmother: You bought a good one. It will be useful when coming up to Seoul…
      Miryeong: And this is an album...
      Grandmother: You bought an album for the baby?
      Taeyeong: (smiling) Nowadays people make albums from the first day the baby is born…

      Gramma: 세상에...

      seysangey
      DM
      'oh my gosh'

                                            (2003. Drama *Nolansonswuken* Ep. 97)

In (24), DM *seysangey* used in a stand-alone position during a speaker's turn reflects Grandmother's surprise at the completely new information given.

The scene in example (25) is set in a hospital, where Mr. Kyung awakens from a coma:

(25) [*Seysangey* in stand-alone position in the positive context]
      Mr. Kyung's wife: Honey? Do you know who I am?
      Mr. Kyung: (slowly looking around, blinking his eyes)
      Mr. Kyung's wife: (holding hands and going close to the face) Honey? Can you see me? Do you know who I am?
      Mr. Kyung: (nodding still)
      Mr. Kyung's wife: 아우, 세상에.

      Awu, **seysangey**
      INTJ, DM
      'Oh! my gosh'

                                            (2005. Drama *Pwuhwal* Ep.18)

In (25), Mr. Kyung's wife's use of *seysangey* reflects her surprise and relief when Mr. Kyung regains consciousness. This stand-alone use highlights *seysangey*'s ability to convey a mix of emotions, enhancing the depth of the speaker's immediate reaction. Although "awu" is an interjection expressing an initial reaction of surprise, the phrase "awu, seysangey" is considered standalone because "awu" serves merely as an interjection and does not affect the discourse marker function of *seysangey*.

The following example is excerpted from a conversation between Eunha, who is about to undergo surgery, and her acquaintance Seoyoung in the hospital room, where Eunha inquires about the well-being of their mutual acquaintances:

(26) [*Seysangey* in stand-alone position in the negative context]
Eunha: [...] Except for Woojin... Is everyone... okay? How about Dr. Jeong?
Seoyoung: (expression changes)
Eunha: (looking worriedly at Seoyoung) What's wrong?
Seoyoung: ... Dr. Jeong passed away. ...it was suicide...

Eunha: 세상에.

seysangey
DM
'Oh my gosh'

(2003. Drama *Lepuleythe* Ep.11–16)

In (26), Eunha's utterance of *seysangey* conveys her shock and sadness upon hearing about Dr. Jeong's suicide. This example underscores the marker's role in expressing strong negative emotions, serving as an immediate and powerful reaction to distressing news.

The stand-alone analysis further confirms that *seysangey* carries both subjective and intersubjective functions regardless of its syntactic position. When used alone, *seysangey* often expresses the speaker's immediate emotional reaction while also drawing the listener into this emotional state. This observation reinforces the idea that *seysangey*'s discourse functions are not confined to a single category (subjective or intersubjective) nor strictly tied to its position within a sentence but rather operate flexibly across different contexts.

In sum, a distributional analysis of *seysangey* shows that it is preferred in negative contexts. However, contrary to the theoretical expectations that LP and RP are distinctly associated with subjective and intersubjective functions respectively, our findings suggest that *seysangey* consistently exhibits both subjective and intersubjective characteristics across different positions, including LP, RP, and stand-alone usage. The data suggests a more complex interaction between context and peripheral functions, emphasizing that *seysangey*'s discourse roles are adaptable and not rigidly determined by its sentence position.

4.4. Answers to the research questions

The study addresses the evolution of *seysang* from a concrete noun meaning 'the world' into a versatile discourse marker, *seysangey*, in contemporary Korean. This transformation illustrates the grammaticalization process whereby *seysang* shifted from a noun describing a physical realm to a discourse marker that expresses various speaker stances, including surprise, admiration, and disapproval.

To answer the first research question, *seysang*'s journey from a noun to a discourse marker involved shifts from describing the external world to marking subjective reactions, aligning with general grammaticalization patterns.

In response to the second question, *seysangey* serves multiple pragmatic functions and conveys stances reflecting emotional and evaluative responses, depending on context. The marker's flexibility—appearing at the left periphery (LP), right periphery (RP), or as a stand-alone utterance—enhances its expressive function and often invites listener engagement.

Lastly, regarding the third question, the grammaticalization of *seysang* reflects broader patterns of language change in Korean. This process, involving subjectification and intersubjectification, aligns with observed trends in other East Asian languages where nouns and concrete expressions evolve into pragmatic markers that convey speaker attitudes and interactional stances.

5. Discussion

The developmental course of *seysang* and its derivative *seysangey* as a DM in Korean reveals theoretically notable aspects: decategorialization, subjectification, intersubjectification, stance-marking, and metaphoric transfer.

5.1. Decategorialization

Examples (9) and (10) illustrate that *seysang* has experienced significant semantic and functional changes over time, transitioning from a primarily nominal role to include adverbial functions as well. The transition of *seysang* from a nominal word to a degree modifier as an adverb in (9) and to an NPI in (10) suggests a reduction in the nominal properties characteristic of *seysang* over time, a process known as decategorialization (cf. Hopper 1991). Historically, *seysang* was nominally oriented, frequently accompanied by case-marking particles such as *-i*, *-ey*, and *-ul*, as well as the topic marker *-un*, all of which are typically attached to nominal expressions. Over the centuries, these nominal properties have gradually diminished in specific contexts. While *seysang* still retains its nominal form and meaning in many uses, in examples (9) and (10) it has undergone decategorialization and functions as an adverb.

As evidence of decategorialization, *seysang* in its noun form is no longer able to be modified by adjectives, as is typical for nouns. Consequently, *seysang* in its adverbial role does not allow the modification of adjectives, as seen in the hypothetical construction 나쁜* 세상 몹쓸 (bad* seysang wicked). Similarly, as an NPI, *seysang* functions as an adverb and cannot be modified by adjectives, resulting in ungrammatical constructions like 아름다운* 세상 이게 되겠어? (beautiful* seysang is this going to work?).

5.2. Subjectification, intersubjectification and stance marking

The transformation of *seysang* into DM *seysangey* illustrates its grammaticalization from a noun through an adverb to a discourse marker. This process involved both subjectification, where the lexeme acquired more subjective meanings, and intersubjectification, where it began to play a role in managing interpersonal relations.

Subjectification refers to the process by which linguistic expressions become increasingly based on the speaker's internal perspective, attitudes, or beliefs (Traugott 1982, 1995, Traugott & König 1991). *Seysang*, a society people live in, is a physical domain for a speaker to live and experience things, so it naturally becomes the epistemic domain for a speaker to judge an event. *Seysang* in (6) means that the people will feel happiness in their hearts. This indicates a shift from a physical world to an emotional state. In addition, *seysang* is employed to reflect the speaker's subjective stance toward the intensity or degree of a property. It has the interpretive meanings of 'very' in modifying adjectives as in (9) and 'at all' as in (10) where an anticipated event does not happen in a way a speaker wants. Through the subjective evaluation of an event or state of affairs, this transfer is made from the physical entity to the abstract entity or subject evaluation.

Intersubjectification, as explained by Traugott (2008), involves the speaker's awareness of the addressee's attitudes and beliefs. This concept emphasizes the interactive nature of communication and the shared understanding that arises from it. Du Bois's (2007) stance triangle further elucidates how speakers use language to position themselves, align with or against others, and evaluate objects of discourse. The stance triangle consists of three components: stance-taking, alignment, and evaluation. When the speaker uses *seysangey*, the speaker is simultaneously evaluating the event, positioning themselves about the event, and aligning with or against the interlocutor's statement. This multifaceted process is exemplified in the usage of *seysangey* to convey not only surprise but also an implicit judgment or reaction to the state of events, as seen in 4.3.

Stance marking is evident in the degree modifier function of *seysang,* as the speaker uses *seysang* to emphasize the degree or intensity of quality, reflecting the awareness of the addressee's attitudes and beliefs. As shown in 4.3, *seysangey* as a DM serves to highlight the notable and surprising value of the information given, meaning something like 'I can't believe it happens in the world'. When the speaker uses *seysangey,* the speaker is showing his/her interpretive judgment on the given statement by an interlocutor or state of events the speaker faces, exemplified in 4.3.

According to Rhee (2011: 405), stance markers can be categorized into attitudinal, epistemic, emotional, and evidential types. The use of *seysangey* in section 4.3 aligns with the emotional stance category, where the marker reflects the speaker's positive, negative, or neutral emotions toward a proposition or event. For example, in negative contexts, *seysangey* often conveys shock or disapproval, reflecting a negative emotional stance. In positive contexts, it may express delight or amazement, aligning with a positive emotional stance. In neutral contexts,

seysangey can indicate mere surprise, showing its role in marking a speaker's emotional response without strong evaluative connotations. This categorization highlights *seysangey*'s function as a versatile stance marker, capable of conveying a range of emotional reactions across different contexts.

5.3. Metaphorical transfer

Heine et al. (1991) have characterized a unidirectional metaphorical shift as "PERSON > OBJECT > PERFORMANCE > SPACE > TIME > QUALITY." Each category represents a domain of conceptualization, and any category to the right is used to conceptualize those to the left.

The source lexeme (世) *sey* of *seysang* denotes 'human' and the meaning of 'human' in the physical domain. So, *seysang* is metaphorically likened to the meaning of human-related concepts such as one's life (PERSON > TIME) and people's hearts (PERSON > QUALITY). However, *seysang* itself, meaning 'the world,' can also be seen as SPACE. There are meaning changes from the lexical meaning to a degree modifier ('very'), negative polarity item (NPI, 'at all'), and an emphatic marker as a DM, expressing surprise or disbelief (SPACE > QUALITY). The semantic change of *seysang* through the metaphorical transfer in their meanings can be schematized as below:

(26)   PERSON > SPACE   > TIME   >   QUALITY
       seysang   ┌ society (inclusive   lifetime   ┌ people's hearts
                    │   or exclusive)   │ degree modifier ('very')
                    └ on the earth   │ NPI, 'at all'
                                                          └ emphatic marker as DM
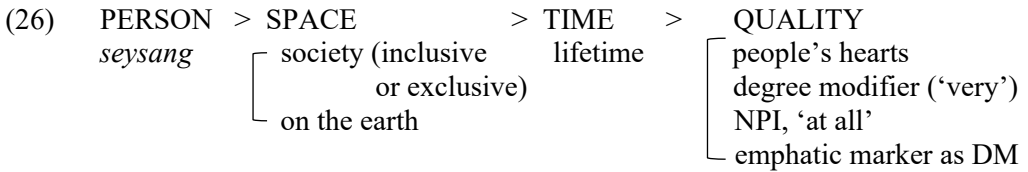
The term *seysang* originally refers to 'the world where people live' in a broad sense, which includes all of society. In some contexts, it is used to describe the 'outer world' as opposed to secluded environments like temples, monasteries, or prisons. Thus, 'exclusive' in the context of *seysang* can be interpreted as 'outer world.' This distinction helps to clarify the different nuances of the term in various metaphorical and grammatical contexts.

6. Conclusion

This research highlights the semantic and functional evolution of *seysang* from a concrete noun to a versatile discourse marker in Korean, using grammaticalization as a framework for qualitative analysis. The study integrates corpus data to quantitatively explore *seysangey*'s usage as a discourse marker in contemporary Korean.

A historical survey shows that the first appearance of *seysang* is attested in the 15th century. The word consists of two logographs: *sey* (世) meaning 'human beings' and *sang* (上) meaning 'top.' Initially, *seysang* had a concrete meaning, but

over time, its meaning extended to multiple abstracted meanings, including intensifiers such as 'very' and 'at all.'

In the course of semantic change, *seysang* and its derivative *seysangey* exhibit subjectification, shifting from a 'physical' to an 'abstract' meaning. These terms mark the speaker's stance, encompassing both positive and negative judgments, and express a wide range of emotions such as joy, sadness, anxiety, and disappointment. *Seysangey* also exhibits intersubjectification, indicating that the speaker signals the addressee to share their evaluation of the situation. *Seysangey* (literally meaning 'in the world') as a DM tends to appear freely in left-periphery (LP), right-periphery (RP), and even in stand-alone positions, often signaling counter-expectation or surprise. The survey reveals that it carries subjective and intersubjective meanings and is frequently used in contexts of both negative and positive emotional responses.

Future research will explore related expressions such as *seysangepsi* (세상없이) 'without the world,' *seysangeupseto* (세상없어도) 'even without the world,' and *seysangmoluta* (세상모르다) 'not know the world,' all of which are related to *seysang* 'the world.'

The study offers insights into the dynamics of language change influenced by past contact with a dominant language like Chinese. Such investigations can enhance our appreciation of the complexities involved in language evolution and the role of external influences in shaping the lexicon and grammar of a language.

Acknowledgement

Abbreviations

ACC = accusative; ADN = adnominal; COND = conditional; CONN = connective; DEC = declarative; DM = discourse marker; END = sentence-ender; FUT = future; GEN = genitive; HON = honorific; INTJ = interjection; INST = instrument; LOC = locative; NOM = nominative; PL = plural; PST = past; Q = question; TOP = topic

References

Adamson, Sylvia M. 2000. A lovely little example: Word order options and category shift in the premodifying string. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English,* 39–66. Amsterdam: Benjamins.
Beeching, Kate & Ulrich Detges. 2014. Introduction. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change,* 1–23. Leiden: Brill.
Degand, Liesbeth. 2014. So very fast then: Discourse markers at left and right periphery in spoken French. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left*

and right periphery: Crosslinguistic investigations of language use and language change, 151–178. Leiden: Brill.

Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction,* 139–182. Amsterdam & Philadelphia: John Benjamins.

Eom, Sujin & Seongha Rhee. 2021. The rise and fall of a discourse marker: The case of *kisil* in Korean. *East Asian Pragmatics* 6 (3). 381–401.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 83–101. Amsterdam: Benjamins.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Higashiizumi, Yuko & Keiko Takahashi. 2021. The emergence of the discourse marker *sonojitsu* 'as a matter of fact' in Japanese. *East Asian Pragmatics* 6 (3). 355–379

Irwin, Mark & Matthew Zisk. 2019. *Japanese Linguistics.* Tokyo: Asakura Publishing.

Kim, Hansaem. 2005. *Hyentaykwuke sayong pinto cosa 2* (*A Frequency Analysis of Modern Korean 2*). Seoul: National Institute of the Korean Language.

Kuryłowicz, Jerzy. 1965. The evolution of grammatical categories. *Diogenes* 13 (51). 55–71.

Lee, Ki-Moon. 2006 [1961] *Kwukesakaysel* (An Outline of the History of the Korean). Seoul: Thaehaksa.

Lee, Woon-Young. 2002. *Phyocwun kwuke taysacen yenkwu pwunsek* (A study on *A complete standard Korean dictionary*). Seoul: National Institute of the Korean Language.

Lehmann, Christian. 2015[1995]. *Thoughts on Grammaticalization*. Berlin: Language Science Press.

Narrog, Heiko & Seongha Rhee. 2013. Grammaticalization of space in Korean and Japanese. In Martine Robbeets & Hubert Cuyckens (eds.), *Shared grammaticalization with special focus on the Transeurasian languages*, 287–315. Amsterdam: Benjamins.

Narrog, Heiko, Seongha Rhee & John Whitman. 2018. Grammaticalization in Japanese and Korean. In Heiko Narrog & Bernd Heine (eds.), *Grammaticalization from a typological perspective,* 166–188. Oxford: Oxford University Press.

Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.

Onodera, Noriko O. 2007. Interplay of (inter)subjectivity and social norm. *Journal of Historical Pragmatics* 8 (2). 239–267.

Onodera, Noriko O. & Traugott Elizabeth C. 2016. Periphery: Diachronic and cross-linguistics approaches. *Journal of Historical Pragmatics* 17 (2). 163–177.

Phyocwun kwuke taysacen. 1992. *(A complete standard Korean dictionary).* Seoul: The National Institute of the Korean Language.

Pangenhaksacen. 2003. *(Dialectology Dictionary).* Seoul: Thaehaksa.

Rhee, Seongha. 1996. *Semantics of verbs and grammaticalization: The development of Korean from a cross-linguistic perspective.* Austin, TX: The University of Texas at Austin dissertation. Seoul: Hankook Publisher.

Rhee, Seongha. 2011. Nominalization and stance marking in Korean. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives*, 393–422. Amsterdam: John Benjamins Publishing Company.

Rhee, Seongha. 2020. From object to text to stance: The case of *kyelkwa* in Korean. 日本認知言語學會論文集 第 20 卷 (*Papers from the 20th National Conference of the Japanese Cognitive Linguistics Association* [JCLA] 20). 471–476.

Rhee, Seongha. 2021. From truth to reality to effect: The journey of *sasilsang* in Korean. *East Asian Pragmatics* 6 (3). 331–353.

Rhee, Seongha, Reijirou Shibasaki & Xinren Chen (eds.). 2021. Special issue, Grammaticalization of discourse markers in East Asian Languages. *East Asian Pragmatics* 6 (3).

Shibasaki, Reijirou. 2021. Discourse markers in the making: Pragmatic difference of *jijitsujoo* from *jijitsu* in Modern through Present Day Japanese. Special issue, Grammaticalization of discourse markers in East Asian Languages. *East Asian Pragmatics* 6 (3). 303–329.

Higashiizumi, Yuko & Reijirou Shibasaki (eds.) in preparation. *The Emergence of Pragmatic Markers from Chinese Compounds in Chinese, Japanese and Korean: Perspectives from East Asian Languages and beyond*. Leiden: Brill.

Sohn, Ho-min. 2001[1999]. *The Korean Language.* Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics,* 245–271. Amsterdam: John Benjamins.

Traugott, Elizabeth C. 1995. Subjectification and grammaticalisation. In Dieter Stein & Susan Wright (eds.), *Subjectivity and subjectivisation: Linguistic perspectives,* 31–54. Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 2008. Intersubjectification and Grammaticalization. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 19–50. Amsterdam: John Benjamins.

Traugott, Elizabeth C. 2014. Intersubjectification and clause periphery. In Lieselotte Brems, Lobke Ghesquière & Freek Van de Velde (eds.), *Intersubjectivity and intersubjectification in grammar and discourse,* 7–27. Amsterdam: John Benjamins.

Traugott, Elizabeth C. & Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to grammaticalization* vol 1, 189–218. Amsterdam: John Benjamins.

Bionote:
Hyun Sook LEE is a faculty member at Jangan University. She received her PhD from Hankuk University of Foreign Studies in 2011, Republc of Korea. Her research interests include Corpus Linguistics, Discourse Analysis, Semantics and Pragmatics, with a particular focus on the processes of grammaticalization.
e-mail: morningdew1020@jangan.ac.kr
https://orcid.org/0009-0004-6712-3141

Сведения об авторах:
Хён Сук ЛИ преподает в Чанганском университете. Она получила степень PhD в Университете иностранных исследований Ханкук в 2011 г., Республика Корея. Ее научные интересы включают корпусную лингвистику, анализ дискурса и прагматику, с особым акцентом на процессах грамматикализации.
e-mail: morningdew1020@jangan.ac.kr
https://orcid.org/0009-0004-6712-3141

Research article / Научная статья

From 'one morning' to a discourse marker:
The case of *iltan* in Korean

Sunhee YAE ⓘ ✉

Chung-Ang University, Seoul/Anseong, Republic of Korea
✉syae@cau.ac.kr

Abstract
The goal of the paper is to identify the developmental pathways of *iltan* shaped by language contact with China. The study explores how the Sino-Korean term *iltan* evolved from a temporal noun into a discourse marker with distinctive functions. The study uses the perspectives of grammaticalization and pragmatics and employs historical and contemporary Korean corpora. *Iltan* originally meant 'one morning' and was used as a temporal noun in Late Middle Korean. However, its original meaning and noun status only persisted in Early Modern Korean until the 18th century. Its meaning 'one morning' was taken over by another Sino-Korean word, *ilco*, and purely Korean terms *halwu achim* and *enu nal*, which are all attested in the 15th century. As *iltan* shifted from a noun to an adverb, a connective, and a discourse marker, it acquired new meanings and functions associated with priority, short duration, and conditionality in Early Modern Korean. In Present-Day Korean, *iltan* is strategically used for multiple purposes, such as presenting topics and sub-topics, establishing common ground, filling pauses, indicating hesitation, and creating interruptions, thereby exerting its scope over a larger portion of the conversation. This transformation of *iltan* illustrates processes such as grammaticalization, discourse grammar, cooptation, (inter)subjectification, peripheral asymmetry, and source characteristics in comparison with the Chinese equivalent *yīdàn*. The findings provide insights into the Source Determination Hypothesis and its pragmatic implications.
Keyword: *discourse marker, grammaticalization, discourse grammar, (inter)subjectification, iltan, Korean*

От «одного утра» до дискурсивного маркера:
iltan в корейском языке

Сунхи ЯЭ ⓘⒹ ✉

Университет Чунг-Анг, Сеул/Ансон, Республика Корея
✉ syae@cau.ac.kr

Аннотация
Цель статьи – определить пути развития корейского слова *iltan*, происходящего под влиянием контакта с китайским языком. В ней исследуется, как китайско-корейская лексема *iltan* превратилась из существительного в дискурсивный маркер с отличительными функциями. Материалом послужили данные исторических и современных корейских корпусов, которые анализировались с позиции грамматикализации и прагматики. Лексема *Iltan* первоначально имела значение «одно утро» и использовалась как временно́е существительное в позднесреднекорейском языке. Однако ее первоначальное значение и статус существительного сохранялись только в раннем современном корейском языке вплоть до XVIII в. Значение «однажды утром» было заимствовано другим китайско-корейским словом *ilco* и чисто корейскими словами *halwu achim* и *enu nal*, которые были засвидетельствованы в XV в. По мере того как лексема *iltan* превращалась из существительного в наречие, соединительный элемент и дискурсивный маркер, в раннем современном корейском языке она приобретала новые значения и функции, связанные с приоритетом, кратковременностью и обусловленностью. В современном корейском языке *iltan* используется для различных стратегических целей, таких как представление тем и подтем, установление общей точки зрения, заполнение пауз, выражение сомнения и др., тем самым охватывая большую часть дискурса. Данная трансформация *iltan* иллюстрирует процессы, касающиеся грамматикализации, грамматики дискурса, кооптации, (интер)субъективации, периферической асимметрии. Также она позволяет охарактеризовать источник – китайское слово *yīdàn* (一旦). Результаты исследования дают более глубокое представление о гипотезе детерминации источника и ее значении для прагматики.
Ключевые слова: *дискурсивный маркер грамматикализация, грамматика дискурса, (интер)субъективация, iltan, корейский язык*

1. Introduction

Korea has maintained long-standing interaction with China, mainly for geopolitical and cultural reasons, which has resulted in a significant portion of Korean vocabulary being derived from Chinese sources (Narrog & Rhee 2013, Lee 2021, Rhee 2021). The Sino-Korean term, *iltan* (一旦), composed of *il* (一) meaning 'one' and *tan* (旦) meaning 'morning,' has experienced noteworthy semantic and functional transformations. It was first recorded in the 15ᵗʰ century

with the meaning 'one morning' or 'in the morning of one day.' Over time, the lexeme *iltan* has changed both semantically and functionally, with its original meaning of 'one morning' becoming obsolete in modern Korean. Instead, it has evolved into an adverb that can mean 'first of all', 'for a while,' or serve as a conditional connective such as 'once,' or 'in case.' Its later meanings have enabled it to function as a discourse marker (DM) with multiple roles in contemporary Korean. Despite the intriguing nature of these shifts, the lexeme has rarely been studied to date, and this research aims to address that gap using data from historical and contemporary corpora.

The goal of the paper is to identify the developmental pathways of *iltan* shaped by language contact with China. To achieve this, the paper has three main objectives: (i) to explore the diachronic development of *iltan* along with its semantic and functional extensions; (ii) to compare its source characteristics with the Chinese lexeme *yīdàn* induced by language contact; and (iii) to analyze the various functions of *iltan* as a DM in Present-Day Korean (PDK), focusing on the mechanisms of cooptation, (inter)subjectification and peripheral positionality.

To meet these research goals, the paper is structured as follows: Section 2 gives a theoretical background; Section 3 describes the corpus data and methodology; Section 4 accounts for the diachronic development of *iltan*; Section 5 presents its theoretical approach to *iltan*; Section 6 concludes with a summary of the findings.

Korean is classified as a head-final (SOV) language with its agglutinative morphological structure. In everyday conversation, subjects and objects are often dropped when context provides enough information for understanding. Verbs placed at the end of sentences are marked with a variety of morphological indicators reflecting TAM, epistemicity, honorifics, and evidentiality, etc. In terms of orthography, Korean employs a system of inter-lexical spacing, meaning that a word with its morphological elements is separated from another with a white space.

2. Theoretical background

The theoretical framework of this paper includes (i) discourse grammar and (ii) peripheral asymmetry hypothesis. In (i), Heine et al.'s (2017, 2021) dual-level discourse grammar distinguishes between sentence grammar, which organizes propositional concepts, clauses, and their combinations, and thetical grammar, which deals with theticals[1]—discourse units with metatextual functions directly anchored in the discourse context (Heine et al. 2021: 72). Cooptation is a short phase when the expression concerned is transferred from the sentence level to the metatextual level of discourse (Heine et al. 2021: 76).

The emergence of DMs can be explained through the interplay of cooptation and grammaticalization in discourse grammar. The grammaticalization of DMs is

---

[1] Theticals are referred to as parentheticals. The grammatical properties of theticals have been covered in the comprehensive literature regarding parentheticals. However, researchers associate a variety of concepts with the term parentheticals. For more information, see Heine et al. (2021: 73).

categorized into two phases: 'early grammaticalization' which occurs before cooptation, and 'late grammaticalization,' which takes place after cooptation. Early grammaticalization typically involves a transition from lexical to grammatical forms. Late grammaticalization involves the gradual grammaticalization of the coopted unit (with DMs primarily being a subgroup of formulaic theticals[2]) after cooptation; by this phase, the coopted unit has evolved into a DM with a broader range of metatextual functions.

Several scholars, including Beeching & Detges (2014), Traugott (2014a & 2014b), Sohn (2014), Rhee (2016, 2020b, & 2021), and Lee (2021), have examined (ii) the roles of the left-periphery (LP) and right-periphery (RP) in DMs. A common hypothesis among researchers is that LP functions are related to 'dialogual' and subjective roles, while RP functions are connected to 'dialogic' and intersubjective roles.[3]

Based on the approaches of discourse grammar and peripheral positionality, this study aims to elucidate the phases of *iltan*'s grammaticalization, cooptation, and pragmaticalization, classifying its subjective and intersubjective functions as a discourse marker. Insights from discourse grammar and peripheral positionality offer theoretical implications for how *iltan*, as a borrowed term, has evolved uniquely from its original form.

3. Corpus data and methods

The Korean corpora used in this study include: (i) The historical corpus, consisting of 15-million-word texts, provides diachronic data from the historical section of the 21st Century Sejong Project (1998–2006). The texts in the historical section date from 1446 through 1913. (ii) The Drama and Movie Corpus, a contemporary collection of 24 million-words comprising 7,454 scripts from dramas and movies produced between 1992 and 2015, compiled by Min Li from Seoul National University. For data retrieval, a concordance software UNICONC, developed by Jinho Park, was utilized.[4]

The Korean alphabet Hankul *iltan* and the Chinese character 一旦 are identified in historical sources, with attestations retrieved through concordance searches using UNICONC. This software is also employed to gather concordance data from the Drama and Movie Corpus. Each concordance hit is reviewed individually to confirm relevance, excluding homonyms and unrelated lexemes.

---

[2] Heine et al. (2021: 74–75) group theticals into three types: instantaneous theticals, constructional theticals, and formulaic theticals. Formulaic theticals include formulae of social exchange (e.g., *Good morning!*, *hello!, please*) and interjections (e.g., *boy!, fuck!, hell!*). Another example of formulaic theticals is DMs, such as *anyway, however, indeed, in fact, instead, I mean, now, okay, so, still, then, I think, well, what else, you know*.

[3] 'Dialogual' refers to two or more speakers, 'dialogic' to two or more view-points (Onodera and Traugott 2016: 166).

[4] I would like to express my gratitude to the developers of these datasets and the search tool for their generosity in allowing their use for academic research.

4. Analysis and results: semantic and functional shift of *iltan*

This section deals with diachronic changes of *iltan* in terms of semantic and functional extensions. This section also deals with diverse functions of *iltan* as a DM. Transliteration follows the Extended Yale System (Rhee 1996). Additionally, periphrastic forms are denoted with a period to signify word boundaries. When multiple English words in the gloss represent a single morpheme in the examples, they are separated by periods. In accordance with the general convention in Korean historical linguistics, any Chinese characters from the source texts are presented in uppercase letters using Modern Korean pronunciation.

4.1. *Iltan* in Late Middle Korean

The word *iltan* is made up of *il* (一), meaning 'one,' and *tan* (旦), meaning 'morning,' originally forming a noun. The term *iltan* was initially recorded in 1496 in Late Middle Korean (LMK)[5], as shown in example (1). *Iltanay*, composed of the noun *iltan* with the locative postposition *-ay,* is placed in a non-argument position, where it functions as an adverb meaning 'one morning' or 'in the morning of one day.'

(1) 南遯十餘年이리시니   一旦–애
    NAM-TWUN-SIP-YE-NYEN-ili-si-ni   ***ILTAN-ay***
    south-hide-ten-more.than-year-be-HON-CAU   one.morning-LOC
    以非風幡動之機로   觸–開–印宗–正眼–하시니
    I-PI-PHWUNG-PEN-TONG-CI-KI-lo   *CHOK-KAY-INCONG-CENG-AN-hA-si-ni*
    with-NEG-wind-flag-move-of-occasion-INST touch-open-[name]-true-eye-do-HON-CONN
    'When it was for more than 10 years since (the Master Hyenung) was hiding in the south, then one morning, he (helped the monk Injong) open his true eyes with a (profound) understanding that the flag did not wave in the wind, and...'

                                    (1496, *Yukcotankyeng* Preface 1b3-5)

Example (1) is taken from a 15th century text, *Yukcotankyeng*, which is a collection of sermons by a Buddhist master of the Tang Dynasty, Yukcohyeynung (639–713). They were recorded by his pupil Hathayksinhoy. Example (1) describes the event that occurred to the Buddhist master one morning: not a single flag waved in the wind. The adverb *iltanay* in (1) designates 'one morning' with an indefinite reference to 'one (一).'

Another attestation in LMK is dated to 1569, titled *Senkakwikam* as exemplified in (2). *Senkakwikam* was written by Chenghe Hyuceng, a great Buddhist monk of the Joseon Dynasty.

---

[5] The following periodization is adopted in this article: Old Korean (OK; –917), Early Middle Korean (EMK; 918–1445), Late Middle Korean (LMK; 1446–1600), Early Modern Korean (EMoK; 1601–1893), Modern Korean (MoK; 1894–present), and Present-Day Korean (PDK; 2000–present; note that PDK is a subset of MoK).

(2) 古一道人이 一旦 定中에 보니 一孝子이 주검을 소내 바다와 울며 닐오대
   KO-IL-TOIN-i   ***ILTAN***   *CENGC-WUNG-ey*   *po-ni*
   old.days-one-Taoist-NOM one morning   meditation-in.the.middle.of-LOC   see-CAU
   IL-HYOCA-i   *cwukem-ul*   *son-ay*   *patao-a*
   one-filial.child-NOM   dead.body-ACC   hand-LOC   hold.up.and.come-CONN
   wul-mye   *nil-otAy*
   cry-CONN   say-CONN
   'While practicing a meditation one morning, an old Taoist ascetic witnessed a filial child holding up a (his mother's) dead body and saying…'

   (1569, *Senkakwikam* #82)

Example (2) is a story about an ascetic Taoist. It depicts what happened to him one morning: an old Taoist ascetic witnessed a child crying with a dead body in hands. In comparison to the 15th century example (1), *iltan* can function as an adverb without the locative postposition *-ay*, meaning 'one morning' in the 16th century. In LMK, *iltan* changed its grammatical status from a noun to an adverb through zero-derivation, serving as an indefinite time marker meaning 'one morning.' This usage is primarily found in Buddhist scriptures, as illustrated in examples (1) and (2).

4.2. *Iltan* in Early Modern Korean

Significant linguistic developments involving *iltan* occurred in the 18th century in Early Modern Korean (EMoK). At first, the meaning of 'one morning' and its classification as a noun continued until the 18th century, co-existing with the adverb *iltan*, as illustrated in (3).

(3) 一旦之決이 과급(過急)하믈 깨다라
   ILTAN-CI-KYEL-i   *KWAKUP-hA-m-ul*   *skAytal-a*
   one morning-of-decision-NOM   impatience-do-NOMZ-ACC   realize-CONN
   '(He) realized that the decision made one morning was hasty.'
   (17??, *Wanwelhoymayngyen* #18240)[6]

In example (3), *iltanci*, formed from the noun *iltan* and the postposition *-ci* 'of', acts as a modifier for the noun *keyl*. *Iltancikeyl* translates to 'decision of one morning,' which can also be paraphrased as 'the decision made one morning.' The excerpt in (3) suggests a sense of regret for making a decision quickly, without taking enough time to reflect on it.

Second, *iltan,* combined with connectors in the 18th century, evolved from being purely lexical in LMK to being used in textual contexts in EMoK. Example (4) illustrates the combined meanings of *iltan* 'one morning' with a conditional connective *-myen*. This context serves as a bridging context (Heine 2002) or a

---

[6] *Wanwelhoymayngyen* is an epic novel, which is estimated to be written in the 18th century, but the exact year of publication is unknown.

critical context (Diewald 2002), where pragmatic ambiguity leads to grammaticalization of *iltan* from an adverbial to a conditional connective.

(4) 而**一旦**豁然貫通焉하면
     *I-**ILTAN**-HWALYEN-KWANTHONG-EN-hA-myen*
     CONN-one.morning-definitely-be.spiritually.awakened-CONN-do-if
     (a) 'and if (you) are definitely awakened one morning…,' (b) 'and if (you) are definitely awakened one day…,' (c) 'and once (you) are definitely awakened…'

     (1749, *Tayhakyulkokenhay* #24)

In (4) *iltan* is ambiguous in its meaning and function when combined with the conditional connective *-myen*. In EMoK, it is notable that the term *iltan*, originally meaning 'one morning' or 'one day' through semantic extension. Here, the indefinite meaning of *il* ('one') in *iltan* becomes salient, while the meaning of *tan* ('morning') recedes. Combined with the conditional connective *-myen* 'if,' *iltan…myen* developed into a conditional connective meaning 'once' or 'in case.' The developmental path of combined meanings and functions of *iltan* and *-myen* is depicted in (5). Generalization (Bybee et al. 1994) operates on the semantic extension between 'one morning' in (5a) and 'one day' in (5b) Ultimately, the meaning of time becomes opaque, leading to a conditional connection expressed by phrases like 'once' or 'in case' in (5c).

