

Original Paper

Consensus, Controversy, and Prospects in the Study of the Evolution of Discourse Markers

Haixia Wang^{1*}

¹ School of English Studies, Beijing International Studies University, Beijing, China

* Haixia Wang, School of English Studies, Beijing International Studies University, Beijing, China

Received: December 2, 2025 Accepted: December 21, 2025 Online Published: December 31, 2025

doi:10.22158/sll.v9n4p144

URL: <http://dx.doi.org/10.22158/sll.v9n4p144>

Abstract

Research on the evolution of discourse markers has yielded abundant results in recent decades. This study systematically reviews the relevant literature at home and abroad. It first examines the basic views on the evolution mechanism of discourse markers from the perspective of grammaticalization, lexicalization, pragmaticalization and cooptation, and then analyzes the differences between the views and the causes of the divergence. This study also discusses the interconnection and intersection of different mechanisms. Finally it puts forward that scholarly attention should be given to a multilevel discussion on the evolution mechanisms and cross-linguistic study on the formation mechanism of discourse markers in future research.

Keywords

evolution mechanism, grammaticalization, lexicalization, pragmaticalization, cooptation

1. Introduction

Discourse markers, expressions like *actually*, *like*, *well*, and *you know*, can be used to refer to a class of recurrent and formulaic linguistic items that are syntactically optional, generally have little propositional meaning but are multifunctional, operating on textual and (or) interpersonal level in conversation. The growing attention to the crucial role of discourse markers in communication has spurred a significant increase in research within this field, primarily across two dimensions: synchronic and diachronic. The former mainly focuses on the functional description of individual cases or certain types of discourse markers, whereas the latter is primarily based on historical corpora to investigate their origins, evolution paths and formation mechanisms. Research on the evolution of discourse markers has been carried out since the 1990s, with its foundational theories primarily including grammaticalization, lexicalization, pragmaticalization and cooptation. Among them, research grounded

in grammaticalization theory is dominant and the mechanism of cooptation is a new trend in recent years. Notably, synchronic studies of discourse markers have been extensively reviewed, but there is a relative lack of systematic reviews on the studies of the evolution of discourse markers. This paper examines the relevant literature, primarily clarifying the various theoretical perspectives on the evolution of discourse markers, exploring their consensus and controversy, and on this basis, outlining future research directions.

2. Different Evolution Mechanisms of Discourse Markers

2.1 Grammaticalization

Many scholars believe that discourse markers are derived from lexical components or sequences with propositional meaning. Their evolution over time has gone through processes identical to traditional lexical grammaticalization, including reanalysis, decategorialization, desemanticization and phonological reduction while exhibiting the characteristics of unidirectionality and gradualness. Thus, the formation of discourse markers represents a typical phenomenon of grammaticalization (e.g., Traugott, 1995; Traugott & Dasher, 2002; Brinton & Traugott, 2005; Wang, 2005; Brinton, 2008, 2011). Related studies have been conducted mainly in terms of desemanticization, syntactic reanalysis and decategorialization.

Desemanticization in the traditional sense refers to the generalization or abstraction of the original propositional meaning of words (Brinton, 2008). A large number of scholars believe that in the process of grammaticalization discourse markers are desemanticized in the sense that the items not only lose the specific and concrete semantic substance, but also acquire procedural and pragmatic meanings replacing their propositional or referential meanings. Traugott (1982) proposes in her early research that the trajectory of the semantic change in grammaticalization is “propositional (> textual) > expressive”, which later scholars modify into several different semantic-pragmatic evolutionary tendencies, arguing that discourse markers do not have propositional or referential meanings, but intersubjective and procedural meanings instead (Traugott & Dasher, 2002, p. 40). Considering the complex relationship between procedural meaning, intersubjective meaning and non-truth-conditionality, Brinton (2008, p. 26) characterizes the semantic evolution of discourse markers as “referential > non-referential (pragmatic, metalinguistic, procedural)”.