(5) a. if…one morning (temporality): 'if (you) are definitely awakened one morning…'
     b. if…one day (temporality): 'if (you) are definitely awakened one day…'
     c. once / in case (conditionality): 'in case (you) are definitely awakened…'

Therefore, the meaning of *iltan* extended from 'one morning' (temporality) to 'once' or 'in case' (condition).[7] In EMoK *iltan* is grammaticalized from a time-designating adverb to a conditional connective.

Third, the Hankul (Korean alphabet) version of *iltan* is recorded during this period, as demonstrated in example (6). Fourth, the new meanings 'for a while' and 'first of all' for the adverb *iltan* emerged in the 18th century. Example (6) is taken from *Meynguylok* in the 18th century. The adverb *iltan* means 'for a moment,' 'for the present,' 'for now,' or 'for a while,' indicating brevity in time.

(6) 후겸과 닌한이 비록 스사로 규각이 이시나 **일단** 날을 원망하난 마암은 두
     역적이 한 심쟝인
     Hwukyem-kwa Ninhan-i   pilok   susAlo   kyukak-i   isi-na
     [name]-and   [name]-NOM although themselves discord-NOM exist-but

---

[7] Let us look at the source of the conditional connective *-umyen*. Koo & Rhee (2001: 5) argue that *umyen* breaks into a simultaneity temporal connective *-umye* plus a topic marker *-nun*, as shown in (i). Therefore, the evolvement of the conditionality is related with the concept of temporality.

(i) *-umyen* = *-umye* (simultaneity temporal connective) + *-nun* (topic marker)

iltan	*na-lul*	*wenmanghA-nAn*	*mAAm-un*	*twu*	*yekcek-i*
for.a.moment	I-ACC	fear.a.grudge.against-ADN	mind-TOP	two	traitor-NOM

hAn	*simcyang-i-n*
one	heart-be-ADN

'Although (the two traitors,) Hwukyem and Ninhan, are themselves in discord, they are of one mind in bearing a grudge against me for now,…'

<div align="right">(1777, Meynguylok #282)</div>

In the textual context of example (6), the adversative connective *-na* 'but' shows the contrast between the dissonance of the two traitors, Hwukyem and Ninhan and their shared position in 'resenting me.' The adverb *iltan*, combined with *-na* 'but,' also highlights a temporal contrast. Their dissonance on matters 'always' happened while their consonance in resenting me occurred 'for now.'

Fifth, in example (7) *iltan* means 'first of all,' indicating prioritization as an example of a bridging context (Heine 2002) between an adverb and a DM.

(7) 친하며 아니믈 의논치 말고 다 이 **일단** 의심하고 두려오믈 품어시니

chinhA-mye	*ani-m-ul*	*uynon-chi*	*mal-ko*	*ta*	*i*
close-CONN	not-NOMZ-ACC	discuss-do.COMP	stop-CONN	all	this

iltan	*uysimhA-ko*	*twulyeom-ul*	*phwum-es-ini*
first.of.all	doubt-CONN	fear-ACC	bear-PST-CAU

'First of all, they suspect who (they talk to), whether they are close or not, and have fear, so…'

<div align="right">(1782, Yucwungoytaysosinseyunum #25)</div>

Iltan can modify the verb *uysimha* 'doubt' and *twulyeomul phwum* 'fear' as an adverb. It highlights the contrast in priority between *chinhAmey animul uynon-* 'figuring out if (s/he) is close or not' and *uysimhAko twulyeomul phwum-* 'doubting and fearing things.' *Iltan* also functions as a DM, representing speaker's evaluative stance by balancing 'figuring out if (s/he) is close or not' with 'doubting and fearing things.' *Iltan* emphasizes the speaker's priority of on the latter over the former, extending its scope over the discourse.

The last notable point in EMoK is that the form *iltanun* (*iltan* + *-un*), which includes the topic marker *-un*, was used. However, the frequency of *iltanun* in the Historical Corpus is quite low, with only 14 hits, compared to 469 hits for *iltan*. Therefore, its usage was not common. Example (8) is excerpted from *Coyachemcay*, a record of historical events from 1392 (the 1st year of King Thayco's reign) to 1710 (the 36th year of King Swukcong's reign). *Coyachemcay* is believed to have been written during the era of King Swukcong (1661–1720), so it may have been composed in the early 18th century.[8]

---

[8] The variant form *iltanun* was attested in several old Hankul novels such as *Heynssiyangwungssanglinki*, *Wanwelhoymayngyen*, *Yunhacengsammwunchwilok*, and others. The periods of their creations are not known. If proven, the beginning of its usage will be revised.

(8) 홍이(弘禹)랄 유협하다난 말이 잇사온즉 이 **일단은** 맛당이 사실할 배로소이다

hongi-lAl	*yuhyephA-ta-nAn*	*mal-i*	*is-sAo-ncuk*
[name]-ACC	threat-REPT-ADN	word-NOM	exist-HON-CONN

i	***iltan-un***	*mastangi*	*sAsilhA-l*	*pA-i-losoita*
this	first.of.all-TOP	justifiably	investigate-ADN	NMLZ-be-END

'The word that Hongi is being threatening is circulating. First of all, we must conduct a justified investigation into it.'

(17??, *Coyachemcay* #12009)

In example (8), *iltanun* indicates a 'primary' evaluation of the speaker's comments regarding the investigation of the rumor that 'Hongi is being threatening,' suggesting that conducting an investigation into the matter is justified. Therefore, it serves as a DM, representing the speaker's evaluative stance.

In the 18th century during the EMoK period, the term *iltan* began to transition from a lexical term in LMK to its use in textual contexts. It was often seen with connectives such as *iltan -hamyen* meaning 'if…one morning' in example (4), *-na iltan* meaning 'but…for a moment' in example (6), and *-ko iltan* which translates to 'and…first of all' in example (7). The written form *iltan* in Hankul started to become more widely used during this time. Additionally, a variant form, *iltanun*, also began to be disseminated in the EMoK period. The DM functions that developed from *iltan* stemmed from its connotation of 'priority,' as demonstrated in examples (7) and (8).

4.3. *Iltan* in Modern Korean

The semantic designation 'one morning' or 'in the morning one day' of *iltan* was obsolete in Modern Korean (MoK). However, the extended meanings of *iltan* from the 18th century are confirmed in use in MoK: (i) priority (ii) brief temporality (iii) conditionality.

Let us look at the examples, which show the three meanings around the turn of the 20th century from the historical corpus. Example (9) is taken from *Moktanhwa*, which is classified as a *Sinsosel*.[9]

(9) 형셰를 보는 것이 안이요 가품도 보는 것이 안이요 **일단** 신랑 한아만 보고

hyengsyey-lul	*po-nun*	*kes-i*	*an-i-yo*	*kaphwum-to*	*po-nun*
condition-ACC	see-ADN	thing-NOM	not-be-POL	family.tradition-also	see-ADN

kes-i	*an-i-yo*	***iltan***	*sinlang*	*hAna-man*	*po-ko*
thing-NOM	not-be-POL	first.of.all	groom	one-only	see-CONN

'(The official Lee is looking for a marriage partner for his daughter Cengswuk.) He does not consider the groom's circumstances or family traditions. First of all, he considers a groom alone…'

(1911, *Moktanhwa* #204)

---

[9] A *Sinsosel* is a type of novel that was widely distributed around the turn of the 20th century in Korea.

In example (9), the speaker discusses the government official Lee's priority in selecting a marriage partner for his daughter Cengswuk. He emphasizes the groom candidate himself (*sinlang hAnaman po-* 'considering a groom alone') over factors such as the candidate's economic and social status or the family's reputation and traditions (*hyengsyey, kaphwumto ponun kesi aniyo* 'not considering the groom's circumstances or family traditions'). In this context, *iltan* functions as a DM, reflecting the speaker's evaluative stance on the official Lee's criteria for choosing a groom for Cengswuk.

Example (10) is taken from *Twukyenseng*, which is also classified as a *Sinsosel*. *Iltan* in expert (10) means 'for a while.'

(10) 혜경의 요됴한 셩품과 개결한 긔질이 꼭 내 마암에 합당하다 하야 **일단** 졍신이 곳 그 규슈에 잇스나

hyeykyeng-uy yotyohAn syengphwum-kwa kAykyelhAn kuycil-i
name-GEN   elegant   personality-and   upstanding   temperament-NOM
*skok   nAy   mAAm-ey   haptanghAta   hA-ya   **iltan**   cyengsin-i*
surely   my   heart-LOC   fit   do-CAU   for.a.while   mind-NOM
kos   ku   kyusyu-ey   issu-na
right   the   lady-LOC   exist-but
'(Mr. Wang) has had Hyeykyeng in mind for a while because he likes her elegant personality and upstanding temperament, which are congenial with him, but...'

(1912, *Twukyenseng* #218)

In example (10), Mr. Wang has been considering Hyeykyeng 'for a while' because he feels her personality and temperament align well with his own. However, he finds her situation too complicated to accept. The use of *iltan* combined with the adversative conjunction *-na* in (10) highlights his ongoing interest in her 'for a moment' against his awareness of her 'enduring' circumstances.

Example (11),[10] taken from *Tayhanmayilsinpo*, shows *iltan…meyn* meaning 'once' or 'in case'

(11) **일단**국혼이 완젼불멸하**면** 돌연히 일죠에 긔맥이 샹통하야 부강국이 되리로다
iltan.*kwukhon-i*   *wancyenpwulmyel-hA***myen**   *tolyenhi   ilcyo-ey*
once.national.spilit-NOM   total.immortal-do.COND   suddenly   one.day-LOC
kuymAyk-i   syangthonghA-ya   pwukangkwuk-i
the.wavelength(thread)-NOM   communicate-CONN   rich.strong.nation-NOM
toy-l-ilota
become-FUT-END
'In case the national spirit is totally immortal, the thread will suddenly run through one day, and the nation will become rich and strong.'

(1904, *Tayhanmayilsinpo* #18714)

---

[10] In the orthography of PDK, a white space divides *iltankwukhon* into *iltan* and *kwukhon* and *kuymAyk* into *kuy* and *mAyk*. But in attestation of *Tayhanmayilsinpo* they are not broken down by the white space.

The structure in (11) suggests that if the national spirit is vibrant, the nation can achieve wealth and strength when the spirit unexpectedly unites one day. In this instance, *ilcyo,* not *iltan,* signifies 'one morning' or 'one day.' When *iltan* is combined with the conditional connective *-myen*, it does not imply temporality. Instead, in this context, *iltan* with *-myen* indicates the conditional 'once' or 'in case.' It is significant that the combination of *iltan* and *-myen* firmly establishes itself as a conditional form in MoK.

Phyocwun Kwuke Taysacen (A Complete Standard Korean dictionary) by the National Institute of the Korean Language also presents the three meanings in PDK: (i) 'first of all' (ii) 'for a while,' 'for now' (iii) 'once,' 'in case.' The examples in (12), taken from *Phyocwun Kwuke Taysacen*, illustrate three meanings of *iltan* respectively.

(12) a. ***iltan***   *ku kos-ey*   *ka po-ca*
       first.of.all   the place-LOC   go see-HORT
       'First of all, let's go there.'
   b. ***iltan***   *cali-lul*   *phiha-sio*
       for.a.while   place-ACC   avoid-HON.IMP
       'Please, leave the place for a while.'
   c. *aphu-myen* ***iltan***   *peyngwen-pwuthe*   *ka-la*
       sick-if   in.case   hospital-from   go-IMP
       'In case you are sick, please see a doctor.'

In example (12a) *iltan* signifies 'first of all,' that is, priority is put on 'going there.' In example (12b), *iltan* means 'for a while,' indicating that leaving the place is going to continue 'for a short period.' In example (12c), the combination of the conditional connective *-myen* and *iltan* designates 'once' or 'in case,' that is, *aphumyen iltan* 'in case you are sick' assumes a hypothetical situation. The extended meanings of realis, irrealis, and necessary/sufficient conditions are derived from the conditional *-myen* 'if,' which are not illustrated in detail in this article.

4.4. *Iltan* as a Discourse Marker

Iltan developed into diverse discursive functions as a DM in the late of the 18th century within EMoK. The DM *iltan* diversifies its meanings and functions in discourse in PDK. The DM *iltan* exhibits the characteristics typically associated with DMs: they are semantically and syntactically independent from their surrounding context, are prosodically distinct from the rest of the utterance, and serve a metatextual function that is anchored in the discourse situation, facilitating text organization, expressing the speaker's attitudes, and/or enhancing speaker-hearer interaction, following the definitions of DMs by Brinton (1995), Schourup (1999), Andersen (2001), González (2004), and Furkó (2014), among others (cited in Heine et al. 2021: 6).

The multifarious discursive meanings and functions of *iltan* are both subjective and intersubjective. In the examples, lines that do not require morphemic glosses

are given in English translation within square brackets, and supplementary information not present in the original text but helpful for understanding the meaning is included in parentheses in the translation.

4.4.1. Topic presenter

First, let us examine the example of *iltan* functioning as a topic presenter at MED (an utterance-medial position). In excerpt (13), C has been diagnosed with a terminal-stage of pancreatic cancer. C and his wife, J, are sitting in front of the desk of D, a chief doctor of the medical center. Mr. Pwu, a friend of C and J, is with them. D is examining his sonogram, clicking on the computer. J is hoping that D will offer some hope for cancer treatment. Finally, D turns toward C and J, moving away from his computer.

(13)　D: [(Investigating a medical chart without a word)]
　　　J: [(Waiting for the doctor to say something)]
　　　D: (Not casting a glance at her) 부선생 말대로, **일단** 항암치료를 하시는 게,
　　　　*Pwu sensayng　mal-taylo **iltan** hangam chilyo-lul　　ha-si-nun　　ke-y*
　　　　[name] teacher word-as　DM　anticancer treatment-ACC do-HON-ADN　thing-NOM
　　　　'As Mr. Pwu said, let's see, why don't we try anticancer treatment…'
　　　J: [Can he recover completely with anticancer treatment alone? Just, surgery is quicker and cleaner, isn't it? (It is likely that she is out of her mind.)]
　　　　　(2005, Drama *90il Salanghal Sikan* [90 days, the time to love] Episode #04)

In (13) D attempts to speak first after a period of silence. J is waiting for D to say something hopeful about her husband C's medical therapy. As a DM, *iltan* at MED serves as a topic presenter, breaking the silence and paving the way to address the specific topic of 'C's anticancer treatment.' Therefore, *iltan* in (13) is a microscope topic presenter rather than a macro-scope or medial-scope topic presenter because it introduces a very local level discourse topic, such as *kukey mwe-nyamyen* (that-thing-NOM what-HYP.COND, 'speaking of what') and *kukey nwukwu-nyamyen* (that-thing-NOM who-HYP.COND, 'speaking of who') (Rhee 2014: 5).

4.4.2. Subtopic shifter

The next example (14) illustrates a subtopic shift expressed by *iltan* at MED. M, a student majoring in electrical engineering at KAIST (Korea Advanced Institute of Science and Technology), visits Y's factory to inspect the spindle motor. The situation involves testing it for a company that is considering investing in it.

(14)　M: [(Smiling)… If it goes well, will production start in earnest?]
　　　Y: 그야 좋은 결과가 나올 때 얘기지. 그래서 **일단** 모터라도 손봐야할거 같아서. 그래서 학생을 불렀어. 도와줄 수 있겠지?
　　　　ku-ya　　cohun kyelkwa-ka nao-1-ttay　　　　yayki-ci　kulayse
　　　　it-COND　good　result-nom　come.out-ADN-when　talk-SFP　therefore
　　　　***iltan** mothe-lato　sonpo-aya　　　ha-1-ke　　　katha-se.*
　　　　DM　　motor-CONC　take.care.of-NF　do-ADN-thing　be.like-CONN

kulayse   haksayng-ul   pwull-ess-e
therefore   student-ACC   call-PST-SFP
'That's only when we get good results. So, first of all, I thought we needed to check the motor in the first place. That's why I called you. Can you help out?'

(2000, Drama *Khaist* [KAIST] Episode #52)

In excerpt (14), M discusses the prospects of the ongoing project, and Y also anticipates a good result. In preparation for this, he plans to check the motor. The main topic between M and Y is the spindle motor project and checking the motor is a related subtopic. *Kulayse iltan* 'so, first of all' in Y's talk signals a shift to the subtopic.

4.4.3. Interaction-initiator

Iltan in excerpt (15) also plays a role to break silence at the beginning of the discourse. C is a CEO running a business. M is a businessman and the boss of a gang in Thailand. C regards M as his business partner. They are sitting across the table from each other in a back room at a secret restaurant in a traditional Korean-style house. Behind them are their body guards.

(15) C: 일단 말이 통하니 편합니다. 그동안 통역 끼고 일하려니까 원 답답해서...
      iltan *mal-i   thongha-ni   phyenha-pnita   kutongan*
      DM   word-NOM communicate-CAU comfortable-HON.DEC meantime
      thongyek   kkiko   ilha-lye-nikka   wen taptapha-y-se
      interpreter together work-PURP-CAU INTJ frustrate-be-CAU
      'Above all else! It's convenient that we can communicate to each other (in Korean) now. It was frustrating having to rely on interpreters all this time...'
    M: [I feel refreshed for everything, trying to realize my dream in my homeland. Oh, my God! I owed you a great debt before.]
      (2007, Drama *Kaywa Nuktayuy Sikan* [Time for a Dog and a Wolf] Episode #03)

In (15) C starts his talk with the DM *iltan*, breaking the awkward atmosphere in the room. Therefore, the DM *iltan* functions as an interaction-initiator. In this context, *iltan* leads to a phatic communion (Malinowski 1972[1923]), prompting an exchange of turn-taking over conventional small talk to facilitate smooth interaction and enhance social rapport before engaging in serious business conversation. C discusses the advantages of conducting business with M, saying *iltan mali thonghani phyenhapnita* 'You know, it's convenient that we can communicate (in Korean) to each other now.' In response, M expresses his renewed feelings about doing business in his homeland, Korea, and adds his gratitude for C's previous treatment, as if something has suddenly come to mind. Phatic communion consists of conventionalized and formulaic expressions exchanged between interlocutors, showing an interpersonal and attitudinal stance.

4.4.4. Pause-filler

The pause-filler function and the hesitation function are practically indistinguishable because the distinction resides only in the speaker's cognitive state or intention (Rhee 2020a and Rhee 2021). Excerpt (16) shows an example of *iltan* functioning as a pause-filler. H (a male)'s wife has gone to see her mother in a local city, and Y (a female)'s husband has an appointment to see his friends. H is asking Y to have dinner together.

(16) Y: [(Feeling stuffy and getting up)]

    H: [(Startled) Why, why? Are you leaving?]

    Y: [No... I just want to make a phone call at home... Because of my husband, I'm afraid he might be waiting for me. It's been a week since he finished work and came back home... (trying to read his countenance) Well, he mentioned that he's meeting a friend today though...]

    H: [Oh... okay. If he's going to see a friend, let's have dinner together.]

    Y: [?]

    H: [(Mumbling) Um, my wife... Her parents moved their place... Where was it?... right, Cella Province, Swunchang... the famous place for red-pepper paste... She has gone there.]

    Y: 그, 글쎄....**일단** 남편한테...물어보구. 남편이 안된다 그러면 안되구.

    *k, kulssey... **iltan** namphyen-hanthey... mwul-epo-kwu.*

    It, DM.   DM   husband-DAT   ask-ATTP-END.

    namphyen-i   an-toy-n-ta   kule-myen   an-toy-kwu.

    husband-NOM   not-become-PRES-COMP   say.so-if   not-become-END

    'Well... first, I'll ask my husband... If he minds, then it's not okay.'

                                        (Drama, *Kecismal* [A lie] Episode #02)

A pause-filler can provide time for a proper lexical search. In this excerpt, the pause-fillers *kulssey* and *iltan* are used to find suitable excuses to politely reject without being offensive. The DM *kulssey* (Rhee 2015) indicates Y's hesitation to give a clear answer to H's proposal of having dinner together. *Iltan* also functions as a DM, signaling her priority on her husband's schedule (*namphyenhanthey mwulepokwu* 'I'll ask my husband') rather than H's proposal, thereby avoiding an immediate and direct rejection. In this sense, *iltan* acts as a hesitation marker. She expresses her hesitation with the combined effects of the two DMs *kulssey* 'well' and *iltan* 'first of all.' The two hesitation markers are employed to avoid disheartening H in a straightforward way. Therefore, hesitation markers demonstrate politeness to save the addressee's face, an example of intersubjectification.

4.4.5. Echoing response marker

Echoing response repeats what the interlocutor has just said, thereby demonstrating the interest in the ongoing conversation and maintaining the floor. Example (17) is a telephone conversation between K (a female) and Y (K's female friend) about K's boyfriend, J. J is holding a nice birthday party for K, just between

J and K. During her birthday party, K is asking Y for advice over the phone to develop a more intimate relationship with him.

(17) Y: [Well, yeah, that's right, that... huhuhu, Mr. J, he's got some subtle charm.]
K: [(Eyes widen) Is that so?]
Y: [What are you up to?]
K: [(Looking at underwear) I'm contemplating. What should I do?]
Y: 어떡하긴 뭘 어떡해! 일단! 새로 산 속옷있지, 무조건 그거 입어.

   ettek-ha-ki-n   *mwe-l*   *ettek-ha-y!*   *iltan!*   *saylo sa-n*
   how-do-NOMZ-TOP what-ACC how-do-be   DM   newly buy-ADN
   sokos-iss-ci   *mwucoken*   *kuke*   *ip-e*
   underwear-exist-Q without.question it   put.on-IMP
  'What should you do? First of all! You've just bought new underwear, so wear that for sure.'
K: **일단**...입고 나서?...

   iltan*… ip-ko*   *na-se?...*
    DM   wear-CONN come.out-CONN
   'Alright…after I put it on, and then?...'
Y: [The woman who approaches indiscriminately is really not that great.⋯]

                 (2010, Drama *Kayinuy Chwihyang* [A taste of Kayin] Episode #13)

In excerpt (17) Y advises K to wear new underwear. After listening to her advice, K repeats the phrase *iltan ip* 'alright, wear it...' and asks a question, *ko nase?* 'and then?.' In Y's remark *iltan! saylo san sokos…ipe* 'first of all, put on the new underwear that you bought,' *iltan* means 'first of all,' modifying the following utterance *saylo san sokos kuke ipe* 'wear the new underwear you have bought.' In K's remark *iltan* means 'alright' or 'OK.' *Iltan* shows K's response to Y's advice. By using the DM *iltan*, K expresses her involvement and interest in Y's advice and allows Y to continue speaking, functioning as a continuer. Y answers K's question about the next stage. As an echoing response marker, the second DM *iltan* can signal to the speaker that the addressee is actively engaged in the conversation, effectively exchanging turns. Echoing response in this context supports the idea that conversation is a collaborative effort between the speaker and the addressee, illustrating the concept of intersubjectivity.

4.4.6. Common-ground solicitor

Excerpt (18) is another example illustrating intersubjectivity. The drama *Censeluy Kohayng* deals with a story inherited from ancient times. In this episode, there are characters J, K, and A. In the drama, they are conspiring to rob a sedan chair belonging to a high official L, who is currently on a long business journey.

(18) J: [Okay. First of all, let's have A take charge of the high official Lee. Find out where he is, and bring him over here by any means necessary. Will three days be enough? (Looking at K)]
K: [(Excitedly nodding) Three days sounds good.]
A: [Well, that should be sufficient.]

J: 그럼 가마 터는 법을 생각해 보자구. 가마 째로 뒤바꾸는 게 제일 쉬운
법일텐데.. 좋아. 그 방법으로 가자구. **일단**,
kulem kama the-nun pep-ul sayngkak-ha-y (e)po-cakwu.
if.so sedan.chair rob-ADN law-ACC think-do-NF ATTP-HORT.END
kama ccaylo twipakkwu-nun ke-y ceil swiwun
sedan.chair in.whole exchange-ADN thing-nom first easy
pep-i-ltheyntey.., coha. ku pangpep-ulo ka-cakwu. **iltan**,
law-be-CONN OK. it method-INST go- HORT.END DM.
'If so, let's think about robbing him of the sedan chair. It might be the easiest way to
exchange it entirely for another. OK let's go with it. Alright,'
(2008, Drama *Censeluy Kohayng* [A Birthplace of Folk Tales] Episode #7)

In excerpt (18), J, K, and A divide their tasks. A is in charge of finding L and
bringing him here. J proposes the idea of robbing L of his sedan chair, saying, *kama
ccaylo twipakkwunun key ceil swiwun pepiltheyntey* 'the easiest way is to exchange
the sedan chair in whole with another.' In this excerpt, J uses *iltan* at the utterance-
final position to solicit common-ground or consensus on his idea from K and
A. Therefore, it is interpersonal and intersubjective.

4.4.7. Interruption marker

Interrupting the speaker is generally regarded as impolite. However, for various
reasons, discourse markers of interruption (DMIs) are often used by discourse
participants to cut into the discourse and discourage the speaker from continuing.
The DMI is used tactfully and effectively in discourse to achieve the interlocuter's
purpose while minimizing the face-threatening nature of the speech act for the
speaker. Therefore, DMIs often take the form of short fragmentary expressions to
gain immediate attention, such as *guarda* 'look' in Italian (Waltereit 2002), *wait*
(Balteiro 2018, Tagaliamonte 2021), *camkkan*, *kaman*, and *kuman* 'just a moment'
in Korean (Rhee 2024). *Iltan* also functions as a DMI in Korean.

In excerpt (19) there are two characters: a female, T, and a male, S. S wants to
propose marriage to T, while T also has something important to tell him.

(19) T: [Mr. S. I have something to talk to you.]
    S: 잠깐만요. **일단**, 차에 타고 얘기해요. 저도 Miss T 한테 할 말 있어요.
    camkkanmanyo. **iltan**, *cha-ey tha-ko yayki hay-yo. ce-to*
    DMI DMI, car-LOC get.on-CONN talk do-HON I-also
    Miss T-hanthey ha-l mal iss-eyo.
    [name]-DAT say-ADN word exist-POL
    'Just a moment, wait, let's get in the car and then have a talk. I also have something
    to discuss with you.'
    (2003, Drama *1%uy Etten Kes* [A Thing of 1%] Episode #10)

In excerpt (19) S interrupts T's speech using a combination of the DMIs
camkkanmanyo 'just a moment' and *iltan* 'wait' or 'look.' *Camkkanmanyo* is one
of the DMIs (Rhee 2024). The DMIs *camkkanmanyo* and *iltan* suggest that 'getting
in a car' should take priority over her talk, as getting into the car is the most urgent

task at hand. These DMIs *camkkanmanyo* and *iltan* shift the focus to the immediate task, showing S's positive and receptive attitudinal stance toward T's agenda, which will be addressed after they get into the car. In this context, although the DMIs *camkkanmanyo* and *iltan* interrupt T's speech, S's receptive attitude towards T's agenda indicates that *iltan* is not used impolitely here.

5. Discussion

5.1. The source characteristics

The Chinese *yīdàn* (一旦) and the borrowed Korean word *iltan* exhibit both similarities and differences in their meanings and functions. Let us examine the similarities. First, the literal and original meaning of *yīdàn* as 'one morning' is obsolete in Present-Day Chinese (PDC), just as *iltan* is in PDK, according to Long (p.c.) and Yang (p.c.).

This paper argues that the meaning of 'one morning' in *iltan* in Korean competed with, and gave way to the Sino-Korean word *ilco* (一朝: *il* (一) meaning 'one' and *co* (朝) meaning 'morning'), which also means 'one morning,' as well as the pure Korean words *halwu-achim* (a.day-morning) meaning 'one morning' and *enu-nal* (some-day) meaning 'one day' (layering (Hopper & Traugott 2003[1993])). The Chinese character *ILCO* is first attested in the 15th century in Korean Historical Corpus, as seen in (20a), and the written form *ilco* in Hankul is first attested in the 18th century, as seen in (20b). Both the Chinese character word *ILCO* and Hankul *ilco* are still in use in PDK.

(20) a. 百千佛刹알 **一朝**애 革하시니

   PAYCHEN-PWULCHAL-Al　　　　**ILCO**-*ay*　　　 *HYEKhA-si-ni*
   hundred.thousand-Bhuddhist.temple-ACC　 one.morning-LOC　 innovate-HON-CAU
   '(The king) innovates many Buddhist temples one morning.'

   (1447 *Yongpiechenka* #213)

   b. 엇지 **일조**의 사람을 죽이고져 하나냐

   esci　　 ***ilco**-uy*　　　　 *sAlAm-ul*　　 *cwuki-kocye hA-nA-nya*
   how.come one.morning-LOC human-ACC　 kill-PURP　　 do-CR-Q
   'How come you intend to kill the people (you came here with) one morning?'

   (17?? *Ulpeyngyenhaynglok* 4, #551)

Example (11) from Section 2.3, which is revisited in (21), features both *iltan* and *ilcyo* (a variant of *ilco*, which was also first attested in the 18th century and circulated until the turn of the 20th century.) within the same sentence. Despite originating from logographs with the same meaning, 'one morning,' *iltan* conveys a conditional sense, while *ilcyo* means 'one morning' or 'one day.'

(21) **일단**국혼이 완전불멸하면 돌연히 **일죠**에 긔맥이 샹통하야 부강국이 되리로다

   ***iltan**kwukhon-i*　　　　　 *wancyenpwulmyel-hAmeyn tolyenhi **ilcyo**-ey*
   once.national.spilit-NOM　　 total immortal-COND　　　　 suddenly one.day-LOC

> *kuymAyk-i*   *syangthonghA-ya*   *pwukangkwuk-i*
> the.wavelength(thread)-NOM   communicate-CONN   rich.strong.nation-NOM
> *toy-l-ilota*
> become-FUT-END
> 'In case the national spirit is totally immortal, the thread will suddenly run through one day, and the nation will become rich and strong.'

<div align="right">(1904, Tayhanmayilsinpo #18714)</div>

As mentioned in Section 4.1, *iltan* has evolved from an adverb into a conditional connective and DM, losing its original meaning of 'one morning.' Conversely, *ilco* has not undergone   grammaticalization to become a DM and still retains its original meaning of 'one morning' in PDK.

The pure Korean word *halwuachim* is first attested in the 15th century, as seen in (22) and is still in use in PDK.

(22) a. 無量劫으로셔 長常母子이 다외야 오다니 하랏 **아차매** 나외야 몯보잡게 다외요니
>   *MWULYANGKEP-ulo CANGSANG-MOCA-i*   *tAoy-ya*   *o-ta-ni*
>   long.long.time-INS   always-mother.son-NOM   become-NF   come-RETR-CONN
>   ***hAlAs achAm-Ay***   *nAoyya mot-pozAp-key tAoy-yoni*
>   one.day morning-LOC   again   not-see-CONN   become-CAU
>   'For a long long time we have always been a mother and son. One morning, a mother and her son can't see each other again.'

<div align="right">(1447 Sekposangcel 23 #28)</div>

As seen in examples above, the earliest attestations of the Sino-Korean words *iltan* and *ilco*, and the pure Korean words *halwuachim* and *enunal* were all found in the 15th century.

Second, *yīdàn* in PDC also designates 'for a while' or 'in a moment,' according to Long (p.c.). The text in (23) is taken from the Biography of Song Fu in the History of the Late Han Dynasty, which was written by the historian Fan Ye in the 5th century.

(23) 天下非一时之用也,海内非**一且**之功也
>   *tiānxià*   *fēi*   *yīshí*   *zhī yòng yě,*
>   the.world   not   single.moment   of   use   END
>   *hǎinèi*   *fēi*   ***yīdàn***   *zhī gōng*   *yě.*
>   within.the.seas   not   in.an.instant   of   achievement   END
>   'The world cannot be achieved in short time, nor can the stability of the realm be accomplished in short time.' (5th century, Book of Late Han Dynasty, Biography of Song Fu (后汉书·宋浮传))

In (23), the Chinese *yīdàn* means 'in an instant' or 'for a while.' (24) provides another example of *yīdàn* in Chinese, meaning 'in an instant' or 'for a while.'

(24)   毀于**一且**
>   *huǐ*   *yú*   ***yīdàn***
>   destroy   in   short.period.of.time

'(Something) is destroyed in a very short time'

Third, the conditional connective meaning of 'once' or 'in case' is present in both PDC and PDK, according to Zhang (2014: 218–219), Long (p.c.), and Yang (p.c.). Example (25), taken from a Chinese dictionary, illustrates the hypothetical meaning of *yīdàn*.[11]

(25) 一旦发生核战争, 整个地球将会陷入灾难.
　　 ***Yīdàn** fāshēng hé　　 zhànzhēng, zhěnggè dìqiú jiāng huì　 xiànrù　 zāinàn.*
　　 if　　 happens nuclear war　　 entire　 earth　 will　 might　 fall.into disaster
　　 'If an atomic war were to break out, it would be disastrous for the entire world.'

Fourth, regarding grammatical category, both *yīdàn* in PDC and *iltan* in PDK are used generally as adverbs. Ráo & Pān (2011: 48) argue that the adverb *yīdàn* began to emerge in the Eastern Han Dynasty (from 25 AD to 220 AD) and further developed during the Wei, Jin, and Northern and Southern Dynasties (from 220 AD to 589 AD).

The difference is that the Chinese *yīdàn* does not carry the meaning of priority. Long (p.c.) argues that *yīdàn* has evolved in meaning along the path of ['one morning' > 'in a short time' > 'in case']. For the meaning of priority, *lìkè* (立刻) is used in PDC. The term *lìkè* has developed in meaning along the path of ['in a short time' > 'immediately' > 'first of all'].

Although *iltan* in Korean is a borrowed term, its evolution indicates a range of diverse meanings such as 'first of all,' 'for a while,' 'once,' and 'in case,' along with its functions as a DM, which are more varied than those of *yīdàn* in Chinese. *Iltan* in Korean exemplifies 'evolutive' or 'natural' development (Paul 1920[1880], Anderson 1973, Thomason & Kaufman 1988). These scenarios strongly suggest that the cognitive processes driving this development are effective, and that the lexical source significantly influences the direction and outcome of grammaticalization, as proposed by the Source Determination Hypothesis (Bybee et al. 1994, Traugott 2014b).