Reanalysis of syntactic structure goes hand in hand with semantic-pragmatic evolution (Brinton, 2008, p. 24). It is “the development of new out of old structures” (Hopper & Traugott, 1993, p. 41), embodied in different syntactic paths of evolution. Traugott (1982), in an early study of the evolution of *why* from interrogative adverbials to discourse markers, proposes the first cline, i.e., “adverb > conjunction > discourse marker” and later (1995) puts forward the development path “clause-internal adverbial > sentence adverbial > discourse particle”. In addition, Stenström (1998) proposes a more comprehensive path, i.e., “lexeme(s) > sentence connective > discourse marker”. Focusing on the nominal type noun

constructions with *kind/**sort of*, Brems and Davidse (2010) figure out the main evolution paths and multiple inheritance links in the type noun constructional network. Unlike the above scholars, Brinton (2008) proposes a series of clines of evolution for discourse markers originating from different sentence structures, which can be seen as a necessary complement to the evolution paths for the development of words or phrases into markers. The evolution of some first-person matrix clause structures, including *I think*, *I'm sorry*, and some second- and third-person items like *you know*, *you see*, is consistent with the “matrix clause hypothesis” (Brinton, 2008, p. 21), namely, “matrix clause > parenthetical > comment clause (discourse marker)”. Brinton suggests a number of more complex and comprehensive trajectories for discourse markers with different source structures. For example, the evolution path for the items which resemble “a syntactically defective matrix clause” or an imperative matrix clause such as *I say*, *let me see*, *look*, *listen*, etc. is “matrix [subject + verb (present tense)] /matrix [verb (imperative)] > indeterminate matrix/parenthetical > comment clause (discourse marker)” (2008, p. 251).

As to decategorialization, it typically involves shift from a more major to a more minor grammatical class. Some scholars (e.g., Traugott, 1995; Brinton, 1996; Brinton & Traugott, 2005) who hold a grammaticalization view of discourse markers have argued that the items concerned, after undergoing a reanalysis from propositionally meaningful expressions to discourse markers, lose, due to the transformation into parentheticals, their “behavioral characteristics, such as the ability to take adverbial or modal modifiers, to govern a complement, and so on” (Brinton 2008, p. 244), thus undergoing a process of decategorialization.

Overall, incorporating the evolution of discourse markers into grammaticalization has substantially enriched its definition—the process whereby lexical items or structural formulas acquire grammatical functions in a given context and, once grammaticalized, will further develop new grammatical functions (Hopper & Traugott, 1993). It is obvious that grammar and grammatical function here go beyond the traditional meaning of grammaticalization.

2.2 Lexicalization

Brinton and Traugott (2005, p. 96) define lexicalization as “the change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction of the word formation pattern. Over time there may be further loss of internal constituency and the item may become more lexical”, which is different from the ordinary process of word formation and (de) grammaticalization. According to Dong (2007, 2010), the lexicalization of discourse markers involves the chunking and integration of frequently co-occurring components syntactically, as well as the conventionalization and semanticization of utterance meaning semantically.

In terms of the lexicalization of English discourse markers, some scholars including Krug (1998), Wischer (2000), Fischer (2007) etc. have carried out in-depth analyses. Krug (1998), for instance,

maintains that in the process of *is it not?* evolving into the discourse marker *innit* via the intermediate stage *in't it*, lexicalization occurs, that is, a complex phrase becomes a monomorphemic word. This process involves the form becoming invariant, inseparable, and morphologically opaque. There is also desemanticization of *it*, loss of phonological substance, fixing into a semi-institutionalized spelling and acquisition of pragmatic functions (mainly a turn-turning function). Fischer (2007) argues that the formation mechanism of *I think* type discourse markers is lexicalization: as formulaic tokens “they lose some referential content, being narrowed down to a more epistemic, evaluative meaning. In non-standardized languages they are likely to form one lexical unit in the course of time” (2007, p. 116). She also points out that discourse markers retain more of their lexical meaning in their evolution than is typical grammaticalization, and they undergo the process of bonding or fusion, all of which are features of lexicalization rather than grammaticalization.

Dong's (2007) diachronic study of Chinese discourse markers *shui zhi dao* and *bie shuo* shows that they are the result of lexicalization: evolving from the original free phrases into single word-like units, they acquire discourse marker functions through the conventionalization and semanticization of utterance meaning. She further proposes that discourse markers are not syntactically obligatory; they do not tend to be cliticized as some other functional categories; and they can have variants in form, carrying more lexical features. The morphologically unstable discourse markers are lexicalized, but their degree of lexicalization is not yet very high, and they are only at the early stage of lexicalization, that is, idiomatization. This is also demonstrated in her research on the discourse marker *wo gao su ni* derived from complete clausal forms, and she further argues that an idiom undergoes a process of increasing lexicalization from having multiple variants to a reduction in variant forms until they disappear entirely (Dong, 2010).

2.3 Pragmaticalization

Some scholars (e.g., Erman & Kotsinas, 1993; Aijmer, 1997; Frank-Job, 2006; Norde, 2009) have argued that since discourse markers do not belong to any of the known lexical categories, are usually independent of syntactic structure, have non-truth-conditional meaning, but have a pragmatic function, discourse markers may not be part of “grammar proper” and therefore cannot be the result of grammaticalization. They thus put forward the hypothesis of the pragmaticalization of discourse markers, i.e., lexically meaningful constituents develop into discourse markers either without going through an intermediate stage of grammaticalization or undergoing a grammaticalization that is different from the standard process, and ultimately function primarily as textstructuring devices at the discourse level (Erman & Kotsinas, 1993). Obviously, “pragmaticalization” was largely proposed to distinguish it from “grammaticalization”. They are two different ways in which lexical items evolve into functional expressions: while the former leads to the formation of discourse markers, the latter leads to the emergence of grammatical markers (Erman & Kotsinas, 1993; Aijmer, 1997).

Of course, there are differences among scholars in their specific understanding of how discourse markers acquire pragmatic functions through pragmatalization. For example, based on the examination of typical examples include *listen*, *look*, *well*, *okay* and *here now*, Frank-Job places particular emphasis on the role of habitualization and automatization in the formation of discourse markers in metacommunicative contexts, arguing that pragmatalization is “the process by which a syntagma or word form, in a given context, changes its propositional meaning in favor of an essentially metacommunicative, discourse interactional meaning” (2006, p. 361). According to Aijmer (1997), when a lexical item becomes a pragmatic expression, it can be said to have undergone pragmatalization, as illustrated in the case of *I think*, which develops meanings involving the speaker’s attitudes towards the hearer or the message. In her view, the items that undergo pragmatalization have non-truth-conditional meaning, and their occurrence is optional. Otherwise, pragmatalization and grammaticalization are quite similar. Qiu and Sun (2011) contend that pragmatalization conventionalizes conversational implicatures and fossilizes them in linguistic expressions producing pragmatic markers with specific pragmatic functions (including interpersonal interaction, expression of emotion and attitude, discourse functions etc.), and as a result forms the holistic subsystem. Unlike the above scholars, Waltereit’s (2002) explanations of discourse markers do not directly use the term pragmatalization, but are closest to it (see Brinton, 2008). He argues that pragmatic markers are formed as a result of the “abuse” of some expressions with their own rhetorical potential to convey textual and interpersonal meanings in certain contexts. Such uses become more frequent and the “abused” expressions then are reanalyzed in some contexts as pragmatic markers, whose conversational implicatures are conventionalized and used in additional contexts. Overall, considering both the outcome and process of evolution, pragmatalization emphasizes that discourse markers belong to the pragmatic rather than the syntactic category and highlights the acquisition of their pragmatic functions.

2.4 Discourse Grammar

In recent years, within the framework of Discourse Grammar (Kaltenböck et al., 2011; Heine et al., 2013; Kaltenböck & Heine, 2014), some scholars examine the rise of discourse markers from a new perspective. Discourse Grammar consists of two distinguished domains of discourse organization, namely, Sentence Grammar (SG) and Thetical Grammar (TG). Whereas the former is organized in terms of propositional concepts and clauses and their combination, the latter concerns the overall contours of discourse beyond the sentence, relating to all components of the situation of discourse. SG and TG interact in various ways in organizing linguistic discourse and most of all the interaction is via the mechanism of cooptation, a ubiquitous operation whereby a chunk of SG, such as a clause, a phrase a word, or any other unit is deployed for use as a thetical (Kaltenböck et al., 2011). Cooptation can occur at any time and there are almost no restrictions on the form of the units and their position in the main structure.

Scholars holding the cooptation view argue that the formation of discourse markers should not be interpreted as a traditional process such as grammaticalization or pragmaticalization, but should be incorporated into Discourse Grammar, or more specifically as cooptation (e.g., Heine, 2013; Heine et al., 2013, 2017, 2021; Long & Wang, 2014). For the formation of discourse markers, cooptation implies an expansion of the semantic-pragmatic scope of a chunk (from syntax to discourse context), which entails a series of syntactic and semantic changes (Heine, 2013). Firstly, the unit is no longer constrained by the requirements of its previous syntactic function. Secondly, the unit responds to the context, which mainly involves factors such as discourse organization, speaker-listener interaction, and speaker attitudes, which also means that it has a more complex meaning. As a representative scholar of cooptation, Heine (2013) examines in detail the evolution of *what else*, *I mean*, *look*, etc. over time, and finds that the formation of discourse markers is spontaneously accomplished rather than gradually developed. In addition, Heine also obtains strong evidence of the non-unidirectionality of discourse marker formation from the studies of other scholars such as Jucker (1997), whose research of *well* shows that from Old English to Modern English, it has gone through a semantic evolution of “interpersonal—textual—interpersonal”, which is clearly not unidirectional. Long and Wang’s (2014) examination of the Chinese thetical *shi de* demonstrates its instantaneous rather than gradual nature as well, suggesting that its formation should not be construed as pragmaticalization but rather as cooptation. Thus, apart from the features in scope, syntax, semantics, morphophonology, prosody and word order, cooptation places special emphasis on the instantaneous and non-unidirectional nature of the formation of discourse markers, which is obviously different from other views.