5.2. Discourse grammar and cooptation

The findings in Section 4 showed that the development of the Korean DM *iltan* is ascribed to the interplay between cooptation and grammaticalization in discourse grammar (Heine et al. 2017, 2021), discussed in Section 2. This process began with early grammaticalization, which transformed the noun *iltan*, meaning 'one morning,' into an adverb and a conditional connective in LMA. During the early grammaticalization phase, *iltan* undergoes decategorization and desemanticization with a zero morpho-phonological change. This was followed by cooptation, where the coopted unit (the thetical *iltan*) was utilized as a DM in the 18th century within the EMoK. Finally, late grammaticalization occurs within the DM, allowing *iltan*

---

[11] https://www.collinsdictionary.com/zh/dictionary/chinese-english/%E4%B8%80%E6%97%A6

to take on a broad spectrum of subjective and intersubjective stances expressed by the speaker. At the late grammaticalization phase, the DM *iltan* is used in a broader range of situations, including presenting topics (=DM1) as exemplified in (13), shifting subtopics (=DM2) as shown in (14), initiating interaction (=DM3) as in (15), filling pauses or indicating hesitation (=DM4) as in (16), expressing echoing response (=DM5) as in (17), seeking common ground (=DM6) as shown in (18), and causing interruptions (=DM7) as illustrated in (19). The evolution of the DM *iltan* can be illustrated as shown in (26).
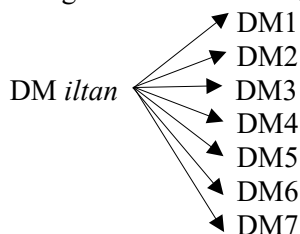
(26) Development of DM *iltan*
    Phase 1: Early grammaticalization (decategorization, desemanticization, zero derivation)

            N. *iltan*   ⟶   Adv. *iltan*
                               Conn. *iltan*
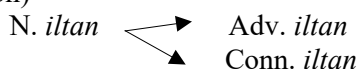
    Phase 2: Cooptation (involving the rise of DM)

           Adv. *iltan* → DM *iltan*

    Phase 3: Late grammaticalization (context extension of pragmatic uses)

                       DM1
                       DM2
        DM *iltan*   DM3
                       DM4
                       DM5
                       DM6
                       DM7

      The transition of *iltan* from an adverb in sentence grammar to a DM at the discourse level involves several changes outlined in (27). These changes together allow *iltan* to broaden its functions, moving from lexical uses at the sentence level to pragmatic applications at the discourse level. Consequently, cooptation signifies a process that highlights this shift, as *iltan* becomes more clearly associated with the discourse context.

(27) Effects of cooptation for the Korean DM *iltan* (adapted from Heine et al. 2021: 208)
    a. Meaning: As a DM, *iltan* no longer contributes to the propositional content of its host utterance.
    b. Function: Its function is metatextual, relating to the attitudes of the speaker and speaker-hearer interaction.
    c. Syntax: It has a syntactic independence from the rest of the sentence.
    d. Prosody: It is marked off prosodically. There may be pauses before and after *iltan*.
    e. Semantic-pragmatic scope: It seems to have distinctly wider scope than the adverb, extending beyond the clause it is associated with.
    f. Placement: It can either be placed at the left or the right periphery of a sentence or in medial position.

5.3. (Inter)subjectification and periphery

'One morning,' the original meaning of *iltan*, was circulated until in the 18th century and became extinct in PDK. But the diverse meanings and functions derived from its source can be illustrated by metonymy, subjectification, and intersubjectification.

Morning comes first, ahead of afternoon and evening. Morning swiftly passes and gives way to afternoon and evening. Therefore, the semantic extension of *iltan* from 'one morning' to priority ('first of all') and brevity ('for a while') is an instance of metonymy based on contiguity.

It is a natural course of the developmental path, driven by the human tendency to attribute subjective value to an event or state of affairs. The emergences of the meanings such as 'first of all,' 'for a while,' and 'once' from the original meaning 'one morning' serve as a clear example of subjectification (Hopper & Traugott 2003[1993]). This process involves evaluating an aspect or entity from a perspective of another aspect or entity.

The development of a DM *iltan* exemplifies both subjectification and intersubjectification. The functions of a subtopic shift, a pause-filling, and a DMI are the instances of subjectification. In general, the DM of interaction-initiator, leading to the phatic communion by exchanging conventional expressions, tend to go further from the proposition level to the interactional level for strategic reasons, i.e., to engage the interlocutor to pay attention to the discourse and share the smooth atmosphere that the speaker designs before they start a serious official topic, thus inducing the function from the subjective to the intersubjective level. Furthermore, the development of the DM *iltan* —as a marker of inviting the addressee into the common ground of shared information or feelings, of signaling hesitance to reveal dispreferred information for the sake of politeness, and of showing an echoing response to encourage the speaker to continue —serves as a good example of intersubjectification (Traugott 2003, 2010, Traugott & Dasher 2002). The developmental path of the semantic-functional extension of the lexeme *iltan* from the source meaning to the DM can be diagrammatically presented in Figure 1.

Regarding the peripheral asymmetry hypothesis (Beeching & Detges 2014) discussed in Section 2, the DM *iltan* appears at LR, RP, and MED, demonstrating both subjective and intersubjective functions. Observations indicate that *iltan* occurs more frequently at LP than at MED or RP, as shown in Section 4.4. When distinguishing subjective and intersubjective functions of *iltan*, its roles as an interruption signal, sub-topic initiator, and pause-filler are subjective, whereas its other functions—such as echoing response, common-ground solicitation, hesitance, and interaction initiation—are intersubjective. The function of a topic-presentation can be both subjective and intersubjective, depending on the context. The (inter)subjectivity variable and the positionality variable with respect to the functions of the DM *iltan* are laid out in Table 1. *Iltan* at LP carries both subjective and intersubjective functions. Thus, the peripheral asymmetry hypothesis, which states that LP is associated with the subjective functions whereas RP is associated

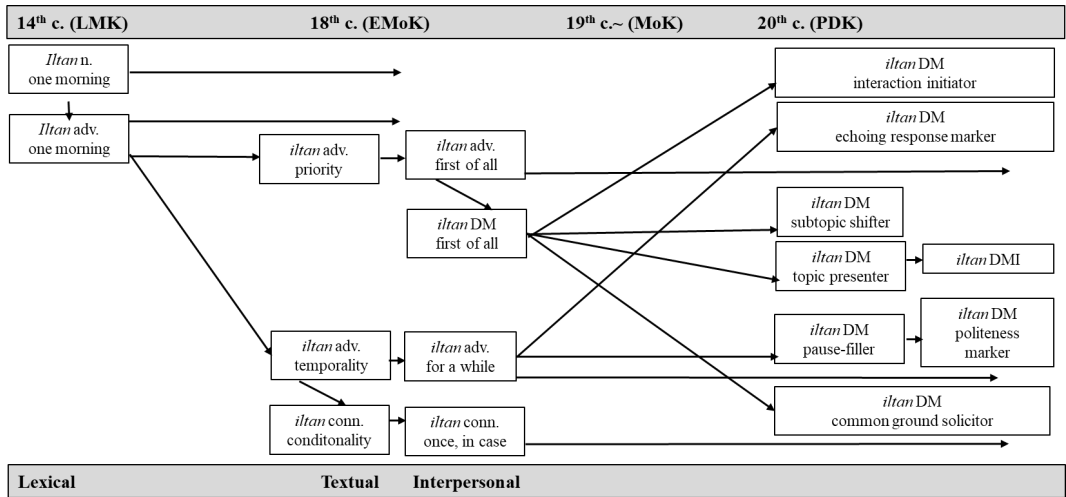with the intersubjective functions, is not supported by the developmental pattern displayed by the DM *iltan*.

Figure 1. Semantic and functional extension patterns of *iltan*

Table 1. **(Inter)subjectivity and periphery of the DM *iltan***

	Subjective function	Intersubjective function
LP	DMI	Interaction-initiation
	Pause-filling	Echoing response
		Hesitance
MED	Sub-topic shift	
	Topic presentation	Topic-presentation
RP		Common-ground solicitation

6. Conclusion

This article has explored the development of the Sino-Korean lexeme *iltan*, focusing specifically on its semantic and structural changes. The word is composed of two logographs: *il* (一) meaning 'one' and *tan* (旦) meaning 'morning'. Its earliest recorded use as a noun dates back to the 15th century in LMK, where it appeared in a non-argument position (in the adverbial phrase) in combination with a locative postposition *-ay*. The adverb *iltan*, dropping the locative postposition *-ay* and undergoing zero derivation, began to be used in the 16th century in LMK. The initial meaning of 'one morning' persisted until the 18th century in EMoK. By the 18th century, new meanings derived from contextual usage began to emerge. Additionally, from this period, *iltan* started to be used with the topic marker *-un*, though less frequently than without it in historical texts. Its role as a DM through the cooptation in the discourse grammar developed in EMoK.

Throughout its semantic evolution, *iltan* demonstrates a shift from temporal sense to notions of priority and conditionality, highlighting the speaker's perspective (subjectification). It also shows intersubjectification, where the speaker

cues the listener to share a certain interactional attitude and evaluation. When considering DMs, the positions at LP, RP, and MED are often analyzed. Although *iltan* can be found at MED and RP, it is more frequently seen at LP, where it conveys both subjective and intersubjective implications. Therefore, the positional asymmetry does not align with the use of *iltan* as a DM. This paper has only briefly outlined the similarities and differences between the Korean *iltan* and the Chinese *yīdàn*. However, the analysis of their source characteristics suggests that this research calls for extensive comparative studies of *iltan* across Asian languages that use Chinese-origin borrowed words.

Acknowledgments

Abbreviations

ACC: accusative; ADN: adnominal; ALL: allative; ATTP: attemptive; CAU; causative; COMP: complementizer; CONC: concessive; COND: conditional; CONN: connective; CR: current-relevance; DAT: dative; DM: discourse marker; DMI: discourse marker of interruption; EMoK: early modern Korean; END: sentential-ending; FUT: future; GEN: genitive; HON: honorific; HORT: hortative; HYP: hypothetical, INST: instrumental; INTJ: interjection; LMK: late middle Korean; LOC: locative; MK: middle Korean; MOK: modern Korean; NEG: negative; NF: non-finite marker; NOM: nominative; NOMZ: nominalizer; OK: Old Korean; PDC: present-day Chinese; PDK: present-day Korean; POL: polite; PRES: present; PST: past; PURP: purposive; REPT: reportative; RETR: retrospective; SFP: sentence-final particle; SUB: subjunctive; and TOP: topic.

References

Andersen, Gisle. 2001. *Pragmatics Markers and Sociolinguistic Variation. A Relevance-theoretic Approach to the Language of Adolescents*. (Pragmatics & Beyond, New Series, 84.]. Amsterdam: Benjamins.

Andersen, Henning. 1973. Abductive and deductive change. *Language* 49 (4). 765–793. https://doi.org/10.2307/412063

Balteiro, Isabel. 2018. Oh wait: English pragmatic markers in Spanish football. *Journal of Pragmatics* 133. 123–133.

Beeching, Kate & Ulrich Detges. 2014. Introduction. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change,* 1–23. Leiden: Brill. https://doi.org/10.1163/9789004274822_002

Brinton, Laurel. 1995. Pragmatic markers in a diachronic perspective. In Jens Ahlers, Lynn Bilmes, Joseph S. Guenter, Betsy A. Kaiser & Jae Namkung (eds.), *Proceedings of the twenty-first annual meeting of the Berkeley linguistics society, February 17-20, 1995.*

General session and parasession on historical issues in Sociolinguistics/ Social Issues in Historical Linguistics, 377–88. Berkeley: Berkeley Linguistics Society.

Bybee, Joan. L., Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World.* Chicago: The University of Chicago Press.

Diewald, Gabriele. 2002. A model for relevant types of contexts in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization,* 103–120. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.09die

Furkó, Bálint Péter. 2014. Cooptation over grammaticalization. *Argumentum* 10. 289–300.

González, Montserrat. 2004. *Pragmatic Markers in Oral Narrative: The Case of English and Catalan*. Amsterdam, Philadelphia: Benjamins.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization,* 83–101. Amsterdam: Benjamins. https://doi.org/10.1075/tsl.49.08hei

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2017. Cooptation as a discourse strategy. *Linguistics* 55 (4). 813–856.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108982856

Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*. Volume 1, 17–35. Amsterdam, Philadelphia: Benjamins.

Hopper, Paul J. & Elizabeth C. Traugott. 2003[1993]. *Grammaticalization*. Cambridge: Cambridge University Press.

Koo, Hyun Jung & Seongha Rhee. 2001. Grammaticalization of a sentential end marker from a conditional marker. *Discourse and Cognition* 8 (1). 1–19.

Lee, Hyun Sook. 2021. From reality to contrast to counterexpectation: The development of *silsang* in Korean. *East Asian Pragmatics* 6 (3). 403–421.

Malinowski, Bronisław. 1972. Phatic communion. In John Laver & Sandy Hutcheson (eds.), *Communication in face to face interaction: Selected readings,* 146–152. Harmondsworth: Penguin. [Excerpt from Malinowski, Bronisław 1923. The problem of meaning in primitive languages. In Charles Kay Ogden & Ivor Armstrong Richards (eds.), *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism*. Supplementary Essays by Bronisław Malinowski & Frederick George Crookshank. 296–336. New York: Harcourt.]

Narrog, Heiko & Seongha Rhee. 2013. Grammaticalization of space in Korean and Japanese. In Martine Robbeets & Hubert Cuyckens (eds.), *Shared grammaticalization. With special focus on the Transeurasian languages,* 287–315. Amsterdam: Benjamins. https://doi.org/10.1075/slcs.132.21nar

Paul, Hermann. 1920[1880]. *Prinzipien der Sprachgeschichte* (5th ed.). Halle: Niemeyer.

Phyocwun Kwuke Taysacen. 1999. (*A complete standard Korean dictionary*). Seoul: The National Institute of Korean Language. https://stdict.korean.go.kr/main/main.do

Ráo, Chūn & Yùhuá Pān. 2011. Yīdàn de yǔfǎhuà kǎochá [A grammaticalization approach on 'yīdàn']. *Journal of Shayang Teachers College* 12 (3). 46–49.

Rhee, Seongha. 1996. *Semantics of verbs and grammaticalization: The development of Korean from a cross-linguistic perspective*. Austin, TX: The University of Texas at Austin dissertation. Seoul: Hankook Publisher.

Rhee, Seongha. 2013. I know I'm shameless to say this: Grammaticalization of the mitigating discourse marker *Makilay* in Korean. *Procedia – Social and Behavioral Sciences* 97. 480–486.

Rhee, Seongha. 2014. "I know you are not, but if you were asking me": On emergence of discourse markers of topic presentation from hypothetical questions. *Journal of Pragmatics* 60. 1–16. https://doi.org/10.1016/j.pragma.2013.10.005

Rhee, Seongha. 2015. On the emergence of Korean markers of agreement. *Journal of Pragmatics* 83. 10–26.

Rhee, Seongha. 2016. LP and RP in the development of discourse markers from "what" in Korean. *Journal of Historical Pragmatics* 17 (2). 255–281.

Rhee, Seongha. 2020a. On determinants of discourse marker functions: Grammaticalization and discourse-analytic perspectives. *Linguistic Research* 37 (2). 289–325. https://doi.org/10.17250/khisli.37.2.202006.00

Rhee, Seongha. 2020b. Pseudo-hortative and the development of the discourse marker *eti poca* ('well, let's see') in Korean. *Journal of Historical Pragmatics* 21 (1). 53–82.

Rhee, Seongha. 2021. From truth to reality to effect: The journey of *sasilsang* in Korean. *East Asian Pragmatics* 6 (3). 331–353.

Rhee, Seongha. 2024. On the emergence and pragmatic functions of discourse markers of interruption: a case in Korean. In Elena Graf & Ulrich Schweier (eds.), *Pragmaticalization: Language change between text and grammar* (Trends in Linguistics studies and monographs; TiLSM), 185–210. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110761139-006

Schourup, Lawrence. 1999. Discourse markers. *Lingua* 107. 227–65.

Sohn, Sung-Ock S. & Stephanie Hyeri Kim. 2016. The interplay of discourse and prosody at the left and right periphery in Korean: An analysis of *kuntey* 'but.' In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change,* 221–249. Leiden: Brill.

Tagaliamonte, Sali. A. 2021. Wait, it's a discourse marker. *American Speech* 96 (4). 424–449.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press. https://doi.org/10.1525/9780520912793

Traugott, Elizabeth C. 2003. From subjectification to intersubjectification. In Raymond Hickey (ed.), *Motives for language change*, 124–139. Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization,* 29–71. Berlin: Mouton De Gruyter. https://doi.org/10.1515/9783110226102.1.29

Traugott, Elizabeth C. 2014a. Intersubjectification and clause periphery. In Lieselotte Brems, Lobke Ghesquière & Freek Van de Velde (eds.), *Intersubjectivity and intersubjectification in grammar and discourse,* 7–27. Amsterdam: John Benjamins. https://doi.org/10.1075/bct.65.02trau

Traugott, Elizabeth C. 2014b. On the function of the epistemic adverbs surely and no doubt at the left and right peripheries of the clause.' In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change,* 72–91. Leiden: Brill.

Traugott, Elizabeth C. & Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486500

Waltereit, Richard. 2002. Imperatives, interruption in conversation, and the rise of discourse markers: A study of Italian *guarda*. *Linguistics* 40 (5). 987–1010. https://doi.org/10.1515/ling.2002.041

Zhang, Xueping. 2014. Yīdàn and wànyī ['Once' and 'if']. *Shìjiè Hànyǔ Jiàoxué* [*World Chinese Language Teaching*] 28 (2). 212–221.

Bionote:
Sunhee YAE is an Associate Professor in the College of General Education at Chung-Ang University in Seoul and Anseong, Republic of Korea. She received her MA in Language and Linguistics from University of Essex in 1998 and her PhD in English Linguistics from Hankuk University of Foreign Studies in 2002. Her research interests include grammaticalization, discourse analysis, pragmatics and cognitive linguistics.
e-mail: syae@cau.ac.kr
https://orcid.org/0000-0002-3313-9725

Сведения об авторе:
Сунхи ЯЭ – доцент Колледжа общего образования в Университете Чунг-Анг в Сеуле и Ансоне, Республика Корея. В 1998 г. она получила степень магистра в области языка и лингвистики в Университете Эссекса, Великобритания, и в 2002 г. – степень PhD в области английской лингвистики в Университете иностранных исследований Хангук, Сеул, Республика Корея. Ее научные интересы включают грамматикализацию, анализ дискурса, прагматику и когнитивную лингвистику.
e-mail: syae@cau.ac.kr
https://orcid.org/0000-0002-3313-9725

Research article / Научная статья

From objective to subjective and to intersubjective functions: The case of the Thai 'truth'-lexeme

Kultida KHAMMEE 🔟 ✉

University of Phayao, Phayao, Thailand
✉annenfants@gmail.com

Abstract

Borrowed lexemes developing into discourse markers (DMs) are uniquely valuable research topics from contact linguistic and grammaticalization perspectives. Although a large number of Sino-Thai lexemes are commonly used, there has not been any serious attempt to analyze such lexemes. The aim of this study is to describe diverse discourse functions of the Sino-Thai lexemes involving *ciŋ* 'true' drawing upon corpus data, to analyze their development with respect to grammaticalization mechanisms, to identify their DM properties, to compare with the developments reported in other languages, and to construct a conceptual-functional network. The data obtained from diverse sources, including historical and contemporary dictionaries, online resources, reference grammars, contemporary corpora, among others show that the lexeme *ciŋ* 'true' denotes 'true, truth, real' (objective meaning) as a lexical word, but it also functions adverbially as an intensifier marking 'surely, definitely' (subjective meaning) which is a natural development of its lexical meaning. In the domain of discourse, the lexeme acquired a number of interactional functions (intersubjective meaning). A review of the Thai DMs based on *ciŋ* 'true' (truth-DMs) in light of grammaticalization parameters shows that changes characterizable as desemanticization, extension, and decategorialization are observable, whereas erosion is either not observed or its reverse is often found. Despite variable degrees of semantic bleaching, the semantic change in the movement from objective to subjective, and further to intersubjective meanings, is prominent. The functional similarity between truth-DMs across languages notwithstanding, the differences are much greater. These findings contribute to the understanding of the role of source semantics as well as its limits in the developmental paths of DMs.

Keywords: *discourse marker, grammaticalization, (inter)subjective functions, Thai, truth*

От объективных функций к субъективным
и интерсубъективным: тайская лексема *ciŋ* «истина»

Культида КХАММИ ⬆ ✉

Университет Пхаяо, Пхаяо, Таиланд
✉annenfants@gmail.com

Аннотация

Заимствованные лексемы, развивающиеся в дискурсивные маркеры (ДМ), являются уникальными лингвистическими феноменами с точки зрения контактной лингвистики и грамматикализации. Несмотря на то, что сино-тайские лексемы достаточно многочисленны и частотны, до сих пор не было серьезных попыток их проанализировать. Целью данного исследования является выявление и описание различных дискурсивных функций сино-тайских лексем, в том числе лексемы *ciŋ* «истинный», с опорой на корпусные данные. Среди наших задач – анализ их развития с точки зрения механизмов грамматикализации, выявление их свойств в качестве ДМ, сравнение с развитием, описанным в других языках, и построение концептуально-функциональной сетки. Данные, полученные из различных источников, включая исторические и современные словари, онлайн-ресурсы, справочные грамматики, современные корпуса и т.д., показывают, что лексема *ciŋ* «истинный» обозначает «истинный, истина, реальный» (объективное значение), но при адвербиальном употреблении в качестве усилителя она обозначает «конечно, определенно» (субъективное значение), что является естественным развитием ее лексического значения. В сфере дискурса лексема выполняет ряд интеракциональных функций (интерсубъективное значение). Рассмотрение тайских ДМ на основе *ciŋ* «истинный» («истина – ДМ») в свете параметров грамматикализации показывает, что изменения, характеризуемые как десемантизация, расширение значения и декатегоризация, лежат на поверхности, их можно наблюдать, тогда как эрозия носит скрытый характер, либо не наблюдается. Несмотря на различную степень десемантизации, семантическое изменение в направлении от объективного значения – к субъективному и далее – к интерсубъективному является значительным. Несмотря на функциональное сходство «истина – ДМ» в разных языках, различия между ними весьма значимы. Полученные результаты способствуют пониманию роли исходной семантики, а также ее ограничений в процессе развития дискурсивных маркеров.

Ключевые слова: *дискурсивный маркер, грамматикализация, (интер)субъективные функции, тайский язык, «истина»*

1. Introduction

Thai and Chinese belong to different language families, with Thai in Kra-Dai (Tai-Kadai) family, and Chinese in the Sino-Tibetan. Although the two languages are not genealogically connected, they share a number of typological features as a result of both geographic propinquity and China's historical cultural-intellectual leadership in the region. As a result, Thai has retained a large number of words borrowed from Middle Chinese (Haarmann 2012[1986]: 165, Suthiwan & Tadmor

2009: 601, among others). One such example is the Thai lexeme *ciŋ* 'true/real' (*ciŋ* per IPA, *ching* per the Royal Thai General System, *jing* per the Enhanced Phonemic Transcription, and จริง in Thai script) which is borrowed from Middle Chinese *cin* 'true, real, sincere' (Suthiwan & Tadmor 2009: 613), and which has been the root of many derived words and expressions that have developed into discourse markers (DMs). Despite the prevalent use and polyfunctionality of the Thai DM *ciŋ* and its formal and functional relatives, collectively referred to as '*ciŋ*-DMs', they have received no scholarly attention to date. This paper, therefore, intends to fill the research gap. The goal of this study is to explore the diverse functions of the *ciŋ*-DMs in contemporary Thai, drawing upon corpus data, to analyze their development from a grammaticalization perspective, to compare the functions with those reported in other languages, and to create a conceptual-functional network. Thus, this paper aims to answer the following research questions: (i) what are the DM functions of the truth-DMs of the Chinese origin, (ii) to what extent does their development conform to grammaticalization parameters, (iii) what properties do they have, (iv) how similar or different are they from the corresponding DMs in other languages, and (v) what kind of conceptual-functional network do they form?

This paper is organized in the following manner. After this introductory section, Section 2 provides a brief description of the theoretical framework; Section 3 describes the research methods; Section 4 exemplifies the lexical uses and DM uses of *ciŋ* and its relatives; Section 5 analyzes the development of *ciŋ*-DMs in view of the grammaticalization parameters, discourse marker properties, crosslinguistic perspectives, and conceptual-functional networks; and Section 6 summarizes major findings and concludes the paper.

2. Theoretical framework

The theoretical framework for this study is grammaticalization theory as pioneered by Meillet (1912), and further elaborated and refined by Kuryłowicz (1975[1965]), Lehmann (2015[1982]), Heine et al. (1991), Heine (1992), Hopper and Traugott (2003), and many other theorists. For analysis of grammaticalization of DMs, the author of this research referenced works by Heine (2013) and Heine et al. (2021), and the discussion of grammaticalization and language contact is drawn largely from Heine and Kuteva (2005), and Shibasaki and Higashiizumi (forthcoming). The analysis of conceptual motivations in the form of networks is due to Narrog and Ito (2007), and Narrog (2010). This synchrony-based reconstruction approach principally draws on the uniformitarian principle (Labov 1994, Romaine 1982), arguably the most fundamental precept in grammaticalization and in historical linguistics more generally. It is particularly notable that Labov (1994: 157) states that "we have no other choice." with respect to methodological uniformitarianism.

3. Data and research methods

The present research is a conceptual, descriptive analysis of linguistic data, and it has involved no experimentation. The data have been collected from diverse sources, including historical and contemporary dictionaries, lexica, online resources, reference grammars, contemporary corpora, and, importantly, native-speaker intuitions of the author.

The corpus data were taken from two major sources: the Thai National Corpus and the Thai Drama Corpus. The Thai National Corpus is an online, searchable, 33.4-million-word corpus developed by Chulalongkorn University in Bangkok, Thailand. The data mostly date between 1988 and 2017. The Thai Drama Corpus is a 163-thousand-word corpus compiled by Kyungeun Park in 2010. It contains the scripts of 94 episodes from the 14 most popular TV dramas aired between 2005 and 2009. [1] As both of these corpora provide only contemporary data, historical information on the Thai language was garnered from other sources such as etymology dictionaries and lexica.

4. Results

Data collection and analysis of usage of *ciŋ* and its related forms rendered the following results at the lexical and discourse levels.

4.1. Thai lexical use of ciŋ

The lexeme *ciŋ* is used primarily as either an adjective denoting 'true, truthful, real, serious, authentic, original, etc.' or an adverb denoting 'really, truthfully, well, etc.' (*Royal Thai Dictionary* 2011). The word is frequently used for quality marking, as in *khɔ̌ɔŋ ciŋ* 'genuine article', *rûaŋ ciŋ* 'true story', *phûut ciŋ* 'serious talk', *tham ciŋ* 'act sincerely', etc. Its other common usage is for degree marking, as in *talòk ciŋ* 'very funny', *sanùk ciŋciŋ* 'very enjoyable', *nâasɔ̌ncay ciŋciŋ* 'very interesting', etc. (As in the last two examples, *ciŋ* is often reduplicated for emphasis.)

The lexical usage of *ciŋ* is largely identical to that of its etymon, the Chinese *cin* (眞/真). According to SEALANG (n.d.; see entry for 'จริง' /*ciŋ*/), the Middle Chinese *ćin* (or *zhen*, *tsyn*) denoted 'true, real, factual, genuine, actual, substantial, really, truly' and 'highest sincerity one is capable of', suggesting that it had adjectival, adverbial, and nominal usages.[2]

---

[1] Special thanks go to Professor Kyungeun Park for granting free use of her corpus for this research.
[2] In Chinese historical periodization, Middle Chinese spans from the 4th to the 12th centuries. The SEALANG data, available at sealang.net/thai/chinese/middle.htm, are based on the works by Thomas Chin, William Baxter, Sergei Starostin, and others. In addition to lexical uses, the SEALANG data indicates the word was used as a family name. According to Bo Hong (p.c.), the Chinese etymon of Thai *ciŋ* is *qing* (情) 'true', which was later replaced by *cin* (眞/真). Despite the possibility of etymological difference, the Thai *ciŋ* being the truth-DM remains unaffected.

4.2. The ciŋ *discourse markers*

Thai *ciŋ*-DMs carry diverse discourse functions. Functional classification may vary greatly depending on the granularity of an analysis, because most functions of a form, by virtue of sharing an origin, are closely related to each other and may be consolidated into a more comprehensive category, or may be divided into multiple categories by means of fine-grained feature analysis. The present research has identified seven related, yet distinct functions of the *ciŋ*-DMs that are reasonably well-defined either by the form's taking different paths of conceptual extension or by the semantic contribution of the forms participating in the construction of polylexemic DMs, including: (dis)agreement response token, sudden realization/remembrance, surprise, confirmation solicitation, perspective shift, elaboration, and emphasis.

4.2.1. (Dis)agreement response token

Among the most common uses of the Thai DM *ciŋ* is its function as an agreement response token. It is widely observed across world languages that lexemes denoting 'true, real, right' often function as markers of agreement or confirmation response. Examples of this phenomenon include the English *right* (Gardner 2001, Bolden et al. 2023), Spanish *en efecto* 'in truth' and *efectivamente* 'truly' (Garcés 2014), German *stimmt* 'right, true' (Betz 2015), Hebrew *naxon* 'right, true' (Maschler & Shapiro 2016), Korean *cincca* '(the) real thing' (Rhee & Zhang 2024), and Korean *maca* '(that's) right', (Seongha Rhee p.c.). The conceptual motivation of [true > 'yes' 'I agree'] is straightforward. Intuitively, the development of agreement-marking DMs that arose from 'true, real, right' indeed seems to be common across languages. This usage of the Thai DM *ciŋ* is exemplified in the following excerpt:[3]

(1) Some villagers are arguing with Seng, a hunter. One couple are against hunting and tell the hunter to respect the lives of animals. The hunter refuses to listen and says that those animals are just animals and that their lives do not matter.
Uncle: [Human beings are not the only species in the world, Seng. There are other species, too.]
Aunt: ***ciŋ***
         DM
         '**DM ((I) agree!)**'
Seng: [Hmm… talking to old people makes me mad. Anyway, I'll be the future leader of this world. Come on! Old people in this village don't know what I have experienced.]

(2008 Drama, *Thida Wanon* Episode #2)

---

[3] In the examples presented in this paper, lines that do not require morphemic glosses are given in English translation within square brackets, and supplementary information not indicated in the original text is provided in parentheses for clarity in translation.

Other uses are closely related to agreement marking. For instance, as observed by Gardner (2001), the agreement DM may be used for back-channeling without claiming the turn, which is also the case in Thai. Furthermore, when the DM *ciŋ* accompanies a negation marker such as *mây* 'no, not', the DM then functions as a marker of disagreement, a conceptually straightforward functional extension with the operator of negation. Thus, the two forms *mây ciŋ* 'not true' and *mây ciŋ rɔ̀ɔk* 'not true' are used as disagreement response tokens. The two forms are nearly identical in function but the particle *rɔ̀ɔk* has the softening effect, thus *mây ciŋ* is more assertive and definitive than *mây ciŋ rɔ̀ɔk*. These two disagreement markers are exemplified in the following excerpts:

(2) Nuan-Prang and Ong-In are female friends. Ong-In comes from a rich family unlike their male classmate Sao, who is from a poor family. Nuan-Prang and another classmate named Motdaeng are asking Ong-In if she loves Sao. Ong-In hides her true feelings.
Nuan-Prang: [Do you love Sao?]
Ong-In: [Who told you that?]
Motdaeng: [Myself. I told myself. I saw your behavior before him. You must be in love with him a lot. Come on, admit it!]
Ong-In: **_mây.ciŋ_** *Mótdɛɛŋ* *chăn* *khɔ̌ɔ* *patìsèet*
          DM   [name]   I   beg   deny
          chăn   *mâydây*   *chɔ̂ɔp*   *kháw*   *sàk*   *nɔ̀y*
          I   cannot   like   him   just   little
          phûuchaay *àray* *mâydây-rûaŋ* *sàk*   *yàaŋ* *bûɯ* *bûɯ*   *thûɯ* *thûɯ*
          man   what incompetent   just   type   dumb dumb   dull   dull
          mây-aw-nǎy *mii* *fɛɛn* *lɛ́ɛw* *ìik-tàaŋhàak* *chăn* *mây* *chɔ̂ɔp*
          useless   have girlfriend PST as.well   I   not like
          '**DM (= No, it's not true!)** Motdaeng! I deny! I don't like him at all. He is a loser, dull and useless. Also, he (already) has a girlfriend. I don't like him.'
Motdaeng: [OK, OK, I believe you. But why are you so angry?]
                              (2007 Drama, *Aphi Mahuema Maha Setthi* Episode #3)

(3) Saranat and Rungthip are exercising in the evening at the front yard of the house. They see Pawi, their son, about to leave the house. They talk to him and find out that Pawi and Nit (their daughter-in-law) have argued. Rungthip walks into the house to talk to Nit who is crying and packing her belongings.[4]
Rungthip: [Oh! where are you going, Nit?]
Nit: [Ah… I am going to stay at sister Ni's house.]
Rungthip: [Umm…you are not a person who always runs away from problems. I know that Pawi loves you.]

[4] A few notable peculiarities of address or person reference exist in Thai. Thais often use nicknames that are typically monosyllabic or truncated names (as *Wa* for *Wathit* in (11)) since most Thai names are polysyllabic and difficult to remember or say. Kinship terms, e.g., *phîi* 'older sibling', are often used in non-kin relations for affection, as in (6) and (12). It is also common to refer to oneself or one's addressee in the third person, or to use one's own (nick)name (as *Nit* in (3)) or addressee's name (as *Phi* in (11)), instead of first person and second person pronouns.

Nit:	***mây.ciŋ.rɔ̀ɔk***	khâ	kháw	ɔ̀ɔkcà	klìat	Nit	cà-taay
	DM	F:PTCL	3SG:M	tend	hate	[name](=I)	extremely

'DM (= It's not true). He really hates me.'

Rungthip: [You two are so stubborn. Anyway, you were hurt so much and now you are crying. I have never seen you cry. Put your clothes back in the closet, please. Trust me.]

(2006 Drama, *Khing Ko Ra Kha Ko Raeng* Episode #9)

4.2.2. Confirmation solicitation

Another function of *ciŋ*-DMs, closely related to agreement marking, is that of confirmation solicitation. The confirmation or agreement function of REAL/TRUE-DMs has been discussed in a number of studies, notably regarding the English *(that's) right* (Schegloff 1996). The confirmation/agreement function can easily be extended to confirmation solicitation by means of question markers or prosodic variation (e.g., *That's right. – Is that right?; Right? – Right.*). In Thai, three variant DMs carry this function of confirmation solicitation: *ciŋ pàaw, ciŋ rɔ̌ɔ* and *ciŋ rɔ̌ɔ pàaw,* all involving markers resembling a 'P or not P' question. It is evident from their formal shape that these three DMs are closely related in form, as they exhibit variable degrees of erosion, and involve, in addition to the central form *ciŋ* 'true'*, rɔ̌ɔ* 'or' (a colloquial form of *rɯ̌ɯ* 'or'), and *pàaw* (an eroded form of *plàaw* 'no, nothing, plain'). Thus, their lexical sources may be traced to *ciŋ rɔ̌ɔ* 'true or', *ciŋ pàaw* 'true (or) not', and *ciŋ rɔ̌ɔ pàaw* 'true or not'. While these DMs may occur at either left-periphery or right-periphery, a survey of their occurrence patterns in the reference corpora shows distributional asymmetry with strong preference for the right-periphery.[5] Solicitation of confirmation, judged from the context, is based on the genuine question '(is it) true or not?' or 'do you agree?', but it may also carry a skeptical overtone when the speaker is uncertain about the veracity of the interlocutor's statement, e.g., 'is that really true?', as in the following excerpt.