Similar questions also arise in Davidse et al.’s (2015) diachronic investigation of *(there/it is /I have) no doubt*. The sudden emergence of *(it/there) is no doubt* with grammatical and epistemic meaning in 1350-1420 is more or less instantaneous rather than gradual and it can be said to be the result of cooptation. Their research on *(I have) no doubt* indicates that the formation of discourse markers should be understood as the result of a combined effect of grammaticalization, lexicalization and TG; the principles of TG is mainly responsible for explaining the positional and scopal flexibility, and the discourse functionality of the items concerned.

3. Main Divergences in Different Mechanisms and the Root Causes

As can be seen from the basic claims of the different mechanisms for the evolution of discourse markers, there is both agreement and disagreement among them. Here we focus on examining the differences between various mechanisms and their underlying causes.

3.1 Main Differences

Regarding whether discourse markers result from grammaticalization or lexicalization, Brinton and Traugott (2005) note significant similarities between the two processes but maintain that discourse markers undergo grammaticalization rather than lexicalization. This is mainly based on the fact that

discourse markers do not belong to any separate lexical category and that the development of discourse markers is characterized by decategorialization, which is specific to grammaticalization. Scholars holding the lexicalization view, on the other hand, put forward that decategorialization occurs in both lexicalization and grammaticalization, and that discourse markers evolve with features unique to lexicalization such as fusion and semantic demotivation. Moreover, they argue that scope extension and syntactic freedom involved in the evolution of discourse markers are contradictory to condensation and fixation, which, according to Lehmann (1982), grammaticalization must undergo. In this regard, Brinton (2008) further suggests that both lexicalization and grammaticalization involve the fossilization of complex structures, but the structures that constitute the main lexical categories (verbs, nouns, adjectives etc.) undergo lexicalization, while the expressions that make up the functional categories undergo grammaticalization. As to some characteristics that do not conform to grammaticalization, Brinton maintains that “not even prototypical cases of grammaticalization (e.g., development of auxiliaries) will exhibit all of Lehmann’s parameters” (2008, p. 50).

Scholars holding the pragmaticalization view question grammaticalization in a number of ways. On the one hand, the evolution of discourse markers occurs only minimally with phonetic and semantic attrition, and does not exhibit paradigmaticization, obligatorification, condensation, coalescence or fixation characteristic of grammaticalization proposed by Lehmann (1982), and is therefore not a process of grammaticalization (Waltereit, 2002). This is basically consistent with the questioning from the perspective of lexicalization. On the other hand, the process of discourse marker formation embodies such pragmaticalization features as syntactic isolation, lack of coalescence, scope extension, non-truth-conditionality, optionality and unique grammatical identity (Aijmer, 1997; Frank-Job, 2006; Norde, 2009; Claridge & Arnovick, 2010). Moreover, semantic-pragmatic evolution occurs during the formation of discourse markers, which is also a feature of pragmaticalization rather than grammaticalization (Claridge & Arnovick, 2010; Van Bogaert, 2011). Of course, from a grammaticalization perspective, the aforementioned objections raised by the pragmaticalization view are not valid. Some scholars believe that grammatical constituents evolved via grammaticalization, including discourse markers, all express non-truth-conditional meanings, and all typically involve some pragmatic function; discourse markers, though carrying scope over more than the sentence, are undoubtedly “part of the grammar” (Traugott & Dasher, 2002, pp. 158-159).

It is thus clear that both lexicalization and pragmaticalization reject the notion that discourse markers are the result of grammaticalization, citing the absence of certain features of canonical grammaticalization in discourse markers. As to the cooptation view, in addition to contending that discourse markers do not exhibit the narrowly defined characteristics of grammaticalization, it puts particular emphasis on the spontaneity of their formation, rejecting the two fundamental features of graduality and unidirectionality that are inherent in the mechanisms of grammaticalization, lexicalization and pragmaticalization. In this regard, Heine (2013) argues that the current diachronic

corpus fails to demonstrate that the evolution of the relevant items follow a process from lower grammaticalization in earlier stages to higher grammaticalization in later stages. For example, *I mean* as a thetical occurred no later than its corresponding SG usage, both appearing almost simultaneously at the end of the 14th century; meanwhile, the sentence adverbial and discourse marker usages of *besides* emerged concurrently in the mid-16th century.