(4) Duean comes to meet her ex-boyfriend Songchai at work. She throws herself into his arms in a very amorous way. She is in fact not interested in him anymore but feigns intimacy in order to win his favor to solve her problem with his help.

Duean: [Songchai.]
Songchai: [Hey, Duean.]
Duean: [I missed you so badly. I wanted to see you.]

Songchai:	***ciŋ.rú.pàaw***	núk	wàa	thamŋaan	phləən	con	lɯɯm	chán	sá-ìik
	DM	think	that	work	enjoy	till	forget	I	PTCL
	léɛw	thîi	bɔ̀ɔk	wâa	mii	thúrá	dùan	nà	rûaŋ	àray
	then	that	tell	that	have	business	urgent	PTCL	subject	what

---

[5] As Reijirou Shibasaki (p.c.) kindly points out, the left/right distinction is problematic with the languages with right-to-left writing conventions, e.g., Arabic, Hebrew, and some traditional styles of writing in East Asian languages. Following the spirit of the researchers who pioneered the notions of left- and right-periphery, the left-periphery can be equated with 'before' the clause and the right-periphery with 'after' the clause.

'**DM (Really?)** I thought you enjoyed your work and forgot me (completely). You said you had an urgent issue to talk with me. What is it?'

Itti: [Umm...the coming issue of *Real Man* has to be finished today. But Mr. Wa has not approved the final draft yet.]

Songchai: [Oh (stressed face)… Where is he now?]

(2005 Drama, *Song Sa-ne-ha* Episode #7)

Ciŋ-DMs can indicate confirmation request, skepticism, disbelief, or even challenge, in which cases they may be translated as 'what are you talking about?', 'get real!', or 'I can't follow your reasoning, etc.' This type of usage is well illustrated in the following excerpt:

(5)   Three students from poor families, Boem, Sao, and Eak, are discussing finding a part-time job to earn money to pay for their study.

Boem: [Hey, this one! Premium Sport Club, the club for high-society people. They are recruiting workers! Well-paid! If you are interested, let's go to apply for a job together this evening.]

Sao: [What? Let me see. What kind of work can we do?]

Boem: [Of course, we will apply for a supervisor position.]

Eak: *hǒo!*   *ciŋ.rɔ̌ɔ*
   INTJ   DM
'Oh, **DM (= What are you talking about?)**'

Boem: [I don't want to be a laborer, of course! Alas, (why are you questioning it?)]

(2007 Drama, *Aphi Mahuema Maha Setthi* Episode #3)

4.2.3. Emphasis

Another function of *ciŋ*-DMs is to add emphasis to an ensuring statement. Considering the source semantics of *ciŋ*-DMs, such a functional extension seems well-motivated, and is widely reported across various languages (cf. the intensifier functions of 'reality, fact' lexemes, Kim 2003, Yaguchi et al. 2010, Gray 2012, Ricca & Visconti 2014, Rhee 2016, 2021, among many others). The *ciŋ*-DM for this function is *khwaam ciŋ* 'truth,' in which *khwaam* is a nominalizing prefix (cf. *khwaam-sùk* [NOMZ-happy] 'happiness'). This development is analogous to the English emphasis markers used at the left-periphery *fact* and *the fact,* which are the truncated forms of *the fact is that* (Kim 2003). As *khwaam ciŋ* typically occurs at the left-periphery, its reference is cataphoric, serving as a preface to noteworthy information that is about to be presented, effectively saying, 'let me tell you this' or 'this is important'. This usage is exemplified in the following excerpt:

(6)   Pawi (= Wi) and Nit got married through an arrangement. Pawi's lover Phiangphen (= Phen) is jealous and does not like Nit. She and her close friend, Bencharat (= Ben), are quarreling with Nit.

Nit: [You spinsters! I think you and Phiangphen might not find any good men to be your husbands.]

Bencharat: [You, b****!]

Nit: [Hey, don't go to Singapore. Go to Cannes because you are going to be an old maid in Cannes*.]

(later…. Phiangphen is calling her lover Pawi to tell him about the quarrel she and her friend had with his wife Nit.)

Phiangphen: *phanrayaa phîi*   *Wi*   *nîa ráaykàat mâak* *ləəy ná-khá*
               wife   brother [name](= you) DM evil   very   really PTCL
               ***khwaam.ciŋ** Phen*   *kôɔ*   *mây yàakcà*   *thoo maa*   *rópkuan*
               DM   [name](=I) also   not   want   call   come disturb
               phîi   *Wi*   *léɛw lâ-khâ tɛ̀ɛ-kɔ̂ɔ òt*   *sŏŋsăan phîi*   *Ben mâydây.*
               brother [name] then   PTCL   but   avoid pity   sister [name] cannot
               thùuk man   *dàa*   *sá*   *sàat-sĭa-thee-sĭa*   *ləəy*   *nâ-khá*
               PASS 3SG:PEJ scold like   severely   really   PTCL
               'Your wife is so mean. **DM (= This is very important),** I didn't want to call to disturb you. But poor Ben (I can't help pitying her)! She got verbally attacked very severely by her (your miserable wife Nit). [Hey! You don't want to say anything?]

Pawi: [I… I have no comment. That's all? I have things to finish.]

Phiangphen: [Yes. That's all.]

               (talks to herself) [He is not jealous about her chasing men. It means that he does not love her. (That's good.)]

                                  (2006 Drama*, Khing Ko Ra Kha Ko Raeng* Episode #9)

(**Cannes* is a homophone of 'house-beam', where men cannot find any woman if she is sitting on it.)

4.2.4. Surprise

The next function of *ciŋ*-DMs is that of marking surprise. Lexemes denoting 'real, true' developing into mirative DMs have been reported by Maschler and Estlein (2008) for Hebrew *be'emet* 'in truth'. Thai *ciŋ*-DMs are similar in that respect. The *ciŋ*-DM with the function is *taay ciŋ*, in which *taay* denotes 'die'. Recruiting 'die' for formation of a DM seems to be similar to the English *dead* used as an intensifier, as in *dead sure, dead right, dead in time* (cf. Blanco-Suárez 2014, Rhee 2016). This usage is illustrated in the following:

(7)   Phatthra is talking on the phone when her sister Pharani comes into the room and overhears some part of the conversation.

Phatthra: *àray ná*   *khá*   ***taay.ciŋ*** *pen*   *àray mâak*   *rú-pàaw*   *khá*
            what PTCL   F:PTCL DM   be   what severe   Q   F:PTCL
            ɔ̆ɔ …   *khɔ̀ɔpkhun*   *mâak khâ*
            DM   thank.you   very   F:PTCL
            'What? **DM (=Oh my God!)** Is she in a grave condition? Oh… Thank you very much.'

Pharani: [What happened, Phat?]

Phatthra: [Nit got clipped by a motorbike last night. But she is alright.]

Pharani: [Oh my God!] (She is shocked and sinks into a sofa.)

                        (2006 Drama, *Khing Ko Ra Kha Ko Raeng* Episode #9)

Since surprise may be of variable strengths and either desirable or undesirable in nature, *taay ciŋ* can mark various kinds of surprise. From a data survey, however, *taay ciŋ* tends to mark a 'happy' surprise, typically occurring in a monologue, unlike the instance illustrated in (7) above. When it is uttered, the prosody includes a soft tone and elongated intonation at the end, e.g., *taay ci~ŋ*.

4.2.5. Sudden realization/remembrance

Another function closely related to the 'surprise' function described above is that of marking sudden realization or remembrance. It has been shown in some studies that certain REAL/TRUE-DMs are used in signaling 'remembering' (e.g., English *that's right,* Heritage 1998, Korean *cham*, Rhee 2021). Since sudden remembrance/realization also involves an element of surprise, the two functions are similar. The difference between them is that the sudden realization/remembrance involves the speaker's memory lapse or inability to perceive something encountered earlier, whereas the surprise function usually involves something the speaker is encountering for the first time. Since sudden realization/remembrance often involves a lapse of memory and thus delinquency, the usage is typically associated with the speaker's feelings of guilt, implying 'It's a shame that only now I remember/realize it.' If the situation involves the speaker's infringement on the interlocutor, the DM carries an apologetic tone, implying 'I'm sorry'. Furthermore, unlike *taay ciŋ* in the surprise function, it tends to be directed to the interlocutor (rather than monologual), spoken fast (rather than slow and elongated) with an excited tone (rather than a soft tone), and apologetic (rather than happy) in the sudden realization/remembrance function. This is exemplified in the following two excerpts:

(8) Sitrang is in front of her daughter's bedroom door. Wondering if her daughter is already asleep, she speaks softly at the door.
Sitrang: [Ong-In, my sweetheart, are you sleeping?]
Ong-In: [I'm sleeping now.]
Sitrang: *ǒo* **taay.ciŋ** **taay.ciŋ** *mɛ̂ɛ* *ləəy* *maa* *rópkuan* *kaan-nɔɔn*
Oh DM DM mother(=I) DM come disturb sleep
khɔ̌ɔŋ *lûuk* *ləəy* *châymǎy* *khá*
of kid(=you) DM Q F:PTCL
'**Oh! DM (I'm sorry)! DM (I'm sorry)!** Did I disturb your sleep?'
Ong-In: [Oh yep. Oops! Nope. You are not disturbing me. But how come you haven't gone to bed this late tonight? Be careful. When you wake up tomorrow, you may not look beautiful (because of lack of sleep).]
Sitrang: [I could not sleep because I'm worried about you.]
(2007 Drama, *Aphi Mahuema Maha Setthi* Episode #3)

(9) Praphot calls Pen, his wife. She forgot to call her brother-in-law Yak to make an arrangement to visit her father-in-law who has been hospitalized. She has been busy working and completely forgot about her promise to call Yak.

Praphot: [Are you OK? Brother Yak said no one called him. (in worried voice)]

Pen: ***taay.ciŋ***   *lɯɯm*   *sĭa*   *sanìt*

              DM   forget   PST   completely

              '**DM (Oh dear!)** (I) completely forgot.'

              (She is disconcerted unable to speak further, and then continues.)

              [I still don't feel well. (I) don't want to spread germs.]

                                    (2009, Fiction, *Phu Ru Phu Tuen Phu Tromtrom*, TNC)

Another type of sudden realization is marked by the periphrastic DM *tὲɛ thîi ciŋ lɛ́ɛw nîa* consisting of *tὲɛ* 'but', *thîi* 'at', *ciŋ* 'true', *lɛ́ɛw* PTCL, and *nîa* 'this:PTCL'. Despite the compositional complexity, this construction is fully univerbated as a single expression which Thai speakers perceive naturally and intuitively rather than analytically. The expression is a unitized DM commonly used to offer an alternative suggestion, translatable as 'Wait, why don't you...?' or 'Oh, I hit upon an idea!' This function closely resembles that of the English REAL/TRUE-DM *actually,* which is used to signal 'I suddenly thought of something better' (cf. Aijmer 2016), as illustrated in the following:

(10) Ong-In wants to go out to work at the university with friends during the weekend. Her mother (Sitrang) tries to convince her to stay home and invite her friends instead, since she is afraid that her daughter will see Sao, the poor man that she does not like.

Ong-In: [Mom, have you seen my sports car key?]

Sitrang: [Where are you going?]

Ong-In: [I have an appointment with friends to work at the university.]

Sitrang: [Friends?]

Ong-In: [Yep. Nuan and Motdaeng... (finds the key) ahh.. I've got it! Thanks, Mom.]

Sitrang: ***tὲɛ.thîi.ciŋ.lɛ́ɛw.nîa*** *thammay*   *nŭu*   *mây*   *chuan*   *yay*   *Nuan kàp*

         DM   why   kid(=you)   not   invite   [title] [name] and

         Mótdɛɛŋ   *maa*   *nâŋlên*   *thîi*   *bâan*   *lɛ́ɛw*   *dĭaw*   *mêɛ*   *hây Aranchara*

         [name]   come   relax   at   home   then   later   mother(=I) ask [name]

         sɯ́ɯ   *khanŏm arɔ̀y*   *arɔ̀y khâwmaa*   *nâŋ*   *thaan kan*   *diikwàa aw mǎy*

         buy   snack   tasty   tasty come   sit   eat   together better   take   Q

         '**DM (Wait, I've got an idea!),** (why don't) you invite Nuan and Motdaeng to come here. I'll ask Aranchara to buy good snack for you guys. (How's that?)'

                            (2007 Drama, *Aphi Mahuema Maha Setthi* Episode #3)

Still another DM of sudden realization/remembrance is *ciŋ sì,* consisting of *ciŋ* 'true' and *sì* PTCL. Its function is nearly identical with the DM *tὲɛ thîi ciŋ lɛ́ɛw nîa* described above, signaling 'I've got an idea' or 'I just remembered'. This function is shown in the following:

(11) At a filming studio, Mr. Daeng, a drama director, is very angry with Philatlak (= Phi) because <u>she forgot to bring the costumes that they need today in shooting. Their assistants, Koen and Chit, are also trying to find solutions.</u>
Koen: [Err… Chit Chit, didn't Lucy say we could use her costumes?]
Chit: [Ah… Lucy said there are no costumes left in her shop because the magazine *Real Man* took them all to shoot their ads.]
Director: [Errrr… This is so disappointing! Cancel everything! No more shooting!]
Koen: [Hey hey hey! Uncle Daeng! Please wait wait wait…]
<div style="padding-left:2em">
*əə   əə   **ciŋ.sì**   nɔ́ɔŋ Phi   naŋsɯ̌ɯ RealMan   nîà pen khɔ̌ɔŋ khun Wathit*
hey hey DM   sister [name] book [name]   DM be of   Mr [name]
chây-máy   raw   nâacà khuy kàp   kháw dâay ná   əə … tὲὲ
Q   we   may   talk with him can PTCL er but
nɔ́ɔŋ Phi   à   yaŋkhoŋ mâydây fɔɔŋ   rûaŋ   bòtsǎmphâat an nán ná
sister [name] PTCL still   not   listen story interview   CLF that PTCL
'Hey hey! **DM (= I've got an idea.)** Phi! The *Real Man* magazine belongs to Wathit? We might talk with him. Er... but you haven't listened to that interview yet, (right?).'
</div>
Philatlak: [<u>Not yet. But I don't get involved</u> with him (her ex-boyfriend) anymore. You can fine me. I'm OK. Or you can deduct it from my pay. Please…please.]

<div style="text-align:right">(2005 Drama, *Song Sa-ne-ha* Episode #7)</div>

4.2.6. Elaboration

Among the most commonly used *ciŋ*-DMs in Thai is *ciŋ ciŋ lέεw,* consisting of *ciŋ* 'true' *ciŋ* 'true', and *lέεw* PTCL. The final *lέεw* is an intensifying particle to mark emphasis, one also used for completion of an event or state, comparable to the English past tense. The DM *ciŋ ciŋ lέεw* signals that an elaboration is to follow, conveying the message 'let me tell you more'. This function also closely resembles that of the English DM *actually*, which signals elaboration, addition, justification, clarification, etc. (Aijmer 2016), as illustrated in the following:

(12) Ong-In (= In) and Sao attend the same university. They like each other but their economic statuses are very different. Ong-In comes from a rich family while Sao is from a poor one. He works part-time at a golf resort, where Ong-In is a VIP member.
Sao: [The world of rich people includes only golf courses and beautiful grass. But in the world of countryside people, plants and soil are their life. That's not a place for high-class girls to hang out in. The volunteer camp in the countryside might not be suitable for you.]
Ong-In: [High-class girls… I really hate that word!]
Sao: **ciŋ.ciŋ.lέεw**   nɔ́ɔŋ In   mây khuancà maa dəən yùu   thîinîi
DM   sister [name] not should come walk at   here
dûaysám nɔ́ɔŋ In   khuancà pay tii-kɔ́ɔp yùu thîi nûun lέεw
even   sister [name] should go play.golf at place that then
phîi   kɔ̂ɔ khuancà pen khon kὲp   lûuk-kɔ́ɔp hây nɔ́ɔŋ In
brother(=I) also should be person collect golf.ball to sister [name]

'**DM (Let me tell you),** you shouldn't be here (the workers' zone). You should go and play golf over there. And I should be the person who collects golf balls for you.'
[If someone sees a VIP customer like you walking with me, it might not be good.]

Ong-In: [By saying this… do you mean that I should get out of here?]

(2007 Drama, *Aphi Mahuema Maha Setthi* Episode #3)

4.2.7. Perspective shift

The last functional category of *ciŋ*-DMs is perspective shift, signaled by the reduplicated *ciŋ ciŋ* 'true true' and *ciŋ ciŋ à,* where *à* is a particle. With these DMs, the speaker signals that they are shifting their perspective from the current one to a new one, whereby they reinstate the crux of the state of affairs, disregarding all peripheral issues, either presented or imagined. It is paraphrasable as 'putting aside all other issues', and may also signal that the speaker is returning to the main issue after digression. Since shift can be highly contrastive on the part of the interlocutor, who is still in the ongoing perspective, this DM is perceived as a signal of forthcoming disalignment. The development of *ciŋ ciŋ* and *ciŋ ciŋ à* is unique in that, in Thai, the general effect of reduplication is intensification rather than disalignment. This function also resembles that of the English DM *actually*, whose core meaning is 'contrast or opposition' (Aijmer 2016) or 'correct[ing] a prior utterance or an implication and emphasiz[ing] divergence' (Haselow 2013). Also notable is that these reduplicative DMs *ciŋ ciŋ* and *ciŋ ciŋ à* carry the function of prefacing not only the shift of the speaker's perspective but also disalignment from the previous speaker, thus often challenging the veracity of the previous speaker's claim or statement. The following excerpts illustrate the functions of perspective shift: one for reinstating the crux of the state of affairs and the other for assuming a new, completely different perspective.

(13) Nuan-Prang invited her friends to a volunteer camp in a rural area because she wanted to see a man there with whom she fell in love. However, she keeps changing her mind as to whether to go or not, and her friends are annoyed.

Ong-In: [What? I just told my dad that we won't go to the camp. Now you want us to register again? Hey! How come you are so capricious?]

Motdaeng: [You keep changing your mind! I'm confused!]

Nuan-Prang: [I am sorry. So, then… I won't go, OK? (feeling guilty)]

Ong-In:
hɤ́əy	*dĭawkɔ̀ɔn*	*dĭawkɔ̀ɔn*	*khɯɯ*	*chăn*	*wâa*	***ciŋ.ciŋ***	*man*	*kɔ̂*
hey	wait	wait	DM	I	think	DM	it	also

lɛ́ɛwtɛ̀ɛ	*thəə*	*ná*	*thâa*	*thəə*	*yàak*	*samàk*	*kɔ̂*	*samàk*
depend	you	PTCL	if	you	want	apply	DM	apply

sì	*aw*	*ləəy*	*dĭaw*	*chăn*	*pay*	*samàk*	*penphɯ̂an*	*kɔ̂dâay*
PTCL	take	PTCL	then	I		go	apply	together	all.right

'Hey! Wait wait! Well, I think **DM (= putting aside all peripheral issues)** it's you who can make a decision. If you want to register to join the camp, do so. Go ahead! I'll go with you.'

(2007 Drama, *Aphi Mahuema Maha Setthi* Episode #3)

(14)   Duean asked Waeo out to settle their conflict.
       Duean: [Hey! Are you still angry with me? I asked you to come here today because
              I wanted to apologize to you for what I did to you.]
              (While talking she takes out a brown envelope from her pocket and puts it
              on the table.)
       Waeo: [What is it?]
       Duean: [It's…a small present.]
       Waeo: [Ah…you rub me on the back after having smacked me on the head.]
       Duean: [I don't mean it. I just want us to be friends like before.]
       Waeo: (Waeo hurriedly puts the envelop in her bag.)

       thîi.ciŋ.à _Duean_ _kàp_ _phîi_   _à_   _kɔ̂ɔ_ _thamŋaan_ _kan_   _maa_
       DM   [name] and   sister(=I) PTCL also   work   together   come
       naan _à-ná_ _man_ _kɔ̌ɔ_ _tɔ̂ŋ_   _mii_ _bâaŋ_ _à-ná_ _ây_   _kràthópkràthâŋ_
       long   PTCL   it   also   must   have   some   PTCL   the   conflict
       kan   _nà_   _tɛ̀ɛ_ _khráŋ_ _nía_ _man_ _rɛɛŋ_   _pay_ _nɔ̀y_   _ná_
       each.other PTCL   but   time   this   it   severe   go   little   PTCL
       'DM (Come to think of it), you and me, we have worked together for a long
       time. Conflict is possible. But this time, it was quite severe.'
       Duean: [I promise that this won't happen again. I have felt guilty all the time.
              Especially, (bad) things that I did to Phi (your close friend)… I feel
              ashamed.]
                                                          (2005 Drama, *Song Sa-ne-ha* Episode #7)

5. Discussion

5.1. Grammaticalization

A number of grammaticalization mechanisms and principles have been proposed in current research, notably by Lehmann (2015[1982]), Heine et al. (1991), Hopper (1991), and Kuteva et al. (2019). This discussion will consider Thai *ciŋ*-DMs in the light of the four parameters proposed in Kuteva et al. (2019), i.e., desemanticization, extension, decategorialization, and erosion. In addition, a brief discussion on (inter)subjectivity will be presented.

5.1.1. Desemanticization

Desemanticization refers to loss of meaning, and is commonly known by its figurative label 'semantic bleaching'. This phenomenon can be observed in some functions of the Thai *ciŋ*-DMs having transformed from the core meaning of *ciŋ* 'true, real, right', to other functions such as sudden realization/remembrance, perspective shift, elaboration, surprise, etc. These functions, however, are not completely devoid of the 'true, real, right' meaning. For instance, the sudden realization/remembrance function is based on the perception that what has been realized or remembered is a truth and thus merits mention. The notion that something is truthful and thus noteworthy also exists in one form or another in the development of the functions of perspective shift, elaboration, and surprise. Despite being related to the original semantics, the named functions are the result of

sufficient semantic bleaching, which has made the relationship largely opaque to contemporary speakers of Thai.

While some functions have undergone substantive desemanticization, some Thai *ciŋ*-DMs still retain much of the source meaning of 'true, real, right', such as those used for agreement (in the form of a statement), disagreement (with a negation marker), emphasis, and confirmation/agreement solicitation (with a question marker). When the source of meaning is retained, it is known as 'persistence' (Hopper 1991), and is common in grammaticalization, suggesting that semantic bleaching is gradual and gradient, and that different functions of the DMs with shared origin may exhibit very different degrees of desemanticization.

5.1.2. Extension

Extension as a grammaticalization parameter refers to expansion of usage context. Usage contexts largely depend on the word class of the lexeme in question. The source lexeme *ciŋ* is mostly used as an adjective and adverb whose primary categorial function is modification of a noun (e.g., *khɔ̌ɔŋ ciŋ ciŋ* 'true article'), an adjective (e.g., *sanùk ciŋ* 'very fun'), or a verb (e.g., *phûut ciŋ* 'talk seriously'). As an adjective, *ciŋ* may occur as a subject complement (e.g., *Man ciŋ* 'It's true' or *Man ciŋ rɯ̌ɯ mây ciŋ?* 'Is it true or not true?') or independently in the contexts of ellipsis (e.g. *Ciŋ* 'True').

The usage contexts of the *ciŋ*-DMs are not restricted to the modifier or complement position. In fact, in line with the properties of DMs discussed in 5.2 below, most instances of *ciŋ*-DMs occur as stand-alone forms. For instance, all excerpts in the foregoing exposition (excluding [13], in which the DM *ciŋ ciŋ* occurs clause-internally, between the main clause and the complement clause [i.e., the medial position]) involve *ciŋ*-DMs either standing alone or at a non-modifying, left-periphery position. Examples such as these make it obvious that *ciŋ* has undergone extension in its development into more various DMs.

5.1.3. Decategorialization

Decategorialization as a grammaticalization parameter refers to the loss of primary category features. The source lexeme *ciŋ* belongs to the main categories adjective and adverb, which notably, belong to the third tier (nouns and verbs being in the first and second tiers, respectively; Heine & Kuteva 2007: 111) in the hierarchical structure of language evolution. Several categorial properties are associated with adjectives and adverbs. In their meticulous description of the parameter, Heine and Kuteva (2007: 40) list salient properties of decategorialization as: (a) loss of ability to be inflected, (b) loss of ability to take on derivational morphology, (c) loss of ability to take modifiers, (d) loss of independence as an autonomous form, increasing dependence on some other form, (e) loss of syntactic freedom, e.g., loss of the ability to be moved around in a sentence in ways that are characteristic of the non-grammaticalized source item, (f) loss of ability to be

referred to anaphorically, and (g) loss of members belonging to the same grammatical paradigm.

Due to the typological features of Thai, which is an isolating and analytic language, only a few of these properties apply, such as (c) the ability to take modifiers, and (e) syntactic freedom. In terms of modifiability, the source lexeme *ciŋ* 'true, real' as a primary category member (adjective, adverb) was once able to take a modifier, typically an adverb (*ciŋ maak* 'very true, very real'), but this ability has been lost with *ciŋ*-DMs. In terms of syntactic freedom, the source lexeme *ciŋ*, as an adjective, was formerly able to be placed in a pre-nominal position, i.e., as a noun modifier, but the *ciŋ*-DMs that may still be classifiable in form as adjectivals can no longer modify a noun. It is noteworthy, however, that the reduplicative *ciŋ ciŋ* is still used as a modifier, but it functions as an intensifying modifier, different from the heterosemous DM *ciŋ ciŋ*.

5.1.4. Erosion

Erosion refers to the loss of phonetic volume. In line with the typological characteristics of an isolating language that is characteristically resistant to formal change, Thai *ciŋ*-DMs evince a minimal level of erosion. Documentation of Thai historical data is limited and thus there is no diachronic grounds for ordering the *ciŋ*-DMs and their functions according to their emergence. In this situation, possible examples of erosion include a set of *ciŋ*-DMs for confirmation solicitation, i.e., *ciŋ rɔ̌ɔ pàaw, ciŋ pàaw,* and *ciŋ rɔ̌ɔ,* if the last two are phonetically eroded from the first, and a pair of the disagreement-marking *ciŋ*-DMs, *mây ciŋ rɔ̀ɔk* and *mây ciŋ,* if the latter is an eroded form of the former. However, this possibility is highly unlikely because the directionality seems to be reversed. This is particularly true with the latter set, *mây ciŋ rɔ̀ɔk* and *mây ciŋ,* because the first is more mitigative than the second, and the difference is due to the presence of the particle *rɔ̀ɔk,* which has a softening effect (see 3.2.2 above). Considering this, *mây ciŋ rɔ̀ɔk* seems to be a modified form of *mây ciŋ,* instead of *mây ciŋ* having been the reduced form of *mây ciŋ rɔ̀ɔk.* Other studies have previously observed that the hypothesized directionality of condensation is reversed with DM developments, namely the study by Méndez-Naya (2006) for the English DM *right*.

There is a more convincing reason to believe that the groups of similar forms involve addition rather than erosion. A prominent, general characteristic of Thai *ciŋ*-DMs is that they often recruit particles and reduplication to create a form with identical, similar, or related functions. This is indeed a common lexicalization and grammaticalization strategy in Thai, called 'polysemy strategy'; examples include 'face' lexicalization (Khammee & Rhee 2022), 'small' lexicalization (Khammee & Rhee 2024), and future markers (Rhee & Khammee 2024), among others. For this idiosyncrasy, although DMs are typically short in form across languages (cf. Brinton 2017, Heine et al. 2021), *ciŋ*-DMs and many grammaticalized forms in general, are often polylexemic in Thai. Therefore, it can be said that the parameter of erosion is not generally observed with Thai *ciŋ*-DMs.

5.1.5. (Inter)subjectification

In terms of semantic change, desemanticization is a common process observable in grammaticalization (see 5.1.1 above). However, as Heine et al. (1991) note, semantic change in grammaticalization cannot be characterized as a uniformly reductive change, because while there is loss there is gain as well, hence 'the loss-and-gain model' (Heine et al. 1991: 110). Many studies have shown that subjectification and intersubjectification are two noteworthy concomitants of grammaticalization, even though these are independent of grammaticalization processes (Traugott 2010). It would appear that what is gained in grammaticalization of DMs is (inter)subjective meanings.

When subjectivity is understood as the relationship to the speaker and his or her beliefs and attitudes, and intersubjectivity as the relationship to the addressee and his or her face, an interesting aspect emerges from the development of *ciŋ*-DMs. The semantics of *ciŋ* 'true, real, right' basically relate to the objective world, i.e., something as existing in reality or as corresponding to the real states of affairs in the world, for example, 'a real person' (*manút ciŋ*) as opposed to a fictional character, 'a true story' (*rûaŋ ciŋ*) as opposed to a fictitious one, 'a real flower' (*dɔ̀ɔkmáay ciŋ*) as opposed to an artificial flower, 'a genuine thing' (*khɔ̌ɔŋ ciŋ*) as opposed to an imitation, etc. The notion of 'real' or 'true' seems to be highly susceptible to the extension into evaluative world, that is, the subjective world. For instance, 'a true person' (*khon ciŋ*) is likely to refer to someone who embodies virtues and desirable qualities. Even *khɔ̌ɔŋ ciŋ* 'a genuine thing' can be used with the evaluative, subjective meaning to refer to a difficult but important event in life such as an entrance exam. Considering that the etymon of *ciŋ* in Middle Chinese denoted 'true, real, factual, genuine, actual, substantial, really, truly' as well as 'highest sincerity one is capable of' (see 4.1 above), it is likely that the lexeme already carried the objective and subjective meanings at the time of borrowing.

Evidently, the objective and subjective meanings have persisted in the development of some functions of *ciŋ*-DMs, for example, in DMs denoting agreement (confirming truthfulness or correctness), emphasis (highlighting focused, genuine qualities), surprise and sudden realization (both highlighting unexpectedness of a newly encountered real-life event or news), etc. As the development of *ciŋ* into *ciŋ*-DMs proceeded further, it acquired diverse interactional meanings, i.e., those of the intersubjective world. For instance, such functions as confirmation solicitation, elaboration, perspective shift, etc. are interactional (by virtue of involving interlocutors) and intersubjective (by virtue of considering the addressee's social and epistemic states). Therefore, from a broad picture, the semantico-functional development of *ciŋ* can be characterized as occurring from objective, to subjective, and on to intersubjective meanings.

5.2. Properties of discourse markers

In their seminal work, Heine et al. (2021: 6) characterize DMs as (a) invariable expressions which are (b) semantically and syntactically independent from their

environment, (c) set off prosodically from the rest of the utterance in some way, and (d) having function that is metatextual, anchored in the situation of discourse, and serving the organization of texts, the attitudes of the speaker, and/or speaker-hearer interaction.

Thai *ciŋ*-DMs are largely invariable expressions, even though some of them contain a seemingly optional but often crucial particle as a component in their construct, e.g., *mây ciŋ* and *mây ciŋ rɔ̀ɔk*, *ciŋ ciŋ* and *ciŋ ciŋ lɛ́ɛw, ciŋ ciŋ* and *ciŋ ciŋ à,* etc., thereby confirming (a). All *ciŋ*-DMs are syntactically independent, typically occurring at left-periphery or as a stand-alone, thus confirming (b). Their prosody as observed in their realization in everyday interaction, although not appearing in the referenced written corpora, is distinctive and set off from the rest of the utterance, thereby confirming (c). As for the metatextuality, Thai *ciŋ*-DMs carry the function of organizing discourse materials, facilitating communicative interaction with the interlocutor, and expressing subjective and intersubjective stances, thus confirming (d). It can be said, therefore, that Thai *ciŋ*-DMs exhibit all diagnostic characteristics of DMs.

5.3. Crosslinguistic comparison

While a full-scale, crosslinguistic comparison is beyond our immediate capacity and is not feasible for the space limitations of this paper, a brief discussion on Thai *ciŋ*-DMs from a crosslinguistic perspective is in order. A number of studies have addressed REAL/TRUE-DMs in individual languages, as briefly listed above in 4.2.1. In particular, most reported cases of REAL/TRUE-DMs carry some functions of the Thai *ciŋ*-DMs, such as agreement (or disagreement with a negation marker, or confirmation solicitation with a question marker), as in the English *(that's) right*, Spanish *en efecto, efectivamente*, German *stimmt*, Hebrew *naxon*, Korean *cengmal, cincca, maca,* etc.[6]

Languages differ with respect to the other, more elaborate functions, largely due to the high level of multifunctionality of DMs and divergent paths taken by the REAL/TRUE-lexemes in the course of their development (cf. Aijmer 1986, Jucker 2002, Aijmer et al. 2006, Defour et al. 2010, Simon-Vandenbergen & Willems 2011, Ricca & Visconti 2014). It is particularly notable that when a language has multiple REAL/TRUE-DMs, their functions may diverge drastically. For instance, the English *actually*, unlike its semantic relative *right*, carries the shift function of disalignment. In other words, English *actually*, like Thai *ciŋ*, signals that there is some incompatibility between the proposition to follow and the one expressed or implied in the context (Haselow 2013, Aijmer 2016). A similar functional development has been reported in Italian with the DM *veramente*, used for a 'mitigated rebuttal' (Ricca & Visconti 2014: 142) and with the French *vraiment* in rebuttal contexts ('contextes réfutatifs', Rodríguez Somolinos 2011). In this context, the discussion of Japanese *shinni* 'truly', developed from the Chinese

---

[6] Solicitation of confirmation or agreement may be marked by an 'appeal intonation contour' (Du Bois et al. 1992: 30, Maschler & Shapiro 2016) instead of a specialized question marker.

etymon 'true', by Higashiizumi et al. (2024) carries an important implication with respect to effects of genres and registers. The authors observe skewing of 'truth'-related words by written and colloquial genres, suggesting that the functional divergence can be observed not only across languages but also within languages.