3.2 The Root of Disagreement

What, then, is the root of the divergence in the evolution mechanism of discourse markers? There are inevitably different viewpoints on the evolution mechanism based on different theories, but this is not the source of disagreement. The divergence mainly stems from the different understanding of some basic theories and concepts in the formation mechanism of discourse markers.

First of all, related research often switches between narrow and broad definitions when questioning or denying other theories. For example, the evolution of discourse markers is an instance of the broad view of grammaticalization (Traugott, 1995; Wang, 2005) or lexicalization (Dong, 2007). However, scholars advocating the broad grammaticalization view tend to deny the lexicalization of discourse markers by citing the features of narrow lexicalization, such as discourse markers not belonging to any major lexical category (Brinton & Traugott, 2005), while scholars holding the broad lexicalization view cite the features of narrow grammaticalization, such as discourse markers not being syntactically obligatory (Dong, 2007), to reject the grammaticalization of discourse markers. A question implied here is whether, for a given mechanism, the evolution of discourse markers needs to satisfy all its conditions.

Secondly, different scholars do not agree on the definitions of grammaticalization, pragmaticalization and lexicalization. Take the disagreement between pragmaticalization and grammaticalization as an example. Scholars such as Traugott (1995), Brinton (1996), Brinton and Traugott (2005) reject pragmaticalization on the grounds that both pragmaticalization and grammaticalization derive from pragmatic reasoning and share the characteristic of subjectification, thus incorporating the former into the latter. Aijmer believes in her earlier research that discourse markers are the result of pragmaticalization (1997), but later clarifies that grammaticalization and discourse markers are interdependent (2002), referring to pragmaticalization as discourse enrichment in grammaticalization. In addition, different scholars may refer to the same process of evolution in terms of grammaticalization and pragmaticalization respectively. Wang (2005), for instance, considers that the gradual loss of the literal meaning of many idiomatic expressions and fixed phrases in a language and the formation of a specific pragmatic function and conventionalized usage in a particular communicative context is grammaticalization in the broader sense, but Hou (2007) defines it as pragmaticalization. As to lexicalization and pragmaticalization, Dong (2010, p. 285) believes that the two can be compatible. Pragmaticalization focuses on the expressive function of linguistic forms. Discourse markers can be regarded as idioms due to the fact that their meanings cannot be inferred

from their forms, and therefore, the formation of discourse markers can be construed as the process of idiomatization of free phrases or clauses, which is actually the primary stage of lexicalization.

Moreover, unlike the disagreement between grammaticalization, lexicalization and pragmaticalization, the divergence between the cooptation view and the other three perspectives appears to be primarily empirical or evidence-based. The cooptation view argues that the diachronic corpus available does not prove that discourse markers evolve in a gradual and unidirectional manner, but is the non-graduality and non-unidirectionality reflected in the development of certain discourse markers to some extent related to the relative lack of diachronic data?

4. Interconnection and Intersection of Different Evolution Mechanisms

Theoretically speaking, grammaticalization, lexicalization, pragmaticalization and cooptation are four relatively independent frameworks, each with its own theoretical basis and norms, which interpret the evolution of discourse markers from different perspectives, and therefore are not necessarily mutually exclusive. For example, the motivations and mechanisms behind lexicalization and grammaticalization are largely consistent: both involve semantic demotivation, both are based on high-frequency usage, and both play a crucial role in the change of meaning through pragmatic reasoning (Brinton & Traugott, 2005; Dong, 2007). In addition, different perspectives overlap in identifying the typical features of discourse marker formation, meaning that a given feature functions not only within a single mechanism but also exists across two or more mechanisms. It is not difficult to discern from the relevant studies discussed earlier that, for example, fusion and decategorialization are common to grammaticalization and lexicalization, desemanticization is common to grammaticalization, pragmaticalization and lexicalization, subjectification is common to grammaticalization, pragmaticalization and lexicalization, and pragmatic strengthening and conventionalization are common to all four mechanisms.

Therefore, the formation of discourse markers may not be the result of a single mechanism acting alone, but rather the product of multiple mechanisms working together. Put another way, there may be interconnections and overlaps between different mechanisms. Scholars hold varying views on the roles of these mechanisms, primarily falling into the following categories.

Firstly, some scholars believe that more than one mechanism is at work, but the order or importance of the mechanisms may be different. Wischer (2000) examines the evolution of *methinks* in Middle English and argues that it is the result of the combined effects of lexicalization and grammaticalization: once *me þinkeð* is lexicalized into a lexical unit, it immediately assumes a grammatical function at the discourse level, that is, grammaticalization occurs. Similarly, Liu's (2007) study on the evolution of *suan le* and Li's (2012) on *bie shuo*, *wan le*, and *jiu shi* both propose that the formation process of discourse markers should involve lexicalization followed by grammaticalization. Moreover, Li argues that grammaticalization plays a major role, while lexicalization only contributes during the initial stages. The research by Xiang et al. (2016) on the evolution of *sort of* and *I think* demonstrates that

pragmaticalization and grammaticalization are not absolutely mutually exclusive, but rather interconnected complementarily in different stages in the development of the two markers. In the case of *sort of*, it should be first lexicalized and then grammaticalized, and this evolutionary pattern is also observed in the development of other typical pragmatic markers.

Secondly, although some scholars tend to explain the evolution of discourse markers in terms of a certain mechanism, they do not deny the role of other mechanisms. Dong (2007), for instance, advocates exploring the formation of Chinese discourse markers from the perspective of lexicalization, but she also makes it clear that discourse markers can be the product of lexicalization, as in *bie shuo*, and also the product of grammaticalization, as in *hao* used to start a new topic, or the result of the successive effects of lexicalization and grammaticalization, as in *shui zhi dao*. Claridge and Arnovick (2010) argue that pragmaticalization is more appropriate for describing the formation of discourse markers, but Claridge's (2013) examination of the evolution of *as it were*, *so to speak/say* and *if you like* also demonstrates that: the changes involved are seen as instances of pragmaticalization and partly also lexicalization, although they share many features with grammaticalization.

Thirdly, some scholars believe that a certain mechanism is bound to work, but there is uncertainty about the role of other mechanisms. Yin (2012), through the examination of Chinese discourse markers *jiu shi*, *wan le*, *na me*, *wo/ni kan*, etc., proposes that the formation of discourse markers generally involves pragmaticalization, while whether they undergo lexicalization or grammaticalization varies depending on the specific discourse marker. Similarly, as mentioned earlier, Heine (2013, p. 1234) argues that "it is cooptation, rather than grammaticalization, that is prerequisite for the rise of DMs", and most coopted units will never grammaticalize. This, however, does not deny that grammaticalization may indeed play a role in the evolution of certain discourse markers, which is likely to take place prior to undergoing cooptation (working on units of SG) or after cooptation (changing the instantaneous coopted units into formulaic theticals, i.e. discourse markers).

Additionally, some scholars hold the view that the formation of discourse markers does not represent the canonical process of grammaticalization, lexicalization, or pragmaticalization. As Li (2010) proposes, the transformation of *wo shuo*-type phrases or clauses into discourse markers is neither a traditional grammaticalization process nor a typical lexicalization process. The formation of discourse markers, therefore, cannot be reduced to standard grammaticalization, nor is it the same as the lexicalization of other ordinary lexical items. Similarly, Barth-Weingarten and Couper-Kuhlen (2002) argue that discourse markers are neither solely the result of grammaticalization, as commonly assumed, nor entirely the result of pragmaticalization; their formation involves characteristics of both grammaticalization and pragmaticalization.

5. Suggestions for Future Research

To sum up, research on the evolution of discourse markers yields abundant findings, with both consensus and controversy, reflecting that the formation of discourse markers is a complex process. There is still much room for development in this field, which is worth further exploration.

The first is to explore the formation mechanisms of discourse markers from a wider range of dimensions. The ultimate goal of research on the evolution of discourse markers is not merely to trace the developmental trajectory of a particular item, but to uncover the mechanisms driving such evolution. There should be a clearer understanding and definition of the various formation mechanisms currently debated, and some fundamental issues require resolution (Liu, 2010; Degand & Evers-Vermeul, 2015; Zhang, 2019). For example, is it possible to reach a consensus on basic concepts such as grammar, lexicon, lexical units, grammaticality, lexicality etc.? Regarding the evolution of discourse markers, is it necessary to satisfy all conditions of each mechanism? Are there core parameters for grammaticalization, lexicalization, pragmaticalization and cooptation respectively, and is it possible for these mechanisms to form an integrated model or unified framework based on such parameters?

Given that different types of discourse markers follow distinct evolutionary trajectories, their formation mechanisms may also differ (e.g., Li, 2012, 2023; Davidse et al., 2015; Zhong, 2017). Therefore, can we discuss different formation mechanisms based on the type of discourse marker? The difficulty here may lie in determining the classification criteria. For instance, one might consider not only the classification based on the source structure (including words, phrases and clauses), but also further clarification of the subdivision of these categories. In this case, it is particularly important to avoid over-simplifying or over-complicating the classification.