For our purposes, a more interesting investigation of Thai *ciŋ*-DMs is one in comparison with the development of its etymon in Chinese, the donor language. As briefly alluded to in 4.1 above, the lexical uses of the Thai *ciŋ* largely coincide with those of the etymon *ćin* (or *zhen*, *tsyn*) of Middle Chinese. According to Rhee and Zhang (2024), the Chinese etymon *qing/zhen* signified 'true' and the Chinese lexemes comparable to Thai *ciŋ*-DMs are *zhende* (*zhēn* 'true' + *de* 'of/PTCL') and *zhenshi* (*zhēn* 'true' + *shi* 'be so, be correct'). According to the authors, *zhende* occurred very infrequently until Yuan Dynasty times (1271–1368 CE), and more frequently occurred in the Ming dynasty (14th–17th centuries) and Qing dynasty (17th–1912). Its use as a DM is first attested in the 20th century for various functions, such as emphasis, preface to noteworthy information, surprise, and preface to upcoming disalignment. The DM *zhensi* has a longer history, having first occurred in the Warring States period (5th–3rd BCE). The adverbial usage with the meaning of 'truly' occurred in the 17th century. Its DM usage is first attested in the Qing dynasty times, and it became functionally diversified in 20th century Modern Chinese when it began to be used to express discontent, sudden remembrance, reproach, and annoyance. Although the two DMs originated from the same source, their functions diverged significantly.

A comparison of the grammaticalization scenarios of REAL/TRUE-DMs in Thai and Chinese shows that some functions are common to the two languages, such as emphasis, surprise, and remembrance/realization. Notably, certain functions found in Chinese are not found in Thai, such as preface to noteworthy information, preface to upcoming disalignment, discontent, sudden reproach, and annoyance. Conversely, certain functions found in Thai are not found in Chinese, such as agreement, disagreement, confirmation solicitation, elaboration, and perspective shift. These states of affairs suggest that developmental paths of the DMs that evolved from the same etymon may not exhibit a great degree of similarity. They may show some commonalities, possibly attributable to the common reasoning patterns from the source semantics, but the differences are greater, possibly due to taking different paths of reasoning, with selective focus on particular aspects of meaning or influence of the context. This would be consonant with some earlier observations that lexemes from the same etymon in different languages can be widely different in the development of the DMs from such lexemes to the point of being 'false friends' (e.g., English *actually* and French *actuellement*, Defour et al. 2010).

5.4. Conceptual-functional network

Semantic and functional changes are gradual and gradient since changes are actualized through constant semantic negotiation between interlocutors. Therefore, even large gaps between the meaning and functions in the source lexeme and the

resultant grammatical forms are often nothing more than cumulative effects of incremental changes that have occurred through inferences. This phenomenon is well captured by the Metonymic-Metaphorical Model (Heine et al. 1991) as well as the Invited Inferencing Theory of Semantic Change Model (Traugott & Dasher 2002). The effect of the gradient nature of grammaticalization is evident in the relationship of diverse functions that arise from the same form, and the relationship creates a conceptual-functional network (cf. Narrog & Ito 2007, Narrog 2010). Networks (re-)constructed on the basis of the inventory of synchronic functions rather than functions historically ordered through diachronic trajectory are hypothetical in nature, but this approach 'from synchrony to diachrony' can be particularly useful in contexts where historical data are scarce (cf. Givón 1971, 2015: 1–26, among others).

The core meanings of the source lexeme *ciŋ* are 'true' and 'real', and from these the basic functions of 'agreement' and 'emphasis' have emerged. Diverse functions have developed from these two key concepts, together with additional devices such as particles, reduplication, prosody, and supplementary lexemes. With the help of these additional devices, the *ciŋ*-DMs of agreement have further developed the functions of disagreement (with negation) and confirmation solicitation (with appeal contour or question markers). The emphasis marking has developed through two divergent paths, one first towards elaboration and further towards perspective shift, and another first towards surprise and further towards sudden realization. In the path towards surprise, the addition of the lexeme *taay* 'die, death' seems to have played an important role. The network of paths of lexeme development is graphically presented in Figure 1.[7]
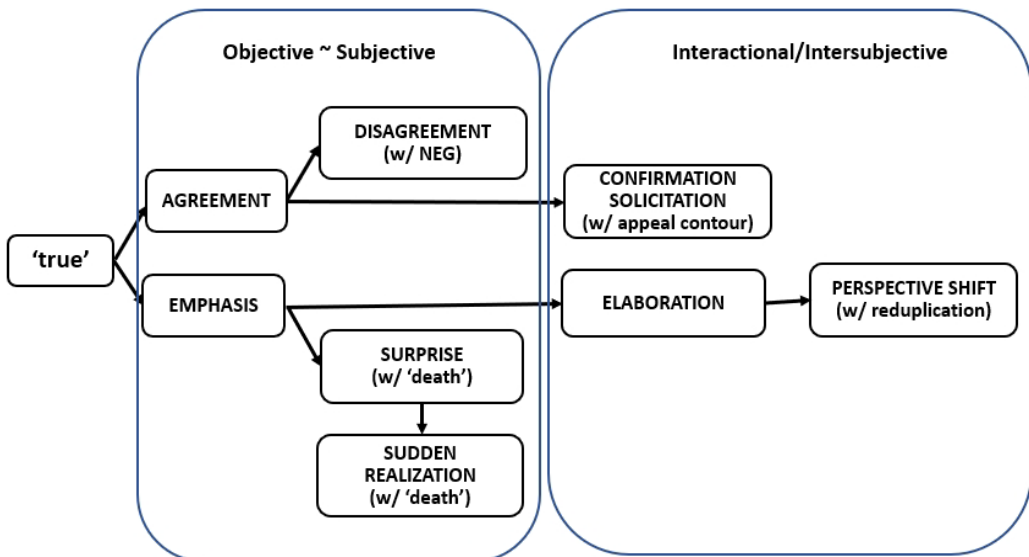

Figure 1. The conceptual-functional network of Thai *ciŋ*-DMs

---

[7] Special thanks go to Reijirou Shibasaki (p.c.) for suggesting a refined network model.

6. Summary and conclusion

This paper aimed to identify the functions of truth-DMs in Thai, to analyze their development, to compare them with corresponding DMs in other languages. The analysis showed that the Thai truth-DMs, developed from the Sino-Thai *ciŋ*, carry the following functions: (dis)agreement response token, confirmation solicitation, emphasis, surprise, sudden realization/remembrance, elaboration, and perspective shift.

A review of *ciŋ*-DMs in light of grammaticalization parameters shows that changes characterizable as desemanticization, extension, and decategorialization are observable, whereas erosion is either not observed or, rather, its reverse is often found. Despite variable degrees of semantic bleaching, the semantic change in the direction from objective to subjective and further to intersubjective meanings, is prominent.

A brief review with other languages, especially Chinese, the donor language, reveals many similarities, supposedly due to the lexical source semantics and common cognitive operations, as well as differences, supposedly due to variation of such cognitive operations or contributions of the participating forms in the periphrastic DMs. An in-depth study comparing the REAL/TRUE-DMs from the same Chinese etymon in other languages is warranted as a future investigation.

Abbreviations

CLF: classifier; DM: discourse marker; F: feminine; INTJ: interjection; PASS: passive; PEJ: pejorative; PST: past; PTCL: particle; Q: question; SG: singular

References

Aijmer, Karin. 1986. Why is *actually* so popular in spoken English? In Gunnel Tottie & Ingegerd Bäcklund (eds.), *English in speech and writing: A symposium,* 119–129. Uppsala: Almqvist & Wilsell International.

Aijmer, Karin. 2016. Revisiting actually in different positions in some national varieties of English. In Francisco Alonso Almeida, Laura Cruz Garcia & Victor M. Gonzales Ruiz (eds.), *Corpus-based studies on language varieties,* 115–143. Bristol: Peter Lang.

Aijmer, Karin, Ad Foolen & Anne-Marie Simon-Vandenbergen. 2006. Pragmatic markers in translation: A methodological proposal. In Kerstin Fischer (ed.), *Approaches to discourse particles*, 101–114. Amsterdam: Elsevier.

Betz, Emma. 2015. Indexing epistemic access through different confirmation formats: Uses of responsive *(das) stimmt* in German interaction. *Journal of Pragmatics* 87. 251–266.

Blanco-Suárez, Zeltia. 2014. The competition between the intensifiers *dead* and *deadly*. In Hilde Hasselgård, Jarle Ebeling & Signe Oksefjell Ebeling (eds.), *Corpus perspectives on patterns of lexis,* 71–90. Amsterdam: John Benjamins.

Bolden, Galina B., Alexa Hepburn & Jenny Mandelbaum. 2023. The distinctive uses of *right* in British and American English interaction. *Journal of Pragmatics* 205. 78–91.

Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: Cambridge University Press.

Defour, Tine, Ulrique D'Hondt, Anne-Marie Simon-Vandenbergen & Dominique Willems. 2010. *In fact, en fait, de fait, au fait:* A contrastive study of the synchronic correspondences and diachronic development of English and French cognates. *Neuphilologische Mitteilungen* 111 (4). 433–463.

Du Bois, John W., Susanna Cumming, Stephan Schuetze-Coburn & Danae Paolino (eds.). 1992. *Discourse Transcription* (Santa Barbara Papers in Linguistics 4). Department of Linguistics, University of California, Santa Barbara, Santa Barbara.

Garcés, María Pilar Gómez. 2014. Gramaticalización y tradiciones discursivas: El proceso de creación de los marcadores de confirmación. *Revue Romane* 29 (2). 264–292.

Gardner, Rod. 2001. *When Listeners Talk: Response Tokens and Listener Stance*. Amsterdam: John Benjamins.

Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society* 7. 394–415.

Givón, Talmy. 2015. *The Diachrony of Grammar*. Amsterdam: John Benjamins.

Gray, Mark. 2012. On the interchangeability of *actually* and *really* in spoken English: Quantitative and qualitative evidence from corpora. *English Language and Linguistics* 16 (1). 151–170.

Haarmann, Harald. 2012[1986]. *Language in Ethnicity: A View of Basic Ecological Relations*. Berlin: Mouton.

Haselow, Alexander. 2013. Arguing for a wide conception of grammar: The case of final particles in spoken discourse. *Folia Linguistica* 47 (2). 375–424.

Heine, Bernd. 1992. Grammaticalization chains. *Studies in Language* 16 (2). 335–368.

Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51 (6). 1205–1247.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework.* Chicago: The University of Chicago Press.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers.* Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2007. *The Genesis of Grammar: A Reconstruction.* Oxford: Oxford University Press.

Heritage, John. 1998. *Oh*-prefaced response to inquiry. *Language in Society* 27. 291–334.

Higashiizumi, Yuko, Reijirou Shibasaki & Keiko Takahashi. 2024. From truth to truly: The case of *shinni* 'truly' in Japanese. *Russian Journal of Linguistics* 28 (4). 843–864. https://doi.org/10.22363/2687-0088-40518

Hopper, Paul J. 1991. On some principles of grammaticalization. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to grammaticalization,* 2 vols. Vol. 1, 17–35. Amsterdam: John Benjamins.

Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Jucker, Andreas H. 2002. Discourse markers in Early Modern English. In Richard Watts & Peter Trudgill (eds.), *Alternative histories of English*, 210–230. London: Routledge.

Khammee, Kultida & Seongha Rhee. 2022. Same and different ways of seeing faces: The cases of Korean and Thai. *The Journal of Linguistic Science* 103. 361–381.

Khammee, Kultida & Seongha Rhee. 2024. Cultural meanings of 'small': Similar yet different semantic networks of diminutives in Thai and Korean. *Selected Conference Proceedings Papers of the SOAS GLOCAL 2022 Conference in Asia.* 188–200.

Kim, Hi-Jean. 2003. *The-Fact-Is-That and related constructions: A usage-based approach and implications in English education.* Doctoral dissertation, Hankuk University of Foreign Studies, Seoul.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World Lexicon of Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.

Kuryłowicz, Jerzy. 1975[1965]. The evolution of grammatical categories. In Eugenio Coseriu (ed.), *Esquisses linguistiques* II, 38–54. Munich: Fink.

Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors.* Oxford: Blackwell.

Lehmann, Christian. 2015[1982]. *Thoughts on Grammaticalization* (3rd ed.). Berlin: Language Science Press.

Maschler, Yael & Roi Estlein. 2008. Stance-taking in Hebrew casual conversation via *beʾemet* ('really, actually, indeed', lit. 'in truth'). *Discourse Studies* 10 (4). 283–316.

Maschler, Yael & Carmit Miller Shaprio. 2016. The role of prosody in the grammaticization of Hebrew *naxon* ('right/true'): Synchronic and diachronic aspects. *Journal of Pragmatics* 92. 43–73.

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* 12. (Reprinted in Meillet. 1948. Linguistique Historique et Linguistique Générale 1. 130–148. Paris: Edouard Champion).

Méndez-Naya, Belén. 2006. Adjunct, modifier, discourse marker: On the various functions of *right* in the history of English. *Folia Linguistica Historica* 40. 141–169.

Narrog, Heiko. 2010. A diachronic dimension in maps of case functions. *Linguistic Discovery* 8 (1). 233–254.

Narrog, Heiko & Shinya Ito. 2007. Re-constructing semantic maps: The comitative-instrumental area. *Sprachtypolgie und Universalienforschung (STUF)* 60 (4). 273–292.

Rhee, Seongha. 2016. On the emergence of the stance-marking function of English adverbs: A case of intensifiers. *Linguistic Research* 33 (3). 395–436.

Rhee, Seongha. 2021. On grammaticalization of truthfulness-based emphatic discourse markers. In Fuxiang Wu, Yonglong Yang & Haiping Long (eds.), *Grammaticalization and studies of grammar* 10, 461–499. Beijing: Commercial Press.

Rhee, Seongha. 2022. Where to go at the end: Polylexicalization and polygrammaticalization of Kaz 'edge' in Korean. *Russian Journal of Linguistics* 26 (3). 571–595. https://doi.org/10.22363/2687-0088-30616

Rhee, Seongha & Kultida Khammee. 2024. Grammaticalization of future-time reference markers in Korean and Thai. *GEMA: Journal of Language Studies* 24 (3). 1–19.

Rhee, Seongha & Lin Zhang. 2024. The way of the truth: The case of Korean DM *cincca* in comparison with Chinese DMs *zhenshi* and *zhende*. *Russian Journal of Linguistics* 28 (4). 818–842. https://doi.org/10.22363/2687-0088-40500

Ricca, Davide & Jacqueline Visconti. 2014. On the development of the Italian truth adverbs *davvero* and *veramente*. In Irma Taavitsainen, Andreas H. Jucker & Jukka Tuominen (eds.), *Diachronic corpus pragmatics,* 133–154. Amsterdam: John Benjamins.

Rodríguez Somolinos, Amalia. 2011. Assertion de la vérité et engagement du locuteur: L'évolution de *voirement* et de *vraiment* en français (XIIe–XVIIe siècle). *Langages* 184. 91–110.

Romaine, Suzanne. 1982. *Socio-Historical Linguistics: Its Status and Methodology.* Cambridge: Cambridge University Press.

Schegloff, Emanuel A. 1996. Confirming allusions: Toward an empirical account of action. *The American Journal of Sociology* 102 (1). 161–216.

Shibasaki, Reijirou & Yuko Higashiizumi. Forthcoming. An introduction. In Yuko Higashiizumi & Reijirou Shibasaki (eds.), *The emergence of pragmatic markers in East Asian languages* (Studies in pragmatics). Leiden: Brill.

Simon-Vandenbergen, Anne-Marie & Dominique Willems. 2011. Crosslinguistic data as evidence in the grammaticalization debate: The case of discourse markers. *Linguistics* 41. 1123–1161.

Suthiwan, Titima & Uri Tadmor. 2009. Loanwords in Thai. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the World's languages – a comparative handbook*, 599–616. Berlin: De Gruyter Mouton.

Traugott, Elizabeth C. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization,* 29–71. Berlin: De Gruyter Mouton.

Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in Semantic Change.* Cambridge: Cambridge University Press.

Yaguchi, Michiko, Yoko Iyeiri & Yasumasa Baba. 2010. Speech style and gender distinctions in the use of *very* and *real/really*: An analysis of the corpus of spoken professional American English. *Journal of Pragmatics* 42. 585–597.

Dictionaries and Internet resources

SEALANG. n.d. sealang.net/thai/chinese/middle.htm.

Royal Thai Dictionary, The (aka *The Royal Institute Dictionary).* 2011. Bangkok: The Royal Society of Thailand.

Bionote:
Kultida KHAMMEE earned her PhD in Applied Linguistics from Mahidol University, Thailand, and has been teaching at University of Phayao, Thailand, since 2012. Her research interests include text analysis, cognitive linguistics, cultural studies, and language change from cross-linguistic and typological perspectives. She has authored and co-authored a number of articles published in *Journal of Pragmatics, GEMA Online® Journal of Language Studies, World Journal of English Language*, among others, and a book chapter in *Culture-Bound Syndromes in Popular Culture* (2024, Routledge).
e-mail: annenfants@gmail.com
https://orcid.org/0000-0002-5620-7487

Сведения об авторе:
Культида КХАММИ получила PhD по прикладной лингвистике в Университете Махидол, Таиланд, преподает в Университете Пхаяо, Таиланд, с 2012 г. Ее научные интересы включают анализ текста, когнитивную лингвистику, культурологию, а также изменение языка с кросс-лингвистической и типологической точек зрения. Она является автором и соавтором ряда статей, опубликованных в *Journal of Pragmatics*, *GEMA Online® Journal of Language Studies*, *World Journal of English Language* и др., а также глав книги *Culture-Bound Syndromes in Popular Culture* (2024, Routledge).
e-mail: annenfants@gmail.com
https://orcid.org/0000-0002-5620-7487

Research article / Научная статья

From *truth* to discourse marker:
The case of *thật* in Vietnamese

Mayumi ADACHI ✉

Research Institute for Languages and Cultures of Asia and Africa,
Tokyo University of Foreign Studies, Tokyo, Japan
✉adachi@aa.tufs.ac.jp

Abstract

Vietnamese *thật* (truth)-based expressions are prevalent in a 17th-century text and contemporary data, however there has been a lack of in-depth studies of such Sino-Vietnamese expressions in terms of their grammaticalization into discourse markers (DMs). The aim of this study is to identify whether linguistic elements from the same source develop into DMs in a similar manner. To achieve this, I compared diachronic data in the 17th century to contemporary Vietnamese. The historical data includes the first Romanized Vietnamese prose written in 1651, while the contemporary data consists of written and spoken materials collected from a language consultant, internet quotes, and naturally occurring conversations. The main findings of this research are as follows: (1) *Thật*-related expressions appear to follow the well-attested grammatical path from "TRUE" to "INTENSIVE" and further into DMs. (2) The subjective and intersubjective uses of *thật* forms are distributed complementarily: subjective usage tends to occur in the utterance-final position, while intersubjective usage typically appears in the utterance-initial position. (3) *Thật* has also combined with other elements to form multiword-unit DMs, indicating that DMs do not always undergo formal reduction. (4) Regarding the borrowing of DMs through language contact, examples from Vietnamese and other languages suggest that lexical elements are initially borrowed as in their original forms and then grammaticalized as DMs in ways that are language-specific yet comparable. These findings provide theoretical contributions to our understanding of the emergence and development of DMs in Vietnamese.

Keywords: *language contact, discourse marker, grammaticalization, intersubjectification, Sino-Vietnamese thật*

От «*истины*» к дискурсивному маркеру:
thật во вьетнамском языке

Маюми АДАЧИ ⓘ ✉

Исследовательский институт языков и культур Азии и Африки,
Токийский университет иностранных языков, Токио, Япония
✉adachi@aa.tufs.ac.jp

Аннотация

Выражения, основанные на китайско-вьетнамской лексеме *thật* «истина» (кит. 實), широко употребляются во вьетнамском языке с XVII в. до наших дней, однако наблюдается отсутствие глубоких исследований с точки зрения их грамматикализации в дискурсивные маркеры (ДМ). Цель данного исследования – определить, переходят ли языковые единицы, происходящие из одного и того же источника, в ДМ схожим образом. Для этого мы сравнили диахронические данные XVII в. с современным вьетнамским языком. Источником исторического материала послужила первая романизированная вьетнамская проза, относящаяся к 1651 г.; современный письменный и устный материал был собран из интернет-источников и естественной коммуникации, а также с помощью интервью. Основные выводы исследования следующие: (1) Выражения с *thật* развиваются по уже засвидетельствованному грамматическому пути: от «ИСТИНА» до «УСИЛИТЕЛЬ» и далее – в ДМ. (2) Субъективное и интерсубъективное использование форм с *thật* распределяется следующим образом: субъективное использование, как правило, встречается в конечной позиции высказывания, в то время как интерсубъективное использование обычно наблюдается в начальной позиции. (3) *Thật* также объединяется с другими элементами, образуя многословные ДМ, что указывает на то, что ДМ не всегда подвергаются формальной редукции. (4) Что касается заимствований ДМ, происходящих в результате языкового контакта, то примеры из вьетнамского языка, как и других языков, показывают, что лексические элементы изначально заимствуются в исходных формах, а затем грамматикализуются в культурно-специфичные, но сопоставимые ДМ. Данные выводы вносят теоретический вклад в понимание возникновения и развития ДМ во вьетнамском языке.

Ключевые слова: *языковой контакт, дискурсивный маркер, грамматикализация, интерсубъективация, китайско-вьетнамский thật*

1. Introduction

Vietnamese *thật* and *thật*-derived discourse markers (DMs), collectively termed as "*thật*-based expressions" (mostly DMs), were prevalent in a 17th century text (introduced in Section 3) and in contemporary data, spoken and written. *Thật* originated from a Sino-Vietnamese lexeme 實 meaning 'full, rich, true' as an adjective and 'fruit' as a noun (Thiều Chửu 2009: 178–179); it is a cognate with corresponding expressions in Chinese, Japanese and Korean. Diachronic change of 實-based expressions has attracted much attention in studies of those languages (for further discussion, see Section 2.2). However, there have been no in-depth studies

of the grammaticalization of such Sino-Vietnamese expressions into DMs. To make theoretical contributions, this study determines whether linguistic elements derived from the same source develop into DMs in similar manners to DMs, that is, through cyclicity (Hansen 2018a, b). To achieve this, I compared *thật*-based expressions in diachronic data in the 17th century and contemporary Vietnamese. The historical data includes *Cathechismvs in octo dies diuisus* (Catechisms in eight days), the first Romanized Vietnamese prose written by Alexandre de Rhodes, a French missionary, in 1651 (for details, see Section 3), while the contemporary data consists of written and spoken materials collected from a language consultant, internet quotes, and naturally occurring conversations.

This study is organized as follows. In Section 2, I present history of Vietnamese and give an overview of relevant research. In Section 3, I present an account of the data used in this study. In Section 4, I present results of the study which are discussed in Section 5 in terms of grammaticalization. Section 6 concludes the paper.

2. Research background

2.1. A short history of Vietnamese and its writing system

Vietnamese is classified as an Austro-Asiatic language. It is an isolating language with a verb-medial and head-initial structure, featuring six tonal distinctions. The dialects of Vietnamese are generally categorized into three main groups: the northern dialect, primarily spoken in Hanoi; the central dialect, spoken in Hue; and the southern dialect, spoken in Ho Chi Minh City (Tomita 2000: 17–20). In this study, I examine the northern dialect. Detailed information on the dialect will be provided in Section 3.

The history of writing in Vietnam is influenced by the country's long cultural connection with China. Accordingly, Vietnamese leaders and scholars borrowed a large number of Chinese words from Chinese written forms (Alves 2007a: 357), which are now known as Sino-Vietnamese words. Regarding Sino-Vietnamese words, Nguyễn (1961) states that they account for 60–70% of modern written Vietnamese texts (see also Alves 2001, 2009, Murakami & Imai 2010), while Miyajima (2019: 809) argues that approximately 70% of Vietnamese words are Sino-Vietnamese (see Shibasaki & Higashiizumi (forthcoming), Note 6 for more details).

Edmondson (2006: 432) summarizes the history of writing in Vietnam as follows: Chinese scripts may have been adopted as a learned script in Vietnamese before the Common Era[1]. Subsequently, a demotic script, *Chữ nôm*, to write purely Vietnamese words, may have been in use as early as the 10th century. Ultimately,

---

[1] Tomita (1988: 761) indicates that the direct control of Vietnam by China from 111 B.C. to 938 A.D. led to the spread of Chinese culture and language in Vietnam. Furthermore, according to Kawamoto (2011: 1901), by the 11th century, the phonetic system of Sino-Vietnamese vocabulary had individualized from Chinese.

these two scripts were replaced by *Quốc ngữ* in the 20th century, a Romanized script developed by Jesuit missionaries in the 17th century. *Chữ nôm* characters are based on Chinese ideograms, while the modern Vietnamese writing system (*Quốc ngữ*) is a phonetic script written using Roman characters (Vũ 2005: 7).

In line with Maspero (1912)[2], Nguyễn (1997: 5) categorizes Vietnamese in the 17th century as Middle Vietnamese, as "reflected in the Vietnamese-Portuguese-Latin dictionary by Alexandre de Rhodes," a French missionary. Missionary literature such as dictionaries and catechisms have been examined as materials for missionary linguistics (Kishimoto & Sirai 2022), which are "modeled according to the traditional Greco-Latin framework" (Zwartjes 2018).

2.2. A brief overview of the relevant studies

Hopper & Traugott (1993: xv) define grammaticalization as "the process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions." According to Kuteva et al. (2019: 443), there is a typological tendency for words meaning 'true' to be grammaticalized as intensive markers. With regard to this point, Rhee et al. (2021) highlighted the Chinese etymon 實-based expressions that denoted 'real(ity), true, full, rich, fruit' and examined how they have been grammaticalized into DMs in languages in the Chinese cultural sphere, including Chinese, Japanese, and Korean. Rhee et al. (2021) point out that the developmental paths of DMs containing the Chinese etymon 實 / 实 / 実 in Chinese, Japanese, and Korean such as 事實上/事实上/事実上 and 其實/其実, shows similarities semantic, syntactic, and discursive changes. Additionally, several articles in this special issue explore words including 真 'true' in different languages, including Khammee (2024)'s work on *cing* in Thai, Higashiizumi et al. (2024)'s discussion of 真に (*shinni*) in Japanese, and Rhee & Zhang (2024)'s analysis of *cincca* in Korean and 真是 (*zhenshi*) and 真的 (*zhende*) in Chinese.[3]

Bui (2012, 2015) lists examples of DMs in Vietnamese that have text connective functions as in (1).

---

[2] Maspero, Henri. 1912. Études sur la phonétique historique de la langue annamite. Les initiales. *Bulletin de l'École française d'Extrême-Orient* 12. 1–124.

[3] Interestingly, the Sino-Vietnamese lexeme *chân* meaning 'true,' which is derived from 真, is not used as a DM in Vietnamese. One reason may be that, as noted in Section 2.1, Sino-Vietnamese words have been largely used in written and formal discourse. This fact offers a possible account for the rarity of spoken and colloquial-oriented elements common to Chinese and Vietnamese (Alves 2007a: 357).

(1)   Vietnamese DMs
   a.   *tuy nhiên* (雖 'though' + 然 'so') [4]
        / *tuy thế* (雖 'though' + 'so')
        / *tuy vậy* (雖 'though' + 'so') 'however'   (Bui 2012: 4, Bui 2015: 23–90)
   b.   *vả lại* ('moreover' + 'again')
        / *vả chăng* ('moreover' + 'surely') 'moreover'   (Bui 2012: 4, Bui 2015: 246–302)
   c.   *dù sao* ('however' + 'how') 'anyway'   (Bui 2012: 91–176)
   d.   *thế* ('so') / *vậy* ('so') 'so'   (Bui 2012: 3)

According to the language consultant, in addition to Bui's (2012, 2015) list, other DMs that consist of 'truth'-related elements, as in (2) exist. Note that *thực* is a variant[5] of *thật*.[6]

(2)   Some other Vietnamese DMs
   a.   *thật ra* (實 'real'+ 'out') 'in fact, actually'
   b.   *thực ra* (寔 'real' + out) 'in fact, actually'
   c.   *thật sự* (實 'real' + 事 'matter') 'in fact, actually'
   d.   *thực sự* (寔 'real' + 事 'matter') 'in fact, actually'
   e.   *trên thực tế* (on + 實 'real' + 際 'occasion') 'in fact, actually'
   f.   *thật tình* (實 'real' + 情 'situation') 'in fact, actually'
   g.   *thực tình* (寔 'real' + 情 'situation') 'in fact, actually'
   h.   *quả tình* (果 'fruit' + 情 'situation') 'in fact, actually'
   i.   *quả thế* (果 'fruit' + 'so') 'in fact, actually'
   j.   *quả thực* (果 'fruit' + 寔 'real') 'in fact, actually'
   k.   *quả thật* (果 'fruit' + 實 'reality') 'in fact, actually'
   l.   *quả nhiên* (果 'fruit' + 然 'so') 'in fact, actually'
   m.   *quả vậy* (果 'fruit' + 'so') 'in fact, actually'
   n.   *kỳ thật* (其 'its/that' + 實 'real') 'in fact, actually'
   o.   *kỳ thực* (其 'its/that' + 寔 'real') 'in fact, actually'

Contrarily, the counterpart expressions in Vietnamese to such cognate-like forms in Chinese, Japanese, Korean and Thai, that is, *thật*, have not yet come under close scrutiny, especially from the perspective of grammaticalization. Broadly speaking, 'truth'-related expressions with 眞/真 'truth' in Chinese, Korean, and Thai have developed in a convergent way to DMs, although the Japanese counterpart *shinni* 'truly, genuinely' has not (yet) reached the DM stage. Therefore, this study uncovers pathways of change from 'truth' to a DM in Vietnamese and

---

[4] The author added glossing with Chinese characters and English.

[5] There is another dialectal variant *thiệt* often used in the southern dialect (Phạm et al. 2009: 388).

[6] In the Nguyễn dynasty (19th century), 實 *thật* was changed to 寔, and its pronunciation was changed to *thực*, because 實 violated the naming taboo of Empress *Tá Thiên*, Emperor *Thiệu Trị*'s mother (Chèn 1982: 570–571). Therefore, it is not expected that *thực* would be observed in the documents of the 17th century.

contributes to a better understanding of the ways DMs develop in this particular cultural region.

Before proceeding, let us examine one of common views of DMs. Brinton (2010: 285–286) clarifies the following formal features of DMs:

(3)　DMs
They are phonologically "short" items that preferentially occur in sentence-initial position. They are syntactically independent elements that are loosely attached to (parenthetical to) their host clause and often constitute a separate intonation unit and as such have scope over global units of discourse (beyond the level of the clause). They occur with high frequency, especially in oral discourse, and may be stylistically stigmatized. (Brinton 2010: 285–286)[7]

Notably, DMs also have text-connective and interpersonal functions, including subjective and intersubjective functions. The concept of speaker stance encompasses both subjectivity and intersubjectivity. Subjectivity refers to the speaker's expression of their own attitudes and beliefs (Lyons 1982: 102), whereas intersubjectivity pertains to the speaker's awareness of the addressee's attitudes and beliefs (Traugott 2003: 126, 2010: 33).

3. Data and methodology

The data used for this study are twofold. For diachronic data, I use *Phép giảng tám ngày* (PGTN; *Cathechismvs in octo dies diuisus*; Catechisms in eight days), which is the first Romanized Vietnamese prose written in 1651 by Alexandre de Rhodes. According to Kishimoto (2018: 58), PGTN is a Latin/Vietnamese bilingual translation. In addition, Schreiter (2015: 8) evaluates PGTN as follows: "The Cathechismus he wrote to evangelize the Vietnamese people was the first Vietnamese-language book to be published in the West. And the Roman alphabet with a diacritical marking system he devised became the basis for the alphabet used to this day in the Vietnamese language."[8] Maspero (1912: 9, Note 1) reports that PGTN relates to the Tonkinese dialect, that is the northern dialect. In the *Từ điển Việt–Bồ–La* (TĐVBL; *Dictionarium Annamiticum Lusitanum et Latinum*; The dictionary of Annamese, Lusitanian and Latin) compiled by Rhodes in 1651 (See 2.1 for its significance in the history of the language) along with PGTN, *thật* is spelled as *thặt*, and it is defined as *verdade* (truth) and *certo* (right) in Lusitanian, that is, Portuguese, as well as *veritas* (truth), *atis* (truly), and *verus* (true) in Latin.[9]

---

[7] See Brinton (2017: 2–8) and Heine et al. (2021: 6) for more recent and comprehensive definitions of DMs as well as pragmatic markers (PMs). For a functional differentiation of DMs and PMs see Shibasaki and Higashiizumi (forthcoming).

[8] Shimizu (2024) is known as a linguistic study based on PGTN, which probes into the historical development of the Vietnamese durative/continuous/progressive marker *đang*, which originated from the Sino-Vietnamese word 當.

[9] The original texts regarding *thật* in *TĐVBL* are as follows. The author added English translations in the brackets to them.

The data of PGTN and TĐVBL is sourced from Wikisource, a free-content digital library maintained by the Wikimedia Foundation. For contemporary Vietnamese, I use data from the northern dialect, which comprise written and spoken data obtained from a language consultant,[10] quotes from the Internet, and family conversations.[11]

For the methodology, I will compare examples from the 17th century with contemporary ones to see how *thật*-based DMs develop over time, with a particular focus on their pragmatic functions. This approach will help to reveal part of the grammatical pathways that lead from 'truth' to a DM in Vietnamese, along with the characteristics of this change.

As Vietnamese lacks inflectional morphology, it is necessary to differentiate word classes based on their functions and distribution for diachronic and typological comparison. The following are brief definitions of terms related to word class in Vietnamese grammar. Nominals can be used as the subject, object of a verb, or object of a preposition. Classifiers are sometimes used before or instead of nominals to indicate specificity and definiteness. Adjectivals can modify noun phrases and act as a predicate. Adverbials modify adjectivals and clauses. Notably, utterance-final particles (UFPs) are used to convey pragmatic information, such as the speaker's emotions and epistemic state, and to indicate the connection and relationship between preceding discourse. Some UFPs can be compounds.

4. Results

In this section, I present the following five types of usages of *thật*, namely, (A) nominal, (B) adjectival, (C) adverbial, (D) utterance-final, and (E) DM, with examples from the 17th century and contemporary Vietnamese. Results indicate that *thật* was not used as (D) utterance-final particles or (E) DM in the historical context. Instead, it primarily appeared in prose as (B) adjectival meaning 'true' and (C) adverbial meaning 'truly' or 'really,' particularly in the context of conveying

---

thặt *verdade* [truth], *certo* [right] veritas [truth], atis [truly], verus [true], a, vm, nói thặt *fallar verdade* [tell the truth]: verum loqui [to speak the truth]. cho thặt: *certamente* [certainly]: certè [sure], profectò [surely], ſanè [really]. thề thặt : *iurar verdade* ([to swear the truth]: verum iurate [to swear the truth]. thề ngay, idem [the same]. thặt thà *ſiel* [faithful], *ſincero* [sincerely]): veridicus [truthful], a, vm, ſincerus [sincere], a. vm.