Finally, attention should also be paid to cross-linguistic comparative research. Current studies on the evolution of discourse markers primarily focus on individual languages, lacking comparative analysis across different linguistic systems. In fact, through the diachronic comparative study of discourse markers in English, Chinese, and other languages (such as research objects with similar functions or similar source structures), we can delve into whether the evolution of discourse markers exhibits cross-linguistic commonalities, thereby testing the theory of discourse marker evolution in a more general sense.

References

Aijmer, K. (1997). *I think*—An English modal particle. In T. Swan, & O. J. Westvik (Eds.), *Modality in Germanic languages: Historical and comparative perspectives* (pp. 1-47). Berlin/New York: Mouton de Gruyter.

Aijmer, K. (2002). *English discourse particles: Evidence from a corpus*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Barth-Weingarten, D., & Couper-Kuhlen, E. (2002). On the development of final thought: A case of grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New reflections on grammaticalization*. Amsterdam & Philadelphia: John Benjamins.

Brems, L., & Davidse, K. (2010). The grammaticalization of nominal type noun constructions with *kind/ sort of*: chronology and paths of change. *English Studies*, 91(2), 180-202.

Brinton, L. J. (1996). *Pragmatic markers in English*. Berlin: Mouton de Gruyter.

Brinton, L. J. (2008). *The comment clause in English*. Cambridge: Cambridge University Press.

Brinton, L. J. (2011). The grammaticalization of complex predicates. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.

Brinton, L. J., & Traugott, E. C. (2005). *Lexicalization and language change*. Cambridge: Cambridge University Press.

Claridge, C. (2013). The evolution of three pragmatic markers: *As it were, so to speak/say* and *if you like*. *Journal of Historical Pragmatics*, 14(2), 161-184.

Claridge, C., & Arnovick, L. (2010). Pragmaticalisation and discursisation. In H. Andreas, & I. T. Jucker (Eds.), *Historical pragmatics*. Berlin: Mouton de Gruyter.

Davidse, K., De Wolf, S., & Van Linden, A. (2015). The development of modal and discourse marker uses of *(there/it is/I have) no doubt*. *Journal of Historical Pragmatics*, 16(1), 25-58.

Degand, L., & Evers-Vermeul, J. (2015). Grammaticalization or pragmaticalization of discourse markers?: More than a terminological issue. *Journal of Historical Pragmatic*, 16(1), 59-85.

Dong, X. F. (2007). Lexicalization and the origin of discourse markers. *Chinese Teaching in the World*, (1), 50-61.

Dong, X. F. (2010). A discourse marker derived from clausal form: *wo gao su ni*. *Linguistic Sciences*, 9(3), 279-286.

Erman, B., & Kotsinas, U.-B. (1993). Pragmaticalization: The case of *ba'* and *you know*. *Studies i modern Språkvetenskap* (Acta Universitatis Stockholmiensis, Stockholm Studies in Modern Philology, New Series 10) (pp. 76-93). Stockholm: Almqvist and Wiksell.

Fischer, O. (2007). The development of English parentheticals: A case of grammaticalization? In U. Smit, S. Dollinger, J. Hüttnner, G. Kaltenböck, & U. Lutzky (Eds.), *Tracing English through time: Explorations in language variation*. A Festschrift for Herbert Schendl on the Occasion for his 65th Birthday (Austrian Studies in English, 95) (pp. 103-118). Vienna: Braumuller.

Frank-Job, B. (2006). A dynamic-interactional approach to discourse markers. In K. Fischer (Ed.), *Approaches to discourse particles* (Studies in Pragmatics, 1) (pp. 359-374). Amsterdam: Elsevier.

Heine, B. (2013). On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics*, 51(6), 1205-1247.

Heine , B., Kaltenböck, G., Kuteva, T., & Long, H. P. (2013). An outline of discourse grammar. In Sh. Bischoff, & J. Carmen (Eds.), *Functional approaches to language* (pp. 175-233). Berlin: Mouton

de Gruyter.

Heine, B., Kaltenböck, G., Kuteva, T., & Long, H. P. (2017). Cooptation as a discourse strategy. *Linguistics*, 55(4), 813-855.

Heine, B., Kaltenböck, G., Kuteva, T., & Long, H. P. (2021). *The rise of discourse markers*. Cambridge: Cambridge University Press.

Hopper, P. J. (1991). On some principles of grammaticalization. In E. C. Traugott, & B. Heine (Eds.), *Approaches to grammaticalization* (Vol. 1, pp. 17-35). Amsterdam/Philadelphia: John Benjamins.