[10] The language consultant is a female native Vietnamese speaker who was born in Hanoi, Vietnam. She is in her 50s as of 2024.

[11] Data for the present study comprised an 8-hour audio recording of a collection of family dinner conversations. These conversations were recorded in Hanoi by the author in 2014. Glossed excerpts with Japanese translations were included in Adachi (2021) as an appendix. Table 0 provides background information about the family members, all of whom are native speakers of Vietnamese who were born and raised in Hanoi.

Table 0. Background information of discourse participants
in the Vietnamese family conversation database

Name		Occupation
Father	(F)	Wholesaler
Mother	(M)	Language teacher
Daughter	(D)	Junior high school student

"religious truth" to the Vietnamese at that time. In contrast, today *thật* exhibits a range of uses, including (B) adjectival meaning 'real'; (C) adverbial (sometimes followed by a linker *là*) 'really'; (D) as an utterance-final particle meaning 'indeed'; and (E) DM in utterance-initial position, such as *thật ra* (*thật* + out) meaning 'in fact'.

4.1. Vietnamese in the 17th century[12]

4.1.1. Nominal usage

In the texts from the 17th century, the word *thật* does not exist as a noun. It is preceded by the classifier *ſự*, which originated from the Sino-Vietnamese word 事 (meaning 'matter'), to form the noun *ſự thạt* (*sự thật* with contemporary spelling)*,* rendered in English as 'truth,' as shown in (4). Note that in traditional Vietnamese linguistics, classifiers are functionally treated as nouns.

(4)

ᵉ thì[13]	*phải*	*giữ*	*ba*	*ſự*	*một*	*là*[14]	***ſự***	***thạt,***
LNK	have.to	keep	three	CLF	one	LNK	CLF	THẬT
hai	*là*	*ſự*	*coū bàng,*		*ba*	*là*	*ſự*	*cần:*
two	LNK	CLF	fair		three	LNK	CLF	need

'(The second commandment prescribes us not to offend God with our tongue by swearing by God's name in vain. For an oath in the name of God or of divine things not to be a sin), there are three conditions: **truth**, justice, and necessity. (Lit. then, (we) have to keep three matters, one is real one, two is fair one, three is necessary one.)'

4.1.2. Adjectival usage

As an attributive adjective meaning 'true', *thật* follows a noun as shown in (5).

(5)

ᵐ huắng	*lọ*	*người*	*ở*	*thể*	*nầy*	*thì*	*phải*	*tìm*
much	less	people	in	world	this	LNK	have.to	find
biét	*được*	*ai,*	*làm*	***Chúa***	***thật***	*đã*	*sinh*	
know	POSSIB	who	do	lord	THẬT	PFT	give.birth	
cho	*nên*	*blời,*	*đất,*	*muôn vật,*	*mà*	*thờ*	*đấy.*	
let	should	sky	earth	all.things	LNK	worship	DEM.MED	

'How much more should all in this world find who **the true lord** is and the creator of heaven and the earth and all things in order to adore him.'

---

[12] Regarding examples in 4.1, I provide original texts of PGTN cited from Wikisource in the first line, followed by glossing in the second line. The English translation is quoted from Phan (2015). And if necessary, a literal translation is added by the author.

[13] According to Do-Hurinville & Dao (2019: 58, Note 1), *thì* "functions as a temporal anaphoric conjunction or a topic marker." In this paper, I consider *thì* a polyfunctional linker and gloss it as LNK.

[14] The word *là* "is a copular verb" and also "functions as a consecutive conjunction, a focus marker or a complementizer. (Do-Hurinville & Dao 2019: 58, Note 1)." In this paper, I consider *là* as a polyfunctional linker and gloss it as LNK, similar to *thì* (refer to Note 13 in this paper).

In addition, the following is an example of *thật* as a predicative adjective. In (6), *thật* follows an intensifier *rứt* meaning 'very sincere'.

(6) | *ꭍ Mà* | *đức* | *Chúa* | *Ieʃu* | *rứt* | *khiem nhuầng,* |
|---|---|---|---|---|---|
| LNK | saint | lord | Jesus | INT | modest |
| *và* | ***rứt*** | ***thạt*** | *thưa* | *ràng :* | |
| and | INT | THẬT | answer | COMP | |

'(But the exceedingly good Lord, bearing this injury with great gentleness, did not want to keep silent lest he appeared to have lacked respect for the high priest.) He answered with greatest humility (Lit. And the Lord answered very modestly and **very sincerely** that…): ("If I have spoken wrongly, show me what I have said wrongly; but if I have spoken rightly, why do you strike me?"'

4.1.3. Adverbial usage

As an adverb, *thật* can express that the following proposition is true ('without doubt'), modify the proceeding verb ('really die'), and intensify the following adjective ('very foolish'), as shown in (7), (8) and (9), respectively.

(7) | *ʸ **thật*** | ***là*** | *ma quỉ* | *ở* | *tlao̅* | *địa ngõục* | | *là* | *nhà* | *nó.* |
|---|---|---|---|---|---|---|---|---|---|
| THẬT | LNK | devil | be | inside | hell | | LNK | home | they |

'**Without doubt** the demon is in hell which is his home.'

(8) | *ʰ ʃau hết* | *đần bà* | *có* | *tội* | *lám,* | *vì* | *mlời* | *thật* |
|---|---|---|---|---|---|---|---|
| finally | woman | have | sin | INT | because | sky | THẬT |
| *đức* | *Chúa* | *blời* | *răn,* | *mà* | *chẳng* | *giữ* | *thì* |
| saint | lord | sky | admonish | LNK | NEG | keep | LNK |
| *đe* | ***chết*** | ***thạt,*** | | | | | |
| threaten | die | THẬT | | | | | |

'Moreover, the woman sinned most grievously by doubting the order and the death threat that God had formally announced (Lit. Finally, the woman had much sin, because the true lord in the sky had admonished that (if she would) not follow (the order) then (she would be) threatened to **really die**)).'

(9) | *ᵉ ví bàng* | | *có* | *ai* | *giã* | *nhà,* | *má* | *chẳng* |
|---|---|---|---|---|---|---|---|
| if | | have | someone | farewell | house | LNK | NEG |
| *giã* | *Chúa* | *nhà,* | ***thạt*** | ***là***[15] | ***đại*** | *mà* | *chớ.* |
| farewell | master | house | THẬT | LNK | foolish | LNK | UFP |

---

[15] When the adverbial *thật* modifies a predicate adjective as an intensifier, inserting the marker *là* after *thật* is grammatically optional. Lien (2009: 748) analyzes the copula-derived *si⁷* 是 as a focus marker that follows *sit⁸* 實 'really' based on the following example of early Southern Min. Similarly, in Vietnamese, copula-derived marker *là* that follows *thật* and precedes an adjective can also be considered a focus marker, as demonstrated in (9). For a more in-depth discussion on the multifunctionality of the marker *là*, refer to Note 14 in this paper.

    b. sit⁸ si⁷ chhing¹-khi³ (07.000, WL)
    實 是 清氣
    really FM clean
    'It is really clean.' (Lien 2009: 748)

'(Suppose you enter a house and see that things to eat and drink have been prepared and laid out. After you have eaten and drunk and rested, to whom do you give thanks, to the house or to the master of the house, even though you do not see him?) Would it not **be foolish** (Lit. **very foolish**) to give thanks to the house and not to its master?'

4.2. Contemporary usages

4.2.1. Nominal usage

Similar to the Vietnamese language of the17th-century, *sự thật* (classifier + true) is used to mean denote 'truth, fact' as shown in (10).

(10)   *Đó*   *là*   ***sự***   ***thật***   *đấy*[16].
         DEM.MED   LNK   CLF   true   UFP [sharing information]
         (That's true, for your information [(Lit.) 'That's **a type of truth**'].) (Elicitation[17])

4.2.2. Adjectival usage

As an attributive adjective meaning 'true', *thật* follows a noun as shown in (11).

(11)[18] *Có nên*   *sử dụng* ***tên*** ***thật*** *trên Facebook của mình hay không?*
         Q   should   use   name   real   on   PN   of   I   or   NEG
         'Should I use my **real name** on Facebook or not?' (A quote from the Internet)

Especially in a spoken context, *thật* is used as a type of backchanneling[19] with some UFPs such as *thật không?* 'really?' as in (12) or *thật à?* 'really?' as in (13). In both instances, the speakers express surprise at new information or information that contradicts their prior expectations, and demonstrate interest through these *thật*-based backchannelings. In (12), the daughter is astonished because the mother reveals unexpected information about desserts that the daughter was unaware of. In (13), the mother is surprised when the father informs her that the price of the candles they are discussing is cheaper than she had assumed.

---

[16] UFP *đấy* functions as an assertive marker when the speaker thinks that the information is new to the addressee with the meaning 'which you may not know about' or 'for your information' (Adachi 2021: 81–84).

[17] The elicited examples (10), (18), (26), and (28) were provided by the language consultant. Refer to Note 10 of this paper for the original source.

[18] Công ty TNHH Phương Nam VINA. 2014. Có nên sử dụng tên thật của mình trên facebook hay không? [Should I use my real name on Facebook or not?] https://websitechuyennghiep.vn/co-that-su-dung-ten-that-tren-facebook.html (published on 5 August, 2014; accessed on 9 September 2024)

[19] Mereu (2024: 2) defines backchanellings as "the short productions uttered by one participant in the conversation when the other participant occupies the floor."

(12) M: *có lẽ   tráng miệng   là   những   xoài   đấy*
          perhaps   dessert   LNK   PL   mango   UFP [sharing information]
          'Perhaps the dessert was mangos, for your information. '
    D: ***thật không?***
          true UFP [question]
          'Really?' (Family conversation)

(13) F: *hôm nay mình   xem   cái   cây   nến   kia*
          today   I [self][20] see   CLF [thing] CLF [tree-like object] candle   DEM.DIST
          'I saw that candle (in the store) today.'
          một cây   kia   là   hai   trăm   nghìn
          one   CLF   DEM.DIST   LNK   two   hundred   thousand
          'That candle cost 200,000 Vietnamese dollars (= about 10 US dollars).'
    M: *bao nhiêu?*
          how   much
          'How much?'
    F: *hai   trămnghìn*
          two   hundred   thousand
          '200,000   Vietnamese   dollars.'
    M: ***thật à?***
          real UFP [question with mild surprise]
          'Really?' (Family conversation)

4.2.3. Adverbial usage

   (14) is an example of *thật* as an intensifier modifying an attributive adjective
ngon 'delicious'.

(14) [21] *Làm   món   nướng   **thật**   **ngon***
          make   food   grill   THẬT   delicious
          'Make **really delicious** grilled dishes.' (A quote from the Internet)

   In sentences (15), (16), and (17), *thật* intensifies the negation of stative verbs
of emotion that follow it such as *nỡ* 'bear', *chịu* 'bear', and *hiểu* 'understand',
meaning 'really unbearable' or 'really incomprehensive'.

(15)[22] *Giờ   phải   chia   xa, tôi   **thật   không nỡ**   chút   nào.*
          time   have.to separate far   I [servant]   THẬT   NEG   bear   little   any
          'When I have to be far away (from my wife), I cannot bear (it) at all.'
                                                        (A quote from the Internet)

---

[20] Supplemental information may be added within brackets.
[21] Khánh Hòa. 2013. Làm món nướng thật ngon [Make really delicious grilled dishes.], *Vnexpress*
https://vnexpress.net/lam-mon-nuong-that-ngon-2755290.html (published on 16 May, 2013
accsessed on 9 September 2024).
[22] *VOV2*. Tôi không nỡ bán vườn, xa quê [I can't bring myself to sell the garden and be far away
from home]. https://vov2.vov.vn/toi-khong-no-ban-vuon-xa-que-12575.vov   (published on 3
February, 2015; accessed on 9 September 2024).

(16)[23] ***Thật   không chịu   nổi****!*
　　THẬT   NEG   bear   able
　　'**Really unbearable**!' (A quote from the Internet)

(17)[24] *"Tôi   **thật   không hiểu   nổi**"*
　　I [servant]   THẬT   NEG   understand   able
　　'**Really incomprehensive** (for me)!' (A quote from the Internet)

The adverbial *thật* can also modify a predicate adjective such as *lãng mạn* 'romantic' as shown in (18).

(18) *Phim   này   **thật   là**[25]   **lãng mạn***.
　　Movie   DEM.PROX   THẬT   LNK   romantic
　　'This movie is **really romantic**.' (Elicitation)

In addition, the adverbial *thật* can modify proceeding verbs *tưởng* 'think' in (19), *nói* 'say' in (20) and (21) , and *sống* 'live' in (21), meaning 'truly', 'candidly' or 'honestly', respectively. Note that, as shown in (19)–(21), *thật* is not formally treated as an object noun like its English translation, 'truth,' but rather as an adverbial modifying the predicate.

(19)[26] *nói   đùa   mà   cứ   **tưởng   thật***
　　say   lie   but   continue   think   THẬT
　　'(I) lied, but (you) **believed it to be true**.'   (A quote from the Internet)

(20)[27] *Không phải   lúc nào cũng   nên   **nói   thật***
　　NEG   right   whenever   should   say   THẬT
　　'**Talking candidly** isn't always the best choice.' (A quote from the Internet)

(21)[28] A quote from the Internet
　　*Nhà văn phải   là   người   **nói   thật,   sống   thật***
　　writer   have.to LNK   person   say   THẬT   live   THẬT
　　'A writer must be someone who **tells the truth** and **lives honestly**.'

　　　　　　　　　　　　　　　　　　　　　　　　(A quote from the Internet)

---

[23] Đăng Huỳnh. Thật không chịu nổi! [Really unbearable!]. *Cần Thơ online*. https://baocantho.com.vn/that-khong-chiu-noi--a69162.html (published on 15 July, 2012; accsessed on 9 September 2024).

[24] Huỳnh Trung Phong. 2024. Declan Rice: "Tôi thật không hiểu nổi" [Declan Rice: "Really incomprehensive!"]. *Tin thể thao* [Sports news]. https://www.tinthethao.com.vn/declan-rice-toi-that-khong-hieu-noi-d766912.html (published on 24 June, 2024; accsessed on 9 September 2024).

[25] See Note 15 in this paper for a focus marker *là*.

[26] Soha tratu. 2024. Thật [true]. http://tratu.soha.vn/dict/vn_vn/Th%E1%BA%ADt (accessed on 9 September 2024).

[27] Hoài Nam. 2024. Không phải lúc nào cũng nên nói thật [Talking candidly isn't always the best choice.] *Dân Trí* [People's Intellectual] https://dantri.com.vn/nhip-song-tre/khong-phai-luc-nao-cung-nen-noi-that-1206978735.htm (published on 1April, 2008; retrieved on 9 September 2024)

[28] Lam Điền. 2022. Nhà văn nói về nghề: Nhà văn phải là người nói thật, sống thật [A writer must be someone who tells the truth and lives honestly.]. *Tuổi trẻ* [Youth]. https://vanchuongthanhphohochiminh.vn/nha-van-noi-ve-nghe-nha-van-phai-la-nguoi-noi-that-song-that (published on 8 June, 2022; retrieved on 9 September 2024)

4.2.4. Utterance-final particle usage

Sometimes, *thật* appears as an UFP with other UFPs, which as a whole plays an intensifier role. For example, *thật đấy* and *ấy thật* are used utterance-finally in (22)–(26) respectively. In other words, *thật* is not fixed but formulaic (or loosely conventionalized). While *thật* as an adverbial to intensify the objective criterion, such as 'better/worse/more/less than normal', it serves as a UFP to express the speaker's subjective evaluation based on direct experience, such as the speaker's fatigue in (22).

(22) M: *mệt*   **thật**   *đấy*
       tired   THẬT   UFP [sharing information]
       '(I'm) tired, **indeed**, (for your information).' (Family conversation)

In (23) and (24), the speaker (M) discusses the plot of the theatrical play she attended the day before. She evaluates the playwright's and actor's styles based on her actual experience.

(23) M: *kiểu*   *kịch*   *Lưu Quang Vũ*   *ấy*[29]   **thật**
       style   play   PN   UFP [shared information]   THẬT
       '(That was) exactly the playwright Luu Quang Vu's dramatic style, **indeed**.
                                           (Family conversation)

(24) M: *thằng*   *Xuân Bắc, nó*   *đóng*   *buồn cười*   **thật**   *đấy*
       CLF [young man] PN   he play humorous THẬT UFP [sharing information]
       'Mr. Xuan Bac played a very humorous role (in the drama yesterday, **indeed**).
                                           (Family conversation)

(25) is an interjectional use of *thật*. When the speaker says *chết*[30] *thật* (Lit. 'really dying') to express his or her annoyance or disappointment, however, he or she is not dying in reality (cf. (4)).

(25) **Chết**   **thật**!
     die   THẬT
     '**Oh no!**' (Elicitation)

4.2.5. Discourse marker usage

Notably, DM usage shows a more subjective view of the situation in an utterance-initial position as shown below. This discourse characteristic is in contrast to the UFP usage used in the utterance-final position, as in Section 4.2.4. (see Section 5.2 for relevant discussions). Furthermore, *thật* has developed in diverse ways as DMs, some of which are shown below (see also Section 2.2). In examples (26)–(29), *thật không may* 'unfortunately' and *thật là may* 'fortunately'

---

[29] Utterance-final particle *ấy* is used to remind the addressee of something the addressee already knew ('remember?') (Adachi 2021: 90–99).
[30] *Chết* 'die' alone and *chết rồi* '(die + PFT) Lit. 'have been dead' also have interjectional use meaning 'Oh no!'.

indicate the speaker's subjective attitude toward the situation described, rather than presenting objective facts. These terms help to draw the addressee's attention to the speaker's perspective.

(26)[31]

Thật	***không***	***may,***	*dịch vụ*	*Google Play*	*đã*	*ngừng*
THẬT	NEG	fortunate	service	PN	PFT	stop

'**Unfortunately**, Google Play services has stopped'. (A quote from the Internet)

(27)[32]

Thật	***không***	***may***,	*tựa*	*phần*	*mềm*	*này*		*không*	*còn*
THẬT	NEG	fortunate	title	part	soft	DEM.PROX		NEG	remain

trong	*kho.*
within	store

'**Unfortunately**, this title is permanently out of inventory.' (A quote from the Internet)

(28)[33]

Thật	***không***	***may***		*lại*	*có*	*vấn đề.*
THẬT	NEG	fortunate		again	have	problem

'**Unfortunately**, there is a problem again.' (A quote from the Internet)

(29)[34]

Thật	***là***	***may***,	*một*	*người*	*bạn*	*đã*	*cho*	*cậu*
THẬT	LNK	fortunate	one	person	friend	PFT	let	him [uncle, young man]

mượn	*một*	*chiếc*	*máy bay*, …[snip]
borrow	one	CLF [vehicle]	airplane

'**Fortunately**, a friend lent him an airplane, ⋯[snip]' (A quote from the Internet)

In examples (30)–(33), the phrase *(phải nói) thật là*, which translates to 'I had to choose, to be honest' (literally 'I have to tell the truth that'), indicates that the speaker feels the situation being described is challenging to communicate to the addressee. This concern arises from the speaker's worries about how the addressee might react. It is worth noting that the words *phải* 'have to' and *nói* 'say, tell') can be omitted without losing the meaning.

---

[31] Công ty SONY Electronics Việt Nam. 2024. Thông báo sau xuất hiện thường xuyên: Thật không may, dịch vụ Google Play đã ngừng. [The following message appears frequently: Unfortunately, Google Play services has stopped]. https://www.sony.com.vn/electronics/support/articles/00141861 (published on 1 October, 2016; accessed on 9 September 2024).

[32] Intel Software Advantage Program Support Center. 2024. Thật không may, tựa phần mềm này không còn trong kho [Unfortunately, this title is permanently out of inventory]. https://tgahelp.zendesk.com/hc/vi/articles/13531897639693-Th%E1%BA%ADt-kh%C3%B4ng-may-t%E1%BB%B1a-ph%E1%BA%A7n-m%E1%BB%81m-n%C3%A0y-kh%C3%B4ng-c%C3%B2n-trong-kho (published on 8 March, 2024; accessed on 9 September 2024).

[33] Glosb. 2024. Thật không may [Unfortunately]. https://vi.glosbe.com/vi/en/th%E1%BA%ADt%20kh%C3%B4ng%20may (Accessed on 9 September 2024).

[34] Remy Charlip 2014. Thật Là May – Fortunately. Fahasa.com. https://www.fahasa.com/that-la-may-fortunately.html?srsltid=AfmBOoqbqUesqUr_U7MGVeEN-Y8heeft6TITv6eyUDu-Lq5PfuXTwvDP (Accessed on 9 September 2024).

(30)[35]   ***Phải***   ***nói***   ***thật***   ***là***   *chất lượng*   *xây dựng*   *các*   *tòa nhà*   *tại*
          have.to   say   THẬT   LNK   quality   construction   PL   building   in
          Việt Nam   *quá*   *kém*.
          Vietnam   too   low.
          '**To be honest**, the quality of the building construction in Vietnam is too low.'
                                                                                (A quote from the Internet)

(31)[36]   *"**Nói***   ***thật***   ***là***   *rất*   *khó"*
          say   THẬT   LNK   INT   difficult
          '**To be honest**, (it's) very difficult.' (A quote from the Internet)

(32)[37]   A quote from the Internet
          Nói   ***thật***   ***là***   *không*   *dám*   *ra*   *khỏi*   *nhà*   *luôn,*   [snip]
          say   THẬT   LNK   NEG   dare   go.out   escape   house   immediately
          '**To be honest**, (I) dared to leave (my) house right away.'
                                                                                (A quote from the Internet)

(33)   ***Thật***   ***là***   *anh*[38]   *nhớ*   *em*   *nhưng anh*
          THẬT   LNK   I [brother]   miss   you [sister]   but   I [brother]
          không   *dám*   *nói*   *ra.*
          NEG   dare   say   outwards
          '**To be honest,** I miss you but I cannot tell you.'   (Elicitation)

In (34) and (35), *thật ra* (thật + ra) means 'actually' and indicates that the speaker presents a situation as a fact that contradicts the addressee's expectations.

---

[35] Nhất Nam. 2017. Chuyên gia Nhật: "Phải nói thật là chất lượng xây dựng các tòa nhà tại Việt Nam quá kém." [A Japanese expert: "To be honest, the quality of the building construction in Vietnam is too low."] *Đầu tư Bất động sản* [Real Estate Investment]. https://baodautu.vn/batdongsan/chuyen-gia-nhat-phai-noi-that-la-chat-luong-xay-dung-cac-toa-nha-tai-viet-nam-qua-kem-d59560.html (published on 1 March, 2017; Retrieved on 9 September 2024)

[36] Hà Vũ 2024. Liên kết phát triển miền Trung: "Nói thật là rất khó" [Central region development linkage: "To be honest, it is very difficult"] *VnEconomy*. https://vneconomy.vn/lien-ket-phat-trien-mien-trung-noi-that-la-rat-kho.htm (Accessed on 9 September 2024)

[37] Đại sứ quán Cộng hòa xã hội chủ nghĩa Việt Nam tại Matxcova - Liên bang Nga [The embassy of Socialist Republic of Viet Nam in the Russian Federation]. 2022. Nỗi niềm người Việt sơ tán từ Ucraina [Concerns of Vietnamese evacuees from Ukraine] https://vnembassy-moscow.mofa.gov.vn/vi-vn/News/EmbassyNews/Trang/N%E1%BB%97i-ni%E1%BB%81m-ng%C6%B0%E1%BB%9Di-Vi%E1%BB%87t-s%C6%A1-t%C3%A1n-t%E1%BB%AB-Ucraina.aspx (published on 14 March, 2022; Accessed on 9 September 2024)

[38] In Vietnamese, kinship terms often function as personal pronouns. both among family members and in broader social contexts. These terms vary according to according to gender, age, familiarity, and social relationships. For example, terms like *anh* 'elder brother', *chị* 'elder sister' and *em* 'younger sibling' indicate the power dynamics, intimacy and solidarity between participants in a conversation. Nguyễn (1997: 129) explains that couples, referred to in Vietnamese as *vợ chồng* 'wife and husband', use the terms *anh* 'elder brother' and *em* 'younger sister' to signify their relationship. In this context, *anh* implies 'I' (when the boy or husband is speaking) and *em* implies 'you' (when the girl or wife is speaking), or vice versa.

(34)[39] ***Thật   ra**, không   phải   lúc nào   cố gắng   cũng   là   tốt*
THẬT RA NEG   right   whenever exert.effort   also   LNK   good
'**Actually**, it's not always good to exert effort.' (A quote from the Internet)

(35)[40] ***Thật   ra   rất   đơn giản.***
THẬT RA   INT   simple
'**Actually**, it's really simple.'   (A quote from the Internet)

In (36), *quả thật* (fruit + *thật*) meaning 'in fact' appears in the utterance-initial position, indicating that the speaker perceives the situation as a fact, similar to *thật ra* 'actually' in (34) and (35).

(36)[41] ***Quả thật**,   tôi   không biết   nên   giải quyết*
QUẢ THẬT   I [servant]   NEG   know   should solve
vấn đề   này   như thế nào   và   bắt đầu   từ   đâu?
problem DEM.PROX like   how   and   begin   from   where
'**In fact**, I don't know how (I) should solve this problem and where (should I) start?'
(A quote from the Internet)

What I have shown above can be summarized as in Table 1. First, the nominal usage of *thật* was obsolete, if not possible, in the 17th century. Second, while adjectival and adverbial usages are attested in the 17th century, the DM usage is found only in contemporary Vietnamese. Lastly, the newly derived DM usage is realized in a variety of ways, which I indicates that *thật* has been well grammaticalized in the language.

Table 1. ***Thật**-based expressions in the 17th and 20th centuries*

	17th century	**20th century**
(A) Nominal usage	obsolete	obsolete
(B) Adjectival usage	*Chúa **thật*** 'the true lord' Ex. (5) *rứt **thạt*** 'very sincerely' Ex. (6)	*tên **thật*** 'real name' Ex. (11) ***thật** không?* 'really' Ex. (12)
(C) Adverbial usage	***thật** là* 'without doubt' Ex. (7) *chết **thạt*** 'really die' Ex. (8) ***thạt** là dại* 'very foolish' Ex. (9)	*món nướng **thật** ngon* 'really delicious grilled dishes' Ex. (14) ***thật** là lãng mạn* 'really romantic' Ex . (18) *nói **thật*** 'tell the truth' Ex. (20), (21)
(D) UFP usage	n.a.	*mệt **thật*** 'tired indeed' Ex. (23)

[39] Chùa Hạnh Đức. 2023. Thật ra, không phải lúc nào cố gắng cũng là tốt [Actually, it's not always good to exert effort] https://chuahanhduc.com/bai-viet/5-nep-song-dao/18550-that-ra-khong-phai-luc-nao-co-gang-cung-la-tot (Published on 4 December, 2023; Retrieved on 9 September 2024).
[40] Glosb. 2024. Thật ra [Actually]. https://vi.glosbe.com/vi/en/th%E1%BA%ADt%20ra (accessed on 9 September 2024).
[41] VietJack khoa hoc. 2024. 'Câu hỏi [Question]' https://khoahoc.vietjack.com/question/464607/ qua-that-toi-khong-biet-nen-giai-quyet-van-de-bay-nhu-the-nao-va (published on 14 July, 2024; accessed on 9 September 2024).

	17th century	20th century
(E) DM usage	n.a.	***thật*** *không may* 'unfortunately' Ex. (26), (27), (28) ***nói thật là*** 'to be honest' Ex. (31), (32) ***thật là*** 'to be honest' Ex. (33) ***thật ra*** 'actually' Ex. (34) (35) ***quả thật*** 'in fact' Ex. (36)

5. Discussion

5.1 From "TRUE" to "INTENSIVE" in Sino-Vietnamese

The above observations suggest the following. First, the historical pathway of Sino-Vietnamese 實 *thật* seems to be in line with the well-attested grammaticalization path from "TRUE" to "INTENSIVE" (Kuteva et al. 2019: 443); from the adjectival usage 'real' to the adverbial usage 'really, indeed', subsequently giving rise to DMs in contemporary Vietnamese. In a broad sense, Hansen's (2018a, b) cyclicity hypothesis holds true for Sino-Vietnamese cases. Interestingly, adjectival and adverbial usages can be attested in the 17th-century text PGTN. When the Chinese character 實 for *thật* was borrowed into Vietnamese (presumably as a noun), its adjectival usage appeared first, followed by its adverbial usage, because the reverse is not plausible from the standpoint of language change. However, of course, this is mostly a matter of speculation owing to the lack of any reliable number of historical resources in the language,[42] although it is based on some cross-linguistic evidence (Kuteva et al. 2019: 443).

5.2. Toward (inter)subjectification

Second, the newly attested DMs of *thật* origin in contemporary Vietnamese are likely to convey the speaker's view of the content of the immediate proposition, that is, subjectification. Moreover, and some of them seem to play an interpersonal role of prompting a reaction from the interlocutors, that is, intersubjectification. Consider the following examples. Note that *ấy* implies shared knowledge.

(23)'
    M: *kiểu*   *kịch*   *Lưu Quang Vũ*   *ấy*   ***thật***
            style   play   PN   UFP   THẬT
         '(That was) exactly the playwright Lưu Quang Vu's dramatic style, indeed.'
                                               (Family conversation).

---

[42] Alves (2007b: 219) contends that most Vietnamese intensifiers of Chinese origin were not originally grammatical forms in Chinese, but rather became grammatical after being incorporated into Vietnamese. This supports the argument presented in this paper.

(31)'   A quote from the Internet
    "***Nói   thật   là   rất   khó***"
    say   THẬT LNK   INT   difficult
    '**To be honest**, (it's) very difficult.'

When *thật* is used as an adverbial intensifier, it appears typically in the utterance-final position as in (26).[43] The usage shows that what is being said exceeds a certain standard according to the speaker's judgment: This usage of *thật* can be regarded as a case of subjectification.[44] Contrasting ly, when *thật* is used as a DM, it always appears in the utterance-initial position as in (36).[45] These DM usages serve to introduce the speaker's attitude toward the following content, which is however only partly true. Interestingly, such utterance-initial DMs appear to play a vital role in attracting the attention of potential interlocutors, thus constituting leading to a case of intersubjectification.[46] The fact that some 'truth'-related expressions in Vietnamese have taken on an intersubjective role in discourse leads us to a greater understanding of the way borrowed words of Chinese origin develop into DMs in each language (see Khammee 2024, inter alia).

Further, the functional differentiation of subjective and intersubjective *thật*-based DMs in terms of discourse-syntactic positions deserves attention. Arguably, Beeching and Detges (2014: 11) propose the following hypothesis: linguistic items such as adverbs and DMs occurring before the proposition serve to express something subjective, whereas those occurring after the proposition serve to express something intersubjective. It is true that certain items in some languages provide evidence in favor of Beeching and Detges's (2014) hypothesis (see papers therein). However, Vietnamese DMs, at least those presented in this study, provide evidence against the hypothesis (see also papers in Rhee et al. 2021).

5.3. Morphosyntactic expansion in grammaticalization

Third, the formal feature of this grammaticalization is worth reconsideration. When it comes to DMs and relevant expressions such as general extenders (GEs) in the history of English, Overstreet and Yule (2021: 103) suggest that "there is a pattern of change from a longer form to a shorter form, retaining only the first two constituents of the longer expression. As we discovered, this is a typical direction of change throughout the history of English." While Overstreet and Yule (2021:

---

[43] Of course, it is possible to take this usage as a DM instead of as an adverbial intensifier. However, differentiating from DMs and adverbial intensifiers is a matter I intend to pursue in future research.
[44] See Rhee (2021: 342) fors discussion of mirativity based on 事實上 (*sasilsang*).
[45] As presented in Section 2.2, *thật* is not just used as a single component but often as a part of formulaic sequences such as *nói thật là* 'to be honest'. Note that this particular formulaic expression may have been conventionalized enough to be reduced to *thật là* 'to be honest'. Nevertheless, this does not necessarily mean the prevalence of 'erosion' among Vietnamese DMs (see Section 5.3 for this issue).
[46] This attention-getting function of *(nói) thật là* 'to be honest' may be termed 'projector'. See Hopper & Thompson (2008) and Shibasaki (2014) on projectors.

175–176) summarize some of the structural differences of GEs in several languages,[47] the above-mentioned pathway of change seems generally plausible.[48] In grammaticalization studies, it is called 'erosion,' that is, "loss of phonetic features that a linguistic expression experiences in the course of grammaticalization" (Narrog & Heine 2021: 334, see Givón 1979: 208–209 for a typical reduction scale).

Contrarily, the Vietnamese DMs summarized in (2) show formal expansion instead of formal reduction, i.e., *thật* 實 'real' + *ra* 'out' in (2a), *thực* 寔 'real' + *ra* 'out' in (2b), *quả* 果 'fruit' + *thực* 寔 'real' in (2j), *kì* 其 'its/that' + *thật* 實 'real' in (2n) as well as *thật* 實 'real' + *là* (linker), discussed in this study. Considering the fact that *thật* 實 and its cognates in Japanese, Korean, and Thai are all monosyllabic-like unless particles are attached (e.g., *shinni* [*shin* 'truth' + *-ni* (adverbializer)] 'truly'), it would be natural for a short form to recruit other elements to make different types of DMs over time. In fact, we can see similar patterns of change in Thai (Khammee 2024) and in Japanese (Higashiizumi et al. 2024). Therefore, this study raises the possibility that some languages go through morphosyntactic expansion or modification instead of erosion to produce a cluster of related DMs.

5.4. Borrowed as a lexeme and grammaticalized as a DM[49]

Last, it would be useful to suggest how DMs are derived in Asian languages, an issue especially addressed in this special issue. The borrowability of DMs through language contact has been pursued mostly from a synchronic perspective, notably in the field of bilingualism (e.g., Maschler 2000, Goss & Salmons 2000). Some studies regard DMs (or prefabricated multiword units) as among the most frequently borrowed items (e.g., Muysken 1981, Grant 2012). Suppose that DMs are easily borrowed from one language to another; 'truth'-related DMs of Chinese origin in Japanese, Korean, Thai, and Vietnamese might have been borrowed through either intensive or extensive contact with Chinese at various points in history. Contrary to expectations, however, this is quite unlikely as reported in this special issue (see also Rhee et al. 2021, Higashiizumi & Shibasaki (forthcoming), inter alia). In these languages, certain lexical items were borrowed from Chinese as is, and then, they developed in language-specific (but similar) ways into DMs.[50]

How, then, can we explain the unborrowable nature of DMs in Asian languages spoken in the region of Chinese characters? It is well known that DMs are typically developed and used in spoken discourse. Therefore, one can easily find studies of

---

[47] The languages examined therein are English Creoles (Trinidad and Hawai'i), Brazilian Portuguese and Spanish, Lithuanian, Russian, German, French, Swedish, and Persian. Note that Overstreet and Yule (2021) make no in-depth survey of any of the Asian languages addressed in the special issue.

[48] In addition to Overstreet and Yule (2021), the *it/there is no X* construction with a variety of nouns such as *doubt, question, way, wonder,* etc. in the slot X underwent clausal reduction to *no X* as DMs (e.g., Davidse & de Wolf 2012, Gentens et al. 2016, Nykiel & Shibasaki, forthcoming).

[49] This subsection is developed from Shibasaki and Higashiizumi (forthcoming).

[50] Matras (2007: 57) states that there is no clear evidence of borrowing of DMs in Vietnamese.

borrowing through spoken contact (e.g., Clyne 2003: 225–232 on the use of *well* and *you know* in bilingual and trilingual contexts, that is, English, German, Dutch; see also García Vizcaíno and Martínez-Cabeza (2005) for *well* in English and *bueno* in Spanish). However, such heavy lexical borrowing of Chinese words into Japanese, Korean, Thai, and Vietnamese became possible only through written contact (see Note 1). Even contemporary DMs, if they can be traced to certain borrowed lexemes form Chinese, are likely to have developed after borrowing in their own ways. Moreover, such written contact-based lexemes are now used in spoken contexts in Vietnamese, as shown in Section 4. As this is a research area to be exploited, I plan to do follow-up research on it, based on both information and inputs from Rhee et al. (2021), Higashiizumi and Shibasaki (forthcoming), and this special issue.

6. Conclusion remarks

In this study, I presented a variety of Vietnamese *thật*-based DMs, by analyzing data from the 17th-century and contemporary Vietnam. The following points were derived from the analysis. *Thật*-related expressions seem to have followed the well-attested path from "TRUE" to "INTENSIVE" (Kuteva et al. 2019: 443) and further into DMs (Section 5.1). The subjective and intersubjective usages of *thật* forms are distributed in a complementary manner: The subjective usage appears in utterance-final position, while the intersubjective usage appears at utterance-initial position. Theoretically as well as descriptively, this finding suggests that researchers to reconsider Beeching & Detges's (2014) hypothesis (Section 5.2). *Thật* has also recruited other elements to create multiword-unit DMs, which implies that DMs do not always undergo formal reduction (Section 5.3). As to whether DMs are borrowed through language contact, examples from Vietnamese and from Japanese (Higashiizumi & Shibasaki forthcoming) and Thai (Khammee 2024) suggest that lexical elements are first borrowed as they are and then grammaticalized as DMs in language-specific but similar ways (Section 5.4).

These findings are theoretical contributions to gain a better understanding of the way DMs emerged and developed in Vietnamese, through a descriptive analysis of historical and contemporary data. However, a comparison of data from the 17th century and from present-day Vietnamese is neither balanced nor sufficient: there is a need to examined the 18th- to 19th-century texts[51] written in Chinese scripts, *Chữ nôm* and *Quốc ngữ* (See 2.1) as well. Languages do not change in a uniform way at a uniform rate. Therefore, I will conduct a follow-up review and research in to further test the results of the study.

---

[51] According to Alves (2005: 320), based on data from *Truyện Kiều*, a poem written in *chữ nôm* characters in the 19th century, the adverb *thật* primarily functions before main verbs rather than adjectives. This observation suggests that DM usage such as *thật ra* 'actually' may occur in the late 19th century or possibly after.

Acknowledgments

Abbreviations

CLF	classifier
COMP	complementizer
DEM.DIST	distal demonstrative
DEM.MED	medial demonstrative
DEM.PROX	proximal demonstrative
INT	intensifier
LNK	linker
NEG	negation
PFT	perfect
PL	plural
PN	proper noun
POSSIB	possibility
Q	question
UFP	utterance-final particle

Quoted data

de Rhodes, Alexandre. 1651. *Phép giảng tám ngày* [Catechisms in eight days] https://vi.wikisource.org/wiki/Ph%C3%A9p_gi%E1%BA%A3ng_t%C3%A1m_ng%C3%A0y

de Rhodes, Alexandre. 1651. *Từ điển Việt–Bồ–La* [The dictionary of Annamese, Lucitanian and Latin] https://vi.wikisource.org/wiki/T%E1%BB%AB_%C4%91i%E1%BB%83n_Vi%E1%BB%87t%E2%80%93B%E1%BB%93%E2%80%93La/

Phan, Peter C. 2015. *Mission and Catechesis: Alexandre de Rhodes & Inculturation in Seventeenth-Century Vietnam.* Maryknoll/New York: Orbis Books.

References

Adachi, Mayumi. 2021. *Betonamugo kuukan daikushisu to sono tenkai: shijishi kara bunmatsushi, kandoushi he [Spatial Deixis and its Development in Vietnamese: From Demonstratives to Sentence-final Particles and Interjections*]. Tokyo: Bensei Publishing.

Alves, Mark J. 2001. What's so Chinese about Vietnamese? In Graham W. Thurgood (ed.), *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*, 221–242. Tempe, AZ: Arizona State University, Program for Southeast Asian Studies.

Alves, Mark J. 2005. Sino-Vietnamese grammatical vocabulary and triggers for grammaticalization. In Vietnam National University, Hanoi. (ed.), *Proceedings of the 6th Pan–Asiatic International Symposium on Linguistics*, 315–332. Hanoi: Nhà xuất bản Khoa học Xã hội [Social Sciences Publishing House].

Alves, Mark J. 2007a. Sino-Vietnamese grammatical borrowing: An overview. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 343–361. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110199192.343

Alves, Mark J. 2007b. Categories of grammatical Sino-Vietnamese vocabulary. *The Mon-Khmer Studies Journal* 37. 217–237.

Alves, Mark J. 2009. Loanwords in Vietnamese. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 617–637. Berlin: Mouton de Gruyter.

Beeching, Kate & Ulrich Detges. 2014. Introduction. In Kate Beeching & Ulrich Detges (eds.), *Discourse Functions at the left and right periphery: Crosslinguistic investigations of language use and language change*, 1–24. Leiden: Brill. https://doi.org/10.1163/9789004274822_002

Brinton, Laurel J. 2010. Discourse markers. In Andreas H. Jucker & Irma Taavitsainen (eds.), *Historical pragmatics*, 285–314. Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110214284.5.285

Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316416013

Bui, Thi Hoang Anh. 2012. *Traduire les marqueurs discursifs (MD) du vietnamien en français. Doctoriales en Sciences du Langage (DoSciLa) 2012-Linguistique et métiers de la traduction, Mar 2012, Paris, France.*

Bui, Thi Hoang Anh. 2015. Etude des marqueurs discursifs du vietnamien dans une perspective comparative avec les marqueurs discursifs du français. Thèse de l'Université Paris Diderot.

Chèn, Ching-ho. 1982. Dainanjisturoku to genchoshuhon ni tsuite [On Đại Nam thực lục and Châu bản triều Nguyễn] In Ine, fune, Matsuri kankou sewanin [Rice, boats, festivals publishing liaisons] (ed.), *Ine, fune, Mtsuri: Matsumoto Nobuhiro sensei tsuitou ronbunshuu* [Rice, boats, festivals: Essays marking the death of Professor Matsumoto Nobuhiro], Tokyo: Rokkou shuppan.

Clyne, Michael. 2003. *Dynamics of Language Contact: English and Immigrant Language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511606526

Davidse, Kristin & Simon De Wolf. 2012. Lexicalization and grammaticalization: The development of idioms and grammaticalized expressions with *no question*. *Text & Talk* 32. 569–591. https://doi.org/10.1515/text-2012-0027

Do-Hurinville, Danh-Thành & Huy-Linh Dao. 2019. The Vietnamese polyfunctional marker *mà* as a Generalized linker: A multilevel approach. *Journal of the Southeeast Asian Linguistcs Society* 12 (2). 58–71.

Edmondson, Jerold A. 2006. Vietnamese. In Keith Brown (ed.), *Encyclopedia of language and linguistics* (Second edition). 432–436. Amsterdam/Tokyo: Elsevier.

García Vizcaíno, María José & Miguel A. Martínez-Cabeza. 2005. The pragmatics of *well* and *bueno* in English and Spanish. *Intercultural Pragmatics* 2 (1). 69–92. https://doi.org/10.1515/iprg.2005.2.1.69

Gentens, Caroline, Ditte Kimps, Kristin Davidse, Gilles Jacobs, An Van linden & Lieselotte Brems 2016. Mirativity and rhetorical structure: The development and prosody of disjunct and anaphoric adverbials with *'no' wonder*. In Gunther Kaltenböck, Evelien Keizer & Arne Lohmann (eds.), *Outside the clause*, 125–56. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.178.05gen

Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press. (Revised edition, Amsterdam: John Benjamins, 2018)

Goss, Emily L. & Joseph C. Salmons. 2000. The evolution of a bilingual discourse marking system: Modal particles and English markers in German-American dialects. *International Journal of Bilingualism* 4 (4). 469–84. https://doi.org/10.1177/13670069000040040501

Grant, Anthony P. 2012. Contact, convergence, and conjunctions. In Claudine Chamoreau & Isabelle Léglise (eds.), *Dynamics of contact-induced language change*, 311–358. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110271430.311

Hansen, Maj-Britt Mosegaard. 2018a. Cyclic phenomena in the evolution of pragmatic markers: Examples from Romance. In Salvador Pons Bordería & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers: New issues in the study of language change*, 51–77. Leiden: Brill. https://doi.org/10.1163/9789004375420_004

Hansen, Maj-Britt Mosegaard. 2018b. The role of inferencing in semantic/pragmatic cyclicity: The case of Lain *nunc* and French *or/maintenant*. *Open Linguistics* 4. 127–146. https://doi.org/10.1515/opli-2018-0007

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108982856

Higashiizumi, Yuko, Reijirou Shibasaki & Keiko Takahashi. 2024. From truth to truly: The case of *shinni* 'truly' in Japanese. *Russian Journal of Linguistics* 28(4). 843–864. https://doi.org/10.22363/2687-0088-40518

Higashiizumi, Yuko & Reijirou Shibasaki (eds.). (forthcoming) *The Emergence of Pragmatic Markers in East Asian Languages*. Leiden: Brill.

Hopper, Paul J. & Sandra A. Thompson. 2008. Projectability and clause combining in interaction. In Ritva Laury (ed.), *Cross-linguistic studies of clause combining*, 99–123. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.80.06hop

Hopper, Paul J. & Elizabeth C. Traugott 1993. *Grammaticalization* (1st edition). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Khammee, Kultida. 2024. From objective to subjective and to intersubjective functions: The case of the Thai 'truth'-lexeme. *Russian Journal of Linguistics* 28 (4). 942–965. https://doi.org/10.22363/2687-0088-40496

Kishimoto, Emi. 2018. Kirishitan gogakusho no tenkai: Joan rodorigesu to arekusandoru do roodo [The development of Christian language textbooks: João Rodrigues and Alexandre de Lord], *Gobun [Language]* 110. 52–66.

Kishimoto, Emi & Sirai Jun. 2022. *Kirishitan gogaku nyuumon [Introduction to Missionary Linguistcs]*. Tokyo: Yagi Shoten.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World Lexicon of Grammaticalization* (2nd edition). Cambridge & New York: Cambridge University Press. https://doi.org/10.1017/9781316479704

Lien, Chinfa. 2009. The Focus marker si^7 是 and lexicalization of si^7 mih^8 是乜 into what wh-words in earlier Southern Min texts. *Language and Linguistics* 10 (4). 745–764.

Lyons, John. 1982. Deixis and subjectivity: Loqior, ergo sum? In Robert J. Jarvella & Wolfgang Klein (eds.), *Speech, place and action*, 101–123. Wiley: Chichester.

Maschler, Yael. 2000. What can bilingual conversation tell us about discourse markers?: Introduction. *International Journal of Bilingualism* 4 (4). 437–45. https://doi.org/10.1177/13670069000040040101

Matras, Yaron. 2007. The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–73. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110199192.31

Mereu, Daniela, Francesco Cangemi & Martine Grice. 2024. Backchannels are not always very short utterances. The case of Italian Multi-Unit Backchannels. *Journal of Pragmatics* 228. 1–16. https://doi.org/10.1016/j.pragma.2024.05.003

Miyajima, Tatsuo. 2019. *Gengoshi no keiryōteki kenkyū* [*Quantitative Study of Language History*]. Tokyo: Kasamashoin.

Murakami, Yutaro & Akio Imai. 2010. Gendai Betonamugo ni okeru Kan-etsugo no kenkyuu (1) Betonamu heno wasei kango no denpa Joukyou [A study on Chinese vocabularies in Vietnamese (1) The reception of Chinese vocabularies from modern Japanese in Vietnamese Southeast Asian studies]. *Toukyou Gaidai Tounan Ajiagaku* [*Southeast Asian studies, Tokyo University of Foreign Studies*] 15. 19–32.

Muysken, Pieter. 1981. Halfway between Quechua and Spanish: The case for relexification. In Arnold Highfield & Albert Valdman (eds.), *Historicity and variation in creole studies*, 52–78. Ann Arbor, MI: Karoma.

Narrog, Heiko & Bernd Heine. 2021. *Grammaticalization*. Oxford: Oxford University Press.

Nguyễn, Đình Hoà. 1961. The Vietnamese Language. *Vietnam Culture Series 2*. Saigon: Department of National Education.

Nguyễn, Đình-Hòa 1997. Vietnamese: *Tiếng Việt không son phấn* [Vietnamese without veneer]. *London Oriental and African Language Library 9*. Amsterdam/Philadelphia: John Benjamins.

Nykiel, Jerzy N. & Reijirou Shibasaki. (forthcoming) On the rise and fall of *(it/there is) no nay* in the history of English. *Paper originally presented at the seventh Colloque Bisannuel de la Diachronie de l'Anglais (CBDA-7) online, January 27, 2023.*

Overstreet, Maryann & George Yule. 2021. *General Extenders: The Forms and Functions of a New Linguistic Category*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108938655

Rhee, Seongha. 2021. From truth to reality to effect: The journey of *sasilsang* in Korean. Grammaticalization of discourse markers in East Asian Languages. *East Asian Pragmatics* 6 (3). 331–353. https://doi.org/10.1558/eap.21135

Rhee, Seongha, Reijirou Shibasaki & Xinren Chen 2021. Grammaticalization of discourse markers in East Asian Languages. *East Asian Pragmatics* 6 (3). [Special issue]. 271–280.

Rhee, Seongha & Lin Zhang. 2024. The way of the truth: The case of Korean DM *cincca* in comparison with Chinese DMs *zhenshi* and *zhende*. *Russian Journal of Linguistics* 28 (4). 818–842. https://doi.org/10.22363/2687-0088-40500

Schreiter, Robert J. 2015. Foreword. In Peter. C. Phan (ed.), *Mission and catechesis: Alexandre de Rhodes & inculturation in seventeenth-century Vietnam*, 7–9. Maryknoll, New York: Orbis Books.

Shibasaki, Reiijrou. 2014. On the grammaticalization of *the thing is* and related issues in the history of American English. In Michael Adams, Laurel J. Brinton & Robert D. Fulk (eds.), *Studies in the history of the English language VI*, 100–121. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110345957.99

Shibasaki, Reijirou & Yuko Higashiizumi. (forthcoming) An introduction. In Yuko Higashiizumi & Reijirou Shibasaki (eds.), *The emergence of pragmatic markers in East Asian languages*. Leiden: Brill.

Shimizu, Masaaki. 2024. A philological study of the Vietnamese grammatical marker *đang*. *Taiwan Journal of Linguistics* 22 (1). 47–65. https://doi.org/10.6519/TJL.202401_22(1).0002

Tomita, Kenji. 1988. Vietnamugo. In Takashi Kamei, Rokurou Kouno & Eiichi Chino (eds.), *The Sanseido encyclopedia of linguistics*, *Vol. 1, Languages of the world, Part one*, 759–787. Tokyo: Sanseido.

Tomita, Kenji. 2000. *Vetonamugo no sekai: Vetonamugo kihon bunten* [*The World of Vietnamese: Vietnamese Basic Grammar Book*]. Tokyo: Daigakusyorin.

Traugott, Elizabeth Closs. 2003. From subjectification to intersubjectification. In Raymond Hickey (ed.), *Motives for language change*, 124–140. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486937.009

Traugott, Elizabeth Closs. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Davidse Kristin, Lieven Vandelotte & Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization*, 29–71. Berlin: Mouteon de Gruyter. https://doi.org/10.1515/9783110226102.1.29

Zwartjes, Otto. 2018. *Missionary Grammars, Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.384 (Accessed on 9 September 9 2024)

Dictionaries

Kawamoto, Kunie. 2011. *Shoukai betonamugo jiten* [The detailed Vietnamese dictionary]. Tokyo: Taishukan Publishing.

Phạm, Văn Hảo, Lê Văn Trường, Nguyễn Tài Thái, Trần Thị Liên Minh & Võ Xuân Quế. (eds). 2009. *Từ điển phương ngữ tiếng Việt* [Vietnamese dialect dictionary]. Hà Nội: Nhà xuất bản Khóa học Xã hội [Social Science Publishing House].

Thiều Chửu. 2009. *Hán Việt Từ điển* [The Han-Vietnamese dictionary]. Hà Nôi: Nhà xuất bản Văn Hóa - Thông Tin [Culture-Information Publishing House].

Vũ Văn Kính. 2005. *Đại tự điển chữ Nôm* [Comprehensive dictionary of chu Nom]. Thành phố Hồ Chí Minh: Nhà xuất bản Văn nghệ Thành phố Hồ Chí Minh [Ho Chi Minh City Literature Publisher]; Trung tâm nghiên cứu Quốc Học [Center for Vietnamese Classical Literature].

Bionote:
Mayumi ADACHI is an Associate Professor at the Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. Her research interest covers Vietnamese pragmatics, specifically, demonstratives, in addition to demonstrative-derived sentence-final particles and interjections. In recognition of her research on spatial and discourse deixis in Vietnamese and the sociolinguistic study of the Vietnamese community in Japan, she was awarded the Shinmura Izuru Research Encouragement Prize FY2022.
e-mail: adachi@aa.tufs.ac.jp
https://orcid.org/0009-0007-9174-3932

Сведения об авторе:
Маюми АДАТИ – доцент Научно-исследовательского института языков и культур Азии и Африки Токийского университета иностранных языков. Ее исследовательские интересы включают вьетнамскую прагматику, в частности, указательные местоимения, а также производные от указательных местоимений частицы и междометия. В знак признания ее исследований пространственного и дискурсивного дейксиса во вьетнамском языке и социолингвистического изучения вьетнамской общины в Японии она была награждена премией Shinmura Izuru Research Encouragement Prize FY2022.
e-mail: adachi@aa.tufs.ac.jp
https://orcid.org/0009-0007-9174-3932

Review of Elizabeth C. Traugott. 2022. *Discourse Structuring Markers in English*. Amsterdam: John Benjamins, xviii+274 pp. ISBN 978-90-272-1091-3

Foong Ha YAP[1] ✉ and Mikyung AHN[2]

[1]Chinese University of Hong Kong, Shenzhen, China
[2]Hankuk University of Foreign Studies, Seoul, Republic of Korea
✉foonghayap@cuhk.edu.cn

Discourse Structuring Markers in English (Traugott 2022) is about the diachrony of pragmatic markers, in particular a subtype which Elizabeth Traugott identifies with the syntactic form often referred to as 'Conjuncts'. These connectives are derived from circumstantial adverbials (CircAdvs) and some are further recruited as discourse structuring markers (DSMs). What makes this monograph a valuable read is its attempt to tackle some thorny questions about language change. Among these tough questions are: (1) Is pragmaticization a separate phenomenon from grammaticalization? (2) If pragmatic markers emerge as a result of reanalysis (or 'neoanalysis'), is the process gradual or instantaneous? (3) If new functional categories also emerge via analogy, what kind of language change model can insightfully capture this type of process? In trying to address these challenging questions, Traugott turns to Construction Grammar, offers a diachronic perspective to how new functional categories are constructionalized, and suggests how a network model can account for the interlocking patterns in language change.

The monograph starts off with an Introduction that provides an overview of the book, including the goals and methodology, and in particular clearly delimiting the type of pragmatic markers being analyzed to Conjuncts derived from CircAdvs. The choice of terminology appears to be deliberate. A Conjunct, as Traugott notes, can be more than just a connective, since it also often expresses the speaker's stance in the way it conjoins discourse segments. For example, conclusive adverbial *after*

all has over time also come to be used as a contrastive connective (e.g. *I can't imagine anyone disliking Nala. After all, who can detest a cat named after a character from The Lion King?*[1]).

Following the Introduction, the monograph is then divided into three major parts. In Part 1, Traugott focuses on foundational issues in language change from a cognitive linguistic and construction grammar perspective. Chapter 2 pays special attention to Goldberg's model of construction grammar. This helps familiarize the reader with basic terms such as *micro-constructions* and *(sub)schemas*, which are used in describing functional categories and which are useful in visualizing—metaphorically—the changes in semantic, syntactic and prosodic properties of each construction as they are put to use in different discourse contexts. Special attention is also given to Croft's construction model, which defines constructions as form-meaning pairings that are, paradoxically but wonderfully, 'holistic yet analyzable'. Traugott explains the beauty of this view of constructions as follows. A construction with the form *all the same* can be used holistically with the meaning 'in exactly the same way', but this form-meaning pairing can also be used to mean 'in contrast' to express disagreement, in which case the form remains the same but the meaning changes, especially in a pragmatic way. As Traugott points out, "If a construction were truly unanalyzable, change could not occur" (p. 27). With both Goldberg's and Croft's models of grammar in tow, Traugott proceeded with her analysis of what happens at the semantics-pragmatics interface, which is crucial to our understanding of how pragmatic markers emerge.

In Chapter 3, Traugott focuses on the tenets that are crucial when we adopt a diachronic approach to Construction Grammar. She makes a distinction between 'usage changes' and 'grammar changes', highlighting that what changes is *the way speakers use a construction* and this then results in shifts in the constructional networks within the language. Traugott also distinguishes between 'innovation' and 'change', recognizing that innovation can begin with the individual (or with the group when teamwork is involved, we think) and the process could be abrupt, while change often takes time before a new functional category becomes conventionalized and entrenched within a speech community. Traugott also maintains that change often involves semantic shifts in the form of metonymic mappings and, as such, the trajectory of a construction from primarily contentful meaning to increasingly pragmatic meaning is a continuous rather than discrete trajectory. For this reason, Traugott argues that pragmaticization is but an extension of grammaticalization, and not a separate phenomenon. This semantic shift from more contentful and monofunctional meaning to more pragmatic and multifunctional meaning relies on more than just metonymic mapping. What also motivates this drift toward increasing pragmaticization is the process of (inter)subjectification, whereby there is greater orientation toward the speaker's or/and the addressee's epistemic and emotive stance. Consistent with the view that

---

[1] Retrieved, with thanks, on August 14, 2024 from https://www.uis.edu/learning-hub/writing-re-sources/handouts/learning-hub/transition-words

language change is gradient rather than abrupt, Traugott maintains that degree of (inter)subjectification is also gradient in nature. Interestingly, despite the emphasis on continuity, Traugott makes a clear distinction between 'constructionalization' and 'constructional changes' (the latter now renamed as 'constructional shifts' in Traugott (2023)). As defined in Traugott (2022), 'constructionalization' refers to the stage when a new functional category is recognized by a speech community, while 'constructional changes' (or 'constructional shifts') refer to additional uses (including pragmatic ones) that do not significantly alter the functional category. For example, the construction *after all* can have more than one pragmatic reading, among them a pre-clausal DM use where the subsequent discourse segment (D2) justifies the claim in the prior discourse segment (D1) (e.g., *I'm not going. After all, who is she to tell me what to do?*) and a post-clausal DM use with the meaning 'in contrast to speaker's expectation' (e.g. *She's not coming, after all*). Worth noting is that a different DM use does not change the DM category of the construction. In other words, differences in pragmatic meaning do not arise from 'constructionalization'; they are the result of 'constructional shifts'. Constructionalization, on the other hand, involves category change, for example, when circumstantial adverbs (CircAdvs) are neoanalyzed as Conjuncts. In principle, constructional shifts can precede as well as follows constructionalization. This, as we will discuss later, has interesting parallels with yet another theoretical model of language change, namely, Discourse Grammar.

Chapter 4 first reviews previous studies on pragmatic markers (PMs), giving readers a glimpse of the various definitions of PMs in extant literature, and clarifying how the terms pragmatic markers (PMs) and discourse markers (DMs) are used in the monograph. In a nutshell, the term pragmatic marker is an umbrella term for a broad category of stance markers, which include discourse structuring markers (DSMs), and the latter in turn include discourse markers (DMs). This chapter then goes on to elaborate how circumstantial adverbs (CircAdvs) that are constructionalized as Conjuncts come to serve as discourse structuring markers (DSMs). Recall that the term 'Conjunct' refers to the syntactic form, while the term 'Discourse Structuring Marker' refers to the pragmatic function(s). Traugott notes that, in the case of Conjunct-DSM form-function pairings, the Conjunct form can remain the same (e.g., *after all*, *by the way*) or become fused (e.g., *morever*) or become phonologically reduced (e.g., *eal swa* 'exactly like this' > *also*), while the DSM function(s) can vary from being largely limited to expressing more contentful, monofunctional meaning (e.g., *further, furthermore, moreover*) to being further recruited to express highly pragmatic, multifunctional meanings (e.g., *after all*, *by the way*). Traugott adopts the term '1DSM' for the more contentful and monofunctional DSMs and reserves the term 'DM' for the more highly pragmatic and multifunctional DSMs. Broadly speaking, DSMs = 1DSMs + DMs, with the distinction between the two subcategories being gradient rather than discrete. In other words, the DSM category comprises a wide array of Conjuncts with pragmatic functions that vary along a cline from monofunctional 1DSMs to multifunctional

DMs. To further illustrate from a diachronic perspective, when a circumstantial adverb (CircAdv) such as temporal adverbial *after all (this)* is neoanalyzed (hence 'constructionalized') as conclusive adverbial *after all* (Conjunct), it begins to serve not only connectively (textuality) but also affectively (pragmatically) as a discourse structuring marker (DSM), since the speaker's (inter)subjective stance is manifested in the way the speaker chooses to conjoin discourse segments D1 and D2. Some DSMs (e.g., *finally*) remain largely monofunctional (1DSMs), while some (e.g., *after all*) acquire multiple pragmatic uses as their domains of use are extended to different discourse contexts. The Conjunct *after all*, for instance, not only expresses connectivity but have also come to be further deployed as a contrastive DM or a justificational DM, depending on the context. Traugott refers to these extended uses beyond the constructionalization stage as 'constructional changes' (or 'constructional shifts'). In short, contentful, compositional/analyzable and monofunctional *after all* becomes increasingly pragmatic, non-compositional/unanalyzable and multifunctional as its range of use extends along a cline from 1DSM to multiple DMs.

Chapter 5 gets exciting as it reviews alternative diachronic models of language change, with special attention to differences between Discourse Grammar (DG) and Diachronic Construction Grammar (DCG). Discourse Grammar, proposed by Bernd Heine and colleagues, posits a dual-layered grammar comprising a sentence-level grammar and a thetical-level grammar. The latter is influenced by discourse factors and operates beyond the sentence level. The transition from sentence grammar to thetical grammar is realized through 'cooptation', a process whereby the propositional meaning of a construction at the sentence level shifts to a pragmatic meaning at the discourse level. Structurally, the construction that contributes pragmatic meaning becomes syntactically detached from the sentence, hence its emergence as a 'thetical' (a category which includes not only clause-medial parentheticals, but also detached pragmatic markers found at the left and right peripheries). Initially, there were frequent references to cooptation being 'instantaneous', i.e. an abrupt shift. Not surprisingly, given that Diachronic Construction Grammar views the rise of pragmatic markers as a gradual process, Chapter 5 of this monograph takes issue with the characterization of cooptation as an instantaneous process.

Part II, then, is devoted to case studies that meticulously trace how various types of Conjuncts come to be deployed as discourse structuring markers (DSMs), many of which are used as multifunctional discourse markers (DMs). Chapter 6 is devoted to the development of elaboratives such as *also*, *further(more)* and *moreover*; Chapter 7 to contrastive markers such as *but*, *all the same* and *instead*; Chapter 8 to markers of 'digressive' topic shift such as *by the way*, *incidentally* and *parenthetically*; Chapter 9 to markers of 'return to a prior topic' such as *back to point X*; finally, Chapter 10 to the development of combinations of DMs such as *and also*, *so also*, *now then*, and *Oh, by the way*. The detailed diachronic analyses in these chapters in Part II pay special attention to syntactic positions. These not

only include clause-initial, clause-medial and clause-final positions, but also pre-clausal and post-clausal positions. Recognition of a distinction between pre-clausal and clause-initial positions is theoretically important; likewise, recognition of a distinction between post-clausal and clause-final positions. In this respect, Diachronic Construction Grammar may share some common ground with Discourse Grammar, since there are some parallels between pre-/post-clausal DMs and theticals, particularly in terms of detachment from the clause and semantic-syntactic scope expansion. As Traugott points out, discourse structuring markers (DSMs) in post-clausal position tend to have scope over the preceding discourse, while those in pre-clausal position tend to have scope over the following discourse (2022: 62, 88, 99). On this point, then, both analytical frameworks embrace syntactic scope expansion, each in their own way going beyond traditional views which had initially characterized grammaticalization in terms of syntactic scope reduction (2022: 86).

Convergent findings can also be found in other subfields in linguistic studies. Consider, for example, prosodic studies in Sinitic languages, which are well-known for their sentence final particles (SFPs), many of which are often stacked together (not unlike combinations of DMs in English noted in Chapter 10 mentioned above). Pauses in prosodic analyses from Cantonese, for instance, reveal detachments of topics at the left periphery and detachments of sentence final mood particles at the right periphery (Wee 2024, see also Wakefield 2010). These observations are also consistent with formal syntactic accounts, which identify SFPs with broader scope syntactically (given that SFPs c-command their immediate dominating clause, TP) and pragmatically (since SFPs determine the construction type and much of the pragmatic meaning of the sentence, e.g., question, doubt, surprise, assertion, etc.) (Wee 2024: 2, see also Simpson 2014, Cheng & Tang 2022). In sum, a diachronic view that links the rise of discourse markers (DMs) with semantic and syntactic scope expansion, as espoused in Traugott (2022), is well-supported by recent research findings.

In Part III, the last section of the monograph, Traugott initiates discussion of how we might conceptualize the architecture of a dynamic construction network that can account for the rise of circumstantial adverbials (CircAdvs) as Conjuncts, which function as DSMs (inclusive of 1DSMs and DMs). She raises three issues for open discussion as we go about constructing a pragmatic model that is faithful to diachronic evidence. In Chapter 11, Traugott emphasizes the synergistic interaction of three processes: subjectification, intersubjectification, and also 'textualization'. The latter process involves the speaker paying attention to text-creation (i.e., the process of arranging how discourse segments should be linked to each other) and inviting the addressee to interpret the intended textual relationships (Traugott 2022: 198, 202). These three processes occur simultaneously, with textualization resulting in Conjuncts, while subjectification and intersubjectification may continue in strength to develop into DSMs and DMs. All

three processes—subjectification, intersubjectification, and textualization—need to be incorporated into a network model of pragmatic language change.

In Chapter 12, Traugott addresses the question of whether clausal position determines the pragmatic meaning of connectives. Her analysis for English shows that pre-clausal position is the default position for Conjuncts (e.g., *and*, *also*, *so*, *now*), and also for many DSMs, particularly those that still primarily "link back to prior discourse and forward to upcoming discourse" (p. 208), but with some DSMs also found in other positions (e.g., *first of all*, *by the way*). DMs can be found in various positions (pre-clausal, post-clausal and also clause-medial), with multifunctional DMs "likely to have different pragmatic and discourse effects in different positions" (p. 203). Because DMs can be found in various positions, Traugott concludes that clausal position does not necessarily determine the pragmatic meaning(s) of a connective. From such a statement, one could then infer that clausal position is but one of many variables in the conceptualization of a network model of pragmatic language change.

Traugott dedicates Chapter 13 to the description of a 'network' model to help explain the diachrony of DSMs from a Construction Grammar perspective. This network metaphor is inspired by neural network simulations but is not a computational model. It comprises various types of links between form-function pairings, each of which is a construction with "a rich bundle of features with values" (p. 226). Different features become more prominent in different contexts, with their values changing as a construction comes to be linked to different combinations of constructions at different times. This allows a network to be dynamic and malleable. Traugott focuses on two major types of linkages, namely, 'vertical networks' and 'horizontal networks'. Constructions that are linked via vertical networks yield inheritance relations, with links formed between constructions "at different levels of abstraction" (p. 227), while constructions that are linked via horizontal networks yield resemblance relations, often via analogy.

At this juncture, it is worth returning to the question of whether Traugott's emphasis on gradualness in the rise of pragmatic markers can be reconciled with Heine's notion of abruptness or instantaneity associated with the phase referred to as 'cooptation'. Diachronic evidence supports Traugott's observation that new functional categories often emerge gradually, often via overlapping metonymic mappings that take some time before a newly emerging construction is conventionalized and widely adopted by a speech community. Structural evidence can be used to show that neoanalysis (= reanalysis) has taken place, affirming that a new functional category has emerged, but it cannot resolve the question of whether such change is instantaneous or gradual. To some extent, neural network simulations can help shed some light here.

Such networks have produced scatterplot graphs showing that random inputs over time result in similar inputs being attracted to a common (i.e. shared) region in mental space. Each input is made up of patterns of activations, analogous to the rich bundle of features in Construction Grammar (the latter in fact being inspired

by the former). At a more abstract level, each pattern of activation can be represented as a whole unit. Although initially occurring as random and sporadic activations, over time similar patterns of activation begin to boost each other's prominence and visibility and together they begin to form a category that is distinctive, and in this sense become strong enough to 'break away' from other categories, yet without breaking off links with related categories. Such breakaways are comparable to Traugott's notion of 'constructionalization'. In a sense, this is also roughly equivalent to Heine's notion of cooptation, except that constructionalization is gradual while cooptation is thought to be instantaneous. Neural networks are dynamic systems and each pattern of activation is sensitive to neighboring patterns of activation, which gives these systems their malleability, as well as the potential to form other new categories or constructionalizations. What neural networks are able to additionally capture is that change happens in poly-dimensional space. It is this complex scenario of interlocking (hence context-dependent) neural-like processes that Diachronic Construction Grammar tries to highlight.