Hopper, P. J., & Traugott, E. C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Hou, G. J. (2007). Grammaticalization and pragmaticalization: Investigated from perspectives of cognitive-pragmatics and pragma-translatology. *Journal of Chongqing University*, (5), 117-124.

Ifantidou, E. (2001). *Evidentials and relevance* (Pragmatics & Beyond, IV: 6). Amsterdam/Philadelphia: John Benjamins.

Jucker, A. H. (1997). The discourse marker *well* in the history of English. *English Language and Linguistics*, 1(1), 91-110.

Kaltenböck, G., Heine, B., & Kuteva, T. (2011). On thetical grammar. *Studies in Language*, 35(4), 852-897.

Kaltenböck, G., & Heine, B. (2014). Sentence grammar vs. thetical grammar: Two competing domains. In M. Brian, A. Malchukov, & E. Moravcsik (Eds.), *Competing motivations in grammar and usage* (pp. 348-363). Oxford: Oxford University Press.

Krug, M. (1998). British English is developing a new discourse marker, *innit*? A study in lexicalization based on social, regional and stylistic variation. *Arbeiten aus Anglistik und Amerikanistik*, (23), 145-197.

Lehmann, C. (1982). *Thoughts on grammaticalization* (LINCOM Studies in Theoretical Linguistics, 01.) Revised and expanded version. Munchen and Newcastle: Lincom Europa.

Li, S. X. (2012). Formation of discourse markers from lexicalization and grammaticalization and the problem of the origin of discourse markers. *Chinese Teaching in the World*, (26), 322-337.

Li, S. X. (2023). Formation mechanisms of discourse markers and the frontier topics. *Journal of Foreign Languages*, (3), 70-81.

Li, Z. J. (2010). On the origin of discourse markers with *wo shuo* as an example. *Chinese Teaching in the World*, (2), 192-198.

Liu, H. N. (2007). On the lexicalization and grammaticalization of the non-syntactic structure *suan le*. *Linguistic Sciences*, (6), 11-21.

Long, H. P., & Wang G. (2014). On the formation of thetical *shi de* from the perspective of cooptation. *Chinese Teaching in the World*, (1), 78-87.

Norde, M. (2009). *Degrammaticalization*. Oxford: Oxford University Press.

Qiu, S. D., & Sun, Q. (2011). Pragmaticalization and pragmatic markers. *Foreign Languages in China*, (41), 30-37.

Stenström, A. B. (1998). From sentence to discourse: *cos (because)* in teenage talk. In A. H. Jucker, & Z. Yael (Eds.), *Discourse markers: Descriptions and theory* (pp. 127-146). Amsterdam: John Benjamins.

Tabor, W., & Traugott, E. C. (1998). Structural scope expansion and grammaticalization. In A. G. Ramat, & P. J. Hopper (Eds.), *The limits of grammaticalization* (pp. 229-272). Amsterdam/Philadelphia: John Benjamins.

Traugott, E. C. (1982). From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In W. P. Lehmann, & Y. Malkiel (Eds.), *Perspectives on historical linguistics* (pp. 245-271). Amsterdam/Philadelphia: John Benjamins.

Traugott, E. C. (1995). *The role of discourse markers in a theory of grammaticalization*. Paper presented at the 12th International Conference on Historical Linguistics, Manchester.

Traugott, E. C., & Dasher, R. B. (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.

Van Bogaert, J. (2011). *I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics*, 49(2), 295-332.

Waltereit, R. (2002). Imperatives, interruption in conversation and the rise of discourse markers: A study of Italian *guarda*. *Linguistics*, 40(5), 987-1010.

Wang, Y. (2005). Narrow-sensed and broad-sensed grammaticalization. *Journal of Sichuan International Studies University*, (5), 68-73.

Wischer, I. (2000). Grammaticalization versus lexicalization: “Methinks” there is some confusion. In O. Fischer, A. Rosenbach, & D. Stein (Eds.), *Pathways of change. Grammaticalization in English* (pp. 354-370). Amsterdam/Philadelphia: John Benjamins.

Xiang, M. Y., Lu, Z. Y., & Yang, G. P. (2016). Notes on the controversies over pragmaticalization and grammaticalization. *Modern Foreign Languages*, (2), 158-168.

Yin, S. L. (2012). On evolution of discourse markers. *Journal of Hunan University of Science and Technology*, (2), 133-138.

Zhang, X. S. (2019). The debate on the nature of the development of discourse markers. *Foreign Language Research*, (4), 43-50.

Zhong, X. Y. (2017). Theoretical discussions on the evolution of discourse markers: State of the art and evaluation. *Journal of Yantai University*, 30(6), 114-121.