At the same time, in neural network literature, these rich patterns of activations are sometimes transformed into graphs which provide us with an abstract 2- or 3-dimensional representation of dynamic activations which in fact occur in poly-dimensional space. One such 2-dimensional graph is the sigmoid curve, which shows a relatively flat trajectory that suddenly spikes up in a steep slope before leveling off again. This sigmoid curve helps explain, at a more abstract level, how language change often goes unnoticed in the early stages of grammaticalization, then is sometimes followed by accelerated change (indicated by the steep slope in the sigmoid curve) before the process of change decelerates and plateaus off. The sudden spike is what Heine and colleagues seem to be trying to capture, and which is compatible with what they refer to as 'cooptation'. Heine mentions an early stage of grammaticalization prior to cooptation, and a later stage of grammaticalization following cooptation. These two stages of grammaticalization are comparable to the relatively flat trajectories before and after the sudden spike in the sigmoid curve. In a sense, then, one could say that Heine and colleagues see this sudden spike as the process that flips meaning at the level of sentence grammar to meaning at the level of discourse grammar.

In light of what we know about neural network metaphors, Heine's Discourse Grammar (DG) can be appreciated as a model of pragmatic language change at a more abstract level of analysis, while Traugott's Diachronic Construction Grammar (DCG) models the same phenomenon at a far more fine-grained level of analysis. The former (DG) makes room for researchers to focus more on the discursive aspects of language change, especially for those working from a more socio-interactive perspective. The higher level of abstraction can also be useful for surveying typological tendencies across languages. The latter (DCG) reminds us that pragmatic strengthening is a continuous process that often goes largely unnoticed, essentially because the process is often gradual despite the appearance

at times of (near-)instantaneity when a speech community recognizes the emergence of a new functional category. Given its attention to diachronic details, Diachronic Construction Grammar can further help explain the motivations that influence language-specific variation. Both theoretical models contribute to our growing understanding of the cognitive processes involved in language change; together they expand our views of language at both more neural-like and more abstract levels of analysis.

Chapter 14 of the monograph calls for further explorations into still-unresolved issues. These include a call for new studies investigating discourse structuring markers in other languages, and also diachronic studies within a constructionist framework of pragmatic markers derived from lexical sources other than Conjuncts and CircAdvs. Overall, this 274-page monograph is a valuable addition to the literature on language change and is written with meticulous attention to diachronic data. It should find a place in our library to inspire further diachronic studies that are informed by a deeper understanding of neural network processes. In terms of style, this monograph is intended for fairly advanced language scholars with some background in grammaticalization and pragmaticization studies and interested in theoretical issues based on historical evidence. But it can also be used as a resource book for undergraduate students to trace the development of specific connectives across different chapters, giving them the opportunity to hone their extensive reading and problem-solving skills and mentoring them in diachronic linguistic analysis. In terms of content, this monograph is bound to stimulate further discussions on the processes of language change.

Acknowledgments

References

Cheng, Siu-Pong & Sze-Wing Tang. 2022. Syntax of sentence-final particles in Chinese. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), *Cambridge handbook on Chinese linguistics*, 578–596. Cambridge: Cambridge University Press.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2013. An outline of discourse grammar. In Shannon Bischoff & Carmen Jany (eds.), *Functional approaches to language*, 175–233. Berlin, New York: Mouton de Gruyter.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2017. Cooptation as a discourse strategy. *Linguistics* 55 (4). 813–856.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge, UK: Cambridge University Press.

Heine, Bernd. 2023. *The Grammar of Interactives.* Oxford, UK: Oxford University Press.

Simpson, Andrew. 2014. Sentence final particles. In C.T. James Huang, Y.H. Audrey Li & Andrew Simspon (eds.), *The handbook of Chinese linguistics*, 156–179. Oxford, UK: John Wiley and Sons, Inc.

Traugott, Elizabeth Closs. 2023. *But I hate him just the same*: On the rise of concessive markers with *same*. *Journal of English Linguistics* 51 (4). 311–345.

Wakefield, John. 2010. *The English equivalents of Cantonese sentence final particles: A contrastive analysis.* Hong Kong: Hong Kong Polytechnic University dissertation.

Wee, Lian-Hee. 2024. *The Prosodic Status of the Sentence Final Particle in Hong Kong Cantonese.* Manuscript, Hong Kong Baptist University.

Bionotes:

Foong Ha YAP is Associate Professor at the Chinese University of Hong Kong, Shenzhen. Her research focuses on the development of pragmatic markers. She has numerous publications in peer reviewed journals including *Diachronica, Discourse & Society, Journal of Historical Pragmatics, Journal of Pragmatics, Text & Talk* among others.
e-mail: foonghayap@cuhk.edu.cn
https://orcid.org/0000-0002-7831-0522

Mikyung AHN is Associate Professor at Hankuk University of Foreign Studies, where she received her PhD in Linguistics in 2009. Her research interests include language change, discourse analysis, and cognitive-functional linguistics, with a focus on the development of discourse markers and their functions from both pragmatic and typological perspectives. Her work has been published in *Diachronica*, *Journal of Pragmatics*, *Lingua* and *Text & Talk*./ among others.
e-mail: mkahn@hufs.ac.kr
https://orcid.org/0009-0007-1464-3276

Сведения об авторах:

Фунг Ха ЙЭП — доцент Китайского университета Гонконга в Шэньчжэне. Ее исследования посвящены изменениям прагматических маркеров. Автор многочисленных публикаций в рецензируемых журналах, включая *Diachronica, Discourse & Society, Journal of Historical Pragmatics, Journal of Pragmatics, Text & Talk* и др.
e-mail: foonghayap@cuhk.edu.cn
https://orcid.org/0000-0002-7831-0522

Микюнг АХН — доцент Университета иностранных исследований Ханкук, где она получила степень PhD по лингвистике в 2009 году. Ее исследовательские интересы включают изменение языка, анализ дискурса и когнитивно-функциональную лингвистику с акцентом на изменении дискурсивных маркеров и их функций как с прагматической, так и с типологической точек зрения. Автор статей, опубликованных в рецензируемых журналах *Diachronica, Journal of Pragmatics, Lingua, Text & Talk* и др.
e-mail: mkahn@hufs.ac.kr
https://orcid.org/0009-0007-1464-3276

Review of Bernd Heine, Gunther Kaltenböck, Tania Kuteva and Haiping Long. 2021. *The Rise of Discourse Markers.* Cambridge: Cambridge University Press, xi+308 pp. ISBN 978-1-108-83385-1

Fangqiong ZHAN ⓘ ✉

*Nanyang Technological University National Institute of Education, Singapore*
✉fangqiong.zhan@nie.edu.sg

The Rise of Discourse Markers coauthored by Bernd Heine, Gunther Kaltenböck, Tania Kuteva, and Haiping Long, offers a cross-linguistic examination of the development of discourse markers (henceforth, DMs), highlighting their significant role in communication and their connection to broader sociocultural behaviour. By examining languages from diverse typological backgrounds, the book reveals how DMs facilitate smooth interaction and coherence in communication. Through detailed case studies and theoretical insights, the authors also demonstrate that DMs are not merely linguistic tools but also integral components of human interaction that adapt to the evolving thats of speakers across different cultures and languages. This work highlights the universal and variable aspects of DMs, making it an essential resource for linguists and anyone interested in understanding the dynamics of human communication.

The book presents a novel and thoroughly supported alternative to traditional theories concerning the development of DMs, which have primarily focused on grammaticalization and pragmaticalization. Traditional theories often emphasize how linguistic elements evolve into DMs through changes in grammar or pragmatics. However, this book challenges these perspectives by introducing the concept of "cooptation."

Cooptation is a process where existing linguistic elements are repurposed to serve as DMs, without necessarily undergoing significant grammatical changes. This concept is pivotal to understanding the emergence of DMs and is introduced in the first two chapters: "The Development of Discourse Markers: An
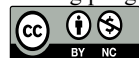
Introduction" and "Concepts of Analysis." In these chapters, the authors lay the theoretical foundation for cooptation, explaining how certain words and phrases are co-opted into discourse as markers that manage conversational flow and structure. This approach allows for a broader and more flexible understanding of how DMs can arise across different languages and contexts. The authors argue that cooptation provides a more accurate account of the development of DMs, as it accounts for the diverse ways in which these markers can emerge and function. By focusing on cooptation, the book shifts the analytical lens from structural changes in grammar to functional adaptations in discourse, offering a fresh perspective that accommodates the complex interplay between language use, social interaction, and cultural influences. This innovative approach provides valuable insights into the dynamic nature of language and the essential role of DMs in communication.

The subsequent four chapters of the book provide an in-depth exploration of the development of discourse markers (DMs) in four languages: English, French, Japanese, and Korean. This section aims to rigorously test the cooptation hypothesis through detailed analyses of each language's unique linguistic landscape and historical development. By selecting languages from different language families and cultural contexts, the authors offer a comprehensive view of how DMs can emerge and evolve through cooptation across diverse linguistic environments. Chapter 7, "Discourse Markers in Language Contact," explores another potential pathway for the evolution of DMs, focusing on how language contact and borrowing can influence the development and spread of DMs, further supporting the cooptation hypothesis by demonstrating how these markers can transcend linguistic boundaries and become integrated into new linguistic contexts. In Chapter 8, "Discussion," the authors address various issues raised in earlier chapters, delving deeper into the factors driving cooptation. The book concludes by summarizing the key points and findings of the study, reinforcing the validity and applicability of the cooptation hypothesis. The authors synthesize the insights gained from the cross-linguistic analyses and discussions, emphasizing the importance of considering functional and contextual factors in the development of DMs. This comprehensive conclusion ties together the book's themes, providing a cohesive understanding of the dynamic nature of DMs and their vital role in human communication.

The book can be divided into three main parts:

Part I Theoretical background and claims (Chapters 1–2): This section lays the groundwork for understanding the authors' approach and introduces the concept of cooptation.

Part II Empirical testing across languages (Chapters 3–6): This part tests the theoretical claims against data from English, French, Japanese, and Korean.

Part III Related issues and broader context (Chapters 7–9): The final section discusses issues related to the earlier chapters, providing a broader context for the findings.

In the following review, I will present the chapters in sequence before offering a general assessment of the book's strengths and contributions.

The book begins with an in-depth introduction to the intricate category of DMs, setting the stage for a thorough exploration of their nature and development. This introduction highlights the significant body of research that has emerged in this dynamic field, reflecting the growing interest in understanding how DMs function within language and communication. The authors define DMs as elements that do not contribute to the semantic content of a sentence or clause; rather, their function is commonly classified as "metatextual." They serve to help speakers organize their thoughts, manage interactions, and signal shifts in topics or perspectives. DMs are fairly independent of the syntactic structure of a sentence and tend to be set off from the rest of the sentence. Their scope is not restricted to any particular part of the proposition; instead, it relates to the proposition as a whole and extends beyond it to the context. While DMs are not confined to positions reserved for adverbials, they can occur in various positions within a sentence, although the left periphery is a preferred position for many of them.

The focus of the book is not simply defining what DMs are, but to explore the processes through which they come into existence. Rather than cataloguing the various forms and functions of DMs, the authors are more interested in understanding the mechanisms that lead to their emergence and development. To illustrate this process, the book shows how certain words evolve from their original non-DM uses to become DMs.

Traditionally, the transition from non-DM uses to DMs has been explained through grammaticalization or, less commonly, pragmaticalization. Landmark studies by Lehmann (1982), Hopper and Traugott (1993), and Brinton (1996) have theorized grammaticalization as a process where lexical items gradually acquire new grammatical functions. However, these features have been less effective in explaining the development of DMs, as the concept of "grammaticalization" is still difficult to explain the main questions: if DMs are of least grammatical effect, can the concept of "grammaticalization" explain items that are not grammatical? How to explain the difference between the formation process of DMs and the standard parameters of grammaticalization (Lehmann 1982). Considering these challenges, pragmaticalization has been proposed as an alternative explanation (Dostie 2009). However, the development of DMs does exhibit some features of grammaticalization, thus, to treat DMs as instances of something other than grammaticalization would be "to obscure its similarities with the more canonical clines" (Traugott 1995: 15).

Compared to other approaches, Diachronic Construction Grammar (DCG) provides a more global perspective of grammatical change (Hilpert 2013, Traugott & Trousdale 2013), a comprehensive understanding of changing grammatical markers in the context of constructional configurations. However, the authors argue that DCG falls short in addressing the central question of the book: How do DMs

evolve from expressions that are deeply embedded in sentence structure to ones that are neither semantically, syntactically, nor prosodically integrated?

Heine et al. (2013) argue that grammaticalization does not adequately explain the development of DMs. Moreover, they propose "cooptation," a process defined as "a cognitive-communicative operation whereby a text segment such as a clause, a phrase, or a word is transferred from the domain or level of sentence grammar and deployed for use on the level of discourse organization" (p. 26). The root "co-opt" of the term "cooptation" originally defined as a verb by the Cambridge English Dictionary as selecting a new member for a group by the current members. The term "cooptation" is used in the literature that there is a shift from one field or body of operation to another. In this book, the term is metaphorically used to involve a shift from the field of sentence organization to that of discourse processing, thereby being integrated in the structure of discourse organization. Cooptation represents a significant shift in a DM's development, moving it from the sentence level to the discourse level, with grammaticalization processes continuing before and after this shift. The proposed schema for DM development, as outlined in Heine 2013, is represented as "(Grammaticalization >) cooptation > grammaticalization" (p. 27).

A key conclusion from the discussions in Chapter 1 is that both grammaticalization and cooptation are crucial in the development of DMs. While grammaticalization, a well-established field, provides significant insights into how linguistic elements evolve, cooptation plays an equally vital role in this process. To address these points, Chapter 2 of the book, "Concepts of analysis" delves into the details of these two mechanisms. The chapter explores how grammaticalization and cooptation contribute separately to the evolution of DMs and examines their interactions in shaping the structure of these markers.

In this book, grammaticalization is defined as the development from lexical to grammatical forms and from grammatical to even more grammatical forms (Heine & Kuteva 2002: 2). Cooptation, on the other hand, is described as "a fully productive operation whereby a chunk of sentence grammar, such as a word, a phrase, a reduced clause, a full clause, or some other piece of text, is deployed for use on the metatextual level of discourse processing, thereby turning into a thetical" (p.67). These two mechanisms are held responsible for the rise of DMs.

The following constructed examples are given to illustrate the process of cooptation at work in English (p. 65):

(1) a. ***You probably know already*** *that our chairman will resign next month.*
     b. *Our chairman**, you probably know already,** will resign next month.*
     c. *Our chairman**, you know,** will resign next month.*

Operated by cooptation, the text piece printed in bold can be derived from a construction like that in (1a) in a principled way and be inserted in another sentence as in (1b). As example (1c) shows, coopted text chunks tend to have the appearance of "elliptical" clauses or phrases. What matters is that any "missing" information be recoverable by the hearer from the situation of discourse (the "context").

Cooptation is a cognitive-communicative operation which enables speakers to switch their perspective from the level of reasoning anchored in the meaning of sentences to a meta-level of reasoning immediately anchored in the situation of discourse (cf. Rhee 2013: 67). Once coopted, DMs may undergo typical effects of grammaticalization including context extension (and new discourse functions), desemanticisation, decategorialisation (and univerbation), erosion, etc. The subsequent four chapters of the book are dedicated to testing the cooptation theory using authentic linguistic data from four distinct languages. The authors aim to evaluate the theory's validity by examining how coopted DMs in different linguistic and cultural contexts undergo these grammaticalization effects. This comparative analysis seeks to confirm whether the patterns observed align with the predictions made by cooptation theory and to explore the generalizability of its claims across various languages.

Chapters 3 through 6 analyze the evolution of DMs from four different languages: English, French, Japanese, and Korean. Each chapter examines the diachronic development of selected DMs, drawing on data from existing research to illustrate how these markers have evolved over time. The analysis is framed within the theoretical approach involving both grammaticalization and cooptation. This approach views the development of DMs as a shift from sentence-level grammar to discourse-level grammar, where markers transition from their original grammatical functions to serve broader metatextual roles.

Chapter 3, titled "English Discourse Markers," explores twelve distinct markers through the lens of the cooptation hypothesis. These markers include *after all*, *anyway*, *I mean*, *if you like*, *if you will*, *instead*, *like*, *no doubt*, *right*, *so to say / so to speak*, *well*, and *what else*. To illustrate the analysis, let's consider the case of *anyway*.

There are three main usages of *anyway*, namely an additive, a dismissive and a resumptive one. Only the second usage is that of a DM, classified by Fraser (2009) as a topic orientation marker. The DM *anyway* is primarily associated with two main usages: at the left periphery of an utterance and at its right periphery. In the former, it signals an upcoming discourse, indicating a particular conception of what will unfold in the emerging discourse relative to the preceding discourse. In the latter usage, *anyway* functions as a modal and intersubjective marker, connecting two units that have already been produced and expressing a conceptual-conditional relationship.

The DM can be traced back to the Old English noun phrase *ænig weg* "any way, path, route," changing to Middle English *any wei/way*, which was grammaticalized to an adverbial meaning "in any way or manner, by any means," already attested in Late Middle English (1350–1500) (Haselow 2015: 172). The text data available suggest that around the mid-nineteenth century, roughly between 1840 and 1865, if not earlier, *anyway* clearly had acquired the main hallmarks of a coopted unit and of a DM. The main evidence showing that *anyway* had turned into

a thetical by the middle of the nineteenth century or earlier is summarized in Haselow (2015: 99).

a Meaning: The meaning of *anyway* is no longer part of the meaning of its host sentence, its validity is not bound to or inferentially linked to the propositional content of its host.

b Function: *Anyway* now seems to function as a metatextual unit, it establishes a sequential relation between discourse units rather than relations within the sentence (Haselow 2015: 180), marking the unit it accompanies as contextually linked to prior discourse.

c Syntax: Changing from sentence-internal to sentence-external expression it now lacks constituent status, it "neither governs any of the constituents of the structural units it refers to, nor is itself governed by any other constituent" (Haselow 2015: 176, 180).

d Prosody: It is likely to be set off prosodically, forming an intonation unit of its own (Haselow 2015: 178).

e Semantic–pragmatic scope: It has wider scope, extending over a sequence of discourse units. At the left periphery, it has scope over larger chunks of discourse while at the right periphery it has scope over two adjacent units of talk (Haselow 2015: 180, 183).

f Placement: It changed from clause-internal adverbial to DM placed at the right or the left periphery of its host.

To conclude, grammaticalization occurred in two distinct phases. First, prior to cooptation, *anyway* gradually transitioned from a noun phrase to a more abstract adverbial, losing its concrete lexical content through desemanticization and its internal word boundary through internal decategorialization. Second, there appears to have been a later phase of grammaticalization following cooptation, during which *anyway* developed a paradigmatic relationship with other DMs–such as *then*, *though*, *but*, and *actually* in the case of right-periphery *anyway* (Thompson & Suzuki 2011, Haselow 2013), and *actually*, *well*, and *so* in the case of left-periphery *anyway* (Haselow 2015: 180–81).

Some scholars including Brinton (2008) propose "indeterminate structures" in Middle English. For example, *I mean* has its full content meaning used in initial position but is followed by a *that*-less complement. In such contexts, it is hypothesized, *I mean* can be interpreted either as a matrix clause introducing a complement clause or as the adverbial 'namely,' which is an example of indeterminate structure. In Section 3.3, the authors argue that the instances of intermediate structure and "ambiguity" identified by Brinton (2008) and others are problematic due to the communicative intent of the speaker or writer, the researcher's specific interpretation, or the absence of sufficient contextual or grammatical information. Thus, generalizations about gradualness and intermediate stages offer valuable tools for reconstructing grammaticalization. However, both must be applied with caution when reconstructing the emergence and evolution of DMs.

Chapter 4, titled "French Discourse Markers," examines seven markers: *à la rigueur, à propos, à ce propos, alors, en fait, au fait,* and *enfin.* The analysis of each marker leads the authors to conclude that grammaticalization alone cannot fully explain the transition from sentence grammar to thetical grammar, making it necessary to hypothesize the sudden cooptation of the item in question.

The hedging DM *à la rigueur* is not attested in the sixteenth and seventeenth centuries, the first instance of it is found in the eighteenth century. Subsequently its use increased dramatically, with the result that in the twenty-first century it is the only one surviving. It appears to have turned into a thetical in the early eighteenth century, if not earlier, and it became a full-fledged DM. There is no grammatical evidence to suggest that the DM properties emerged in a gradual way one after the other in the history of *à la rigueur.*

The case studies of DMs in Japanese, presented in Chapter 5, focus on a selection of seven markers: *dakedo, demo, douride, ga, jijitsu, sate,* and *wake.* The historical development of the DM *dakedo* can be understood in terms of two stages (Onodera 1995, Aoki 2019). Stage I, spanning from the eighteenth to the early twentieth century, marks the emergence of the clause-final construction [V + *kedo*]. The [V + *kedo*] construction continued to be used when sentence-initial *dakedo* emerged in the early twentieth century at Stage II, which covers the period from the early twentieth century to the present. There is no information to suggest that the change from Stage I to Stage II took place gradually, involving stages as they characterize grammaticalization and, as far as the evidence provided by Onodera suggests, the changes appear to have occurred rather simultaneously. Chapter 5 concludes that, overall, there is converging evidence supporting the cooptation hypothesis proposed in this chapter.

Chapter 6 notes that the Korean language displays an extraordinary richness of discourse marker structures, and Korean DMs also stand out with regard to the wide range of functions they express. Another interesting feature of Korean can be seen in the fact that it has developed a range of DMs from interrogative constructions via rhetorical questions, that is, questions that do not seek information or require an answer (Rhee 2004: 413). Furthermore, prosodic marking appears to strongly correlate with the respective functions expressed by a given DM. The case studies focus on four DMs: *icey, makilay, maliya,* and *tul.* In contemporary Korean, *makilay*-DMs are very popular among youngsters – to the extent that between 2011 and 2017 there was a children's TV show in South Korea named the *Makilay Show* (Rhee 2019). The *makilay*-DMs are transparently derived from clause-like constructions – a process whereby constructions like *makwu ilehkey ha-e(yo)* 'reckless like.this say-SFP(POL)' turned into *makilay(yo)* 'DM' (Rhee 2013). The main grammatical changes leading to the rise of *makilay* DMs suggest that these changes align with the hypothesized effects of cooptation, as well as with observations made in the development of other DMs in Korean and other languages.

Chapters 7 and 8 shift focus from case studies to explore alternative pathways in the development of DMs, adopting a broader typological perspective on the issue.

Chapter 7 explores the development of DMs through borrowing in contexts of language contact, presenting this as a common alternative to cooptation. Given that DMs are deeply entrenched in the socio-cultural environment and discourse organization of a language, it might be surprising that they can be influenced by language contact more significantly than many other types of linguistic expressions.

Borrowing involves taking a DM as a form-meaning unit from one language and inserting it into another. Replication, including loan translation, is a more complex process that involves meaning and/or structure but not form (Heine & Kuteva 2005). While borrowing may involve just one or a few DMs, it can also affect an entire repertoire of such markers. These observations lead to the question: Why are DMs frequently borrowed from one language and incorporated into texts of another language? A comparative survey suggests that the following types of factors are most commonly cited to account for the role of DMs in bilingual situations, either as reasons for or contributors to the borrowing of DMs: formal linguistic factors; text organization; attitudes of the speaker; speaker-hearer interaction.

An interesting and critical point raised in the concluding paragraphs of this chapter is that a DM arising from language contact often remains psychologically linked to its donor language. Its use may evoke cultural and/or emotional features associated with the donor language community (p. 235). This connection is likely what most distinguishes borrowed DMs from native ones in a given context.

Chapter 8 addresses issues that emerged in previous chapters but could not be discussed in detail at the time. The first section of the chapter focuses on the dualistic organization of linguistic discourse. DMs were defined as having a metatextual function and the term "metatextual" implies that, in addition to a "textual" level there is a second level of discourse processing. The dualistic model of language activity builds on earlier research (Heine 2019) and ongoing work, distinguishing between a microstructure and a macrostructure of discourse. The microstructure corresponds to sentence grammar, while the macrostructure aligns with thetical grammar (p. 237). The distinction between microstructure and macrostructure is not only supported by evidence from a wide range of data and different research traditions, it is also supported by neurolinguistic findings (Heine, Kuteva & Kaltenböck 2014, Heine et al. 2015, Haselow 2019, Heine 2019). The second section of the chapter examines the factors that lead people to create new DMs and those that facilitate cooptation. It also addresses the types of text segments selected for cooptation and the inherent limitations of this process. Section 8.3 examines constituent-anchored DMs, which modify a specific constituent rather than the entire utterance. Research suggests that these DMs evolve from utterance-anchored DMs, indicating that they may represent a stage of late grammaticalization occurring after the process of cooptation. The fourth section of the chapter addresses imperative-derived DMs. "Canonical imperatives," which are positive singular imperatives directed at a singular referent (such as a hearer, reader, or signee), express commands or requests. These forms are typically non-finite and concise (Aikhenvald 2010: 18) and often exhibit thetical properties. The authors

address imperative-derived DMs separately, noting a key distinction in their development compared to many other DMs: Imperatives are inherently thetical (Heine 2016). Consequently, their evolution into DMs bypasses cooptation and is primarily characterized by late grammaticalization, during which they progressively transform into text-structuring devices. The final section of the chapter looks at fillers or hesitation markers, such as *ano* in Japanese. The authors differentiate between primary fillers, which function solely as fillers and are similar to interjections, and secondary fillers, which some languages use in additional roles (typically as demonstratives). For secondary fillers, the authors propose a potential development through cooptation, analogous to the process described for other DMs in the book.

In the concluding chapter, the authors reaffirm that the primary aim of the book was to address the question: Why do DMs possess the grammatical properties they do? According to the perspective adopted in this book, language users are creative agents who frequently employ linguistic resources in innovative ways. They may extend an expression's use to new contexts for specific cognitive and communicative purposes, gradually grammaticalizing it into a form that serves new functions. Alternatively, through cooptation, they might adapt elements of sentence grammar to the metatextual level of discourse, leading to the emergence of new DMs.

The Rise of Discourse Markers offers a significant contribution to metadiscursive studies, presenting a central hypothesis that DMs develop through cooptation. This process involves a shift from sentence grammar (microstructure) to thetical grammar (macrostructure). While grammaticalization plays a role both before and after cooptation, the authors argue that it alone cannot account for the profound transformation required for an expression to become a DM. The book challenges the traditional reliance on grammaticalization theory and introduces a nuanced approach to understanding the emergence of DMs. By focusing on cooptation as a key factor, it provides a compelling and insightful exploration of the evolution and function of DMs. This approach not only offers significant theoretical implications but also encourages further research into the mechanisms shaping metatextual phenomena.

The authors provide a robust critique of existing explanations that rely solely on grammaticalization or pragmaticalization, highlighting the limitations of these theories. They advocate for a more refined understanding of how metatextual items emerge, supported by extensive research and analysis. The book situates its study within the broader historical development of DMs, enriching readers' understanding of language dynamics and the factors influencing DM evolution. Moreover, *The Rise of Discourse Markers* includes a detailed examination of case studies and empirical data, offering concrete examples of how cooptation functions across different languages. The comparative analysis presented adds depth to the discussion and underscores the universal aspects of the proposed model, making it a valuable addition to the field of metadiscursive studies.

Nevertheless, there has been rapid progress recently in the theoretical framework of the development of DMs, including significant advancements in Diachronic Construction Grammar (DCG). While the study acknowledges DCG, it does not fully explore its potential contributions to understanding linguistic change. The diachronic constructional approach to the formation of DMs is inspired by Cognitive Construction Grammar, as conceptualized in Goldberg (2003, 2006) and Croft (2001). Traugott (2020, 2022) takes up the challenge to find ways of infusing more pragmatics into constructionalist thinking, investigates the development of constructions that can be used to do discourse structuring work, and develops a systematic and nuanced approach to the history of discourse structuring markers (DSMs, which include two types – 1DSMs and DMs – both function as connectors) from DCG (Traugott & Trousdale 2013, Barðdal et al. 2015). Given the recent advancements in DCG, incorporating a more thorough analysis of this framework could enhance the study by providing deeper insights into the evolution of constructions and their role in language change. Future research that integrates DCG more extensively could offer valuable perspectives and contribute to a more comprehensive understanding of linguistic evolution.

Overall *The Rise of Discourse Markers* is a groundbreaking work that offers a fresh perspective on the development of DMs through the process of cooptation. It is a significant contribution to the field of metadiscursive studies and is highly recommended for researchers and scholars interested in the intricacies of language evolution and the role of DMs. The book's depth of analysis and theoretical advancements make it an essential read for anyone engaged in the study of discourse.

Acknowledgements

References

Aoki, Hirofumi. 2019. Setsuzokushi to bunpooka – Chusei kooki 'shoomono shiryou' o chuushin ni [Conjunctions and grammaticalization – With special focus on the Late Medieval "Shoomono data"]. Paper presented at the Workshop on Grammaticalization (GJNL-4), Tohoku University.

Aikhenvald, Alexandra Y. 2010. *Imperatives and Commands (Oxford Studies in Linguistic Theory)*. Oxford: Oxford University Press.

Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.). 2015. *Diachronic Construction Grammar.* Amsterdam: John Benjamins.

Brinton, Laurel. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions (Topics in English Linguistics, 19).* Berlin: Mouton de Gruyter.

Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development (Studies in English Language).* Cambridge: Cambridge University Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective.* Oxford: Oxford University Press.

Dostie, Gaetane 2009. Discourse markers and regional variation in French: A lexico-semantic approach. In Kate Beeching, Nigel Armstrong & Francoise Gadet (eds.), *Sociolinguistic variation in contemporary French,* 201–214. Amsterdam: John Benjamins.

Fraser, Bruce. 2009. Topic orientation markers. *Journal of Pragmatics* 41. 892–898.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7. 219–224.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language.* Oxford: Oxford University Press.

Haselow, Alexander. 2013. Arguing for a wide conception of grammar: The case of final particles in spoken discourse. *Folia Linguistica* 47 (2). 375–424.

Haselow, Alexander. 2015. Left vs. right periphery in grammaticalization: The case of *anyway*. In Andrew D. M. Smith, Graeme Trousdale & Richard Waltereit (eds.), *New directions in grammaticalization research,* 157–186. Amsterdam: John Benjamins.

Haselow, Alexander. 2019. Discourse markers and brain lateralization: Evidence for dual language processing from neurological disorders. Lecture presented at the University of Düsseldorf, December 15.

Hilpert, Martin 2013. *Constructional Change in English: Developments in Allomorphy, Word-Formation and Syntax.* Cambridge: Cambridge University Press.

Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51 (6). 1205–47.

Heine, Bernd. 2016. On non-finiteness and canonical imperatives. In Claudine Chamoreau & Zarina Estrada-Fernández (eds.), *Finiteness and nominalization (typological studies in language 113),* 245–270. Amsterdam: John Benjamins.

Heine, Bernd. 2019. Some observations on the dualistic nature of discourse processing. *Folia Linguistica* 53 (2). 411–442.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long. 2015. *On Some Correlations between Grammar and Brain Lateralization (Oxford Handbooks Online in Linguistics).* New York: Oxford University Press.

Heine, Bernd, Tania Kuteva & Gunther Kaltenböck. 2014. Discourse grammar, the dual process model, and brain lateralization: Some correlations. *Language & Cognition* 6 (1). 146–180.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva & Haiping Long 2013. An outline of discourse grammar. In Shannon Bischoff & Carmen Jany (eds.), *Functional approaches to language,* 175–233. Berlin: Mouton de Gruyter.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change (Cambridge Approaches to Language Contact, 3).* Cambridge: Cambridge University Press.

Hopper, Paul J. & Elizabeth C. Traugott. 2003[1993]. *Grammaticalization.* Cambridge: Cambridge University Press.

Lehmann, Christian. 2015[1982]. *Thoughts on Grammaticalization.* Berlin: Language Science Press.

Onodera, Noriko Okada. 1995. Diachronic analysis of Japanese discourse markers. In Andreas H. Jucker (ed.), *Historical pragmatics*, 393–437. Amsterdam: John Benjamins.

Rhee, Seongha. 2004. From discourse to grammar: Grammaticalization and lexicalization of rhetorical questions in Korean. In Gordon Fulton, William J. Sullivan & Arle R. Lommel (eds.), *LACUS: Forum XXX: Language, thought and reality,* 413–423. Houston: Lacus.

Rhee, Seongha. 2013. "I know I'm shameless to say this": Grammaticalization of the mitigating discourse marker *makilay* in Korean. *Procedia: Social and Behavioral Sciences* 97. 480–486.

Rhee, Seongha. 2020. On determinants of discourse marker functions: Grammaticalization and discourse-analytic perspectives. *Linguistic Research* 37 (2). 289–325.

Thompson, Sandra A. & Ryoko Suzuki. 2011. Grammaticalization of final particles. In Heiko Narrog & Bernd Heine (eds.), *The Oxford handbook of grammaticalization,* 665–677. Oxford: Oxford University Press.

Traugott, Elizabeth C. 1995. Subjectification in grammaticalization. In Dieter Stein & Susan Wright (eds.), *Subjectivity and subjectivisation*, 31–55. Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs. 2020. The development of "digressive" discourse-topic shift markers in English. *Journal of Pragmatics* 156. 121–135.

Traugott, Elizabeth Closs. 2022. *Discourse Structuring Markers in English: A Historical Constructionalist Perspective on Pragmatics*. Amsterdam: John Benjamins.

Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Bionote:

Fangqiong ZHAN is an Assistant Professor at National Institute of Education, Nanyang Technological University, Singapore. She received her PhD in Chinese Linguistics from Stanford University, USA. Her research interests include Chinese linguistics, discourse grammar, and language acquisition and pedagogy. Her numerous publications occur in high-level peer-reviewed journals including *Journal of Pragmatics*, *Chinese Language and Discourse*, *Lingua*, *Diachronica*, *Journal of Historical Pragmatics*, *Studies in Language*, among others.
e-mail: fangqiong.zhan@nie.edu.sg
https://orcid.org/0000-0002-5791-8205

Сведения об авторе:

Фаньцюн ЧЖАНЬ – доцент Национального института образования, Наньянский технологический университет, Сингапур. Получила степень PhD по китайской лингвистике в Стэнфордском университете, США. Ее научные интересы включают китайскую лингвистику, грамматику дискурса, а также педагогику и изучение иностранного языка. Автор многочисленных публикаций рецензируемых журналах, включая *Journal of Pragmatics*, *Chinese Language and Discourse*, *Lingua*, *Diachronica*, *Journal of Historical Pragmatics*, *Studies in Language* и др.
e-mail: fangqiong.zhan@nie.edu.sg
https://orcid.org/0000-0002-5791-8205