The Evolution of $ɦi^{23}kã^{34}$ ('he says') in Shanghainese*

Weifeng Han and Dingxu Shi

The Hong Kong Polytechnic University

This paper presents a novel analysis of the grammaticalization path of the various uses of $ɦi^{23}kã^{34}$ ('he says') in modern Shanghainese. We propose that $ɦi^{23}kã^{34}$ originated from a quotative clause, then developed its reportative use by making the source of information obscure (in the sentence-final position), and finally integrated with the proposition to form a modal particle indicative of the speaker's subjective evaluation. Recently, $ɦi^{23}kã^{34}$ has been raised to topic position to serve as a topic-marking function; specifically, one that expresses a counter-expectation meaning. The whole process can be chronologized as follows: quotative clause > reportative clause > modal particle > topic-marking particle. Such a path is considered to be elegant and practical since not only do the stages involved gracefully follow the principles of grammaticalization as delineated by Hopper (1991, 1996), but it is also helpful for the analysis of similar patterns across languages.

Key words: $ɦi^{23}kã^{34}$, grammaticalization, modal particle, Shanghainese, topic-marking particle

1. Introduction

$ɦi^{23}kã^{34}$ (伊講), or 'he says' in Shanghainese, is basically used as a main clause introducing (in)direct speech, as shown in (1). The quotative clause can be "moved" to the right of the quoted clause, as evidenced by (2).

(1)　　$ɦi^{23}$　$kã^{34}$　　　$və^{ʔ12}hɔ^{34}$
　　　　he　says　　　no
　　　　'He says NO!'

(2)　　$və^{ʔ12}hɔ^{34}$,　　$ɦi^{23}$　　$kã^{34}$
　　　　no　　　　　he　　says
　　　　'He says NO!'

The "right-dislocated" $ɦi^{23}kã^{34}$ is, more often than not, used to mark the speaker's subjective evaluation of the proposition in modern Shanghainese (see Tao & Li 2009), though without changing the proposition's truth value, as illustrated in (3).[1] Such an "evaluative" $ɦi^{23}kã^{34}$,

---

[1]　The following abbreviations are used in this paper: ASP = aspect particle; MOD = modal particle; PRT = particle; QOT = quotative clause; REP = reportative clause; TOP = topic particle.

however, can never be "moved back" to the sentence-initial position,[2] and thus is not a simple "dislocated" element.

(3) $və^{ʔ12}hɔ^{34}$ $ɦi^{23}kã^{34}$
     no   PRT
     'Unexpectedly, there is a NO!'

Tao & Li's paper (2009) was the first—and, to date, the only one—to pay particular attention to the evaluative use of $ɦi^{23}kã^{34}$. How this evaluative use "changed" from its original quotative usage, however, is never mentioned. Moreover, the fact that $ɦi^{23}kã^{34}$ can be inserted as a topic marker in the middle of a clause,[3] as shown in (4), has also been somehow neglected in the past literature.[4]

(4) $ɦi^{23}$ $ɦi^{23}kã^{34}$, $kã^{34}$ $və^{ʔ12}hɔ^{34}$
     he   TOP   says   no
     'Unexpectedly, he says NO!'

This paper sets out to sketch the grammaticalization path of $ɦi^{23}kã^{34}$. The proposal is that there is a transitional reportative use between the quotative and evaluative $ɦi^{23}kã^{34}$, from which the new topic marking $ɦi^{23}kã^{34}$ is derived. The path is believed to be simple and elegant, since it develops through a series of typical stages of grammaticalization, which are cross-linguistically attested (e.g. Aikhenvald 2004; Suzuki 2007; Ahn & Yap 2012; Tamaji & Yap 2013, etc.), as has been proposed for the Japanese *tte* in (5):

(5) *ashita*   *ame*   *ga*   *huru*   *tte*
     tomorrow   rain   NOM   fall   PRT
     a.   'He said(QUOTATIVE), "it will rain tomorrow."'
     b.   'It is said(REPORTATIVE) it will rain tomorrow.'
     (adapted from Yap & Tamaji 2012)

---

[2] Rather, the sentence-final evaluative $ɦi^{23}kã^{34}$ can co-occur with the sentence-initial quotative $ɦi^{23}kã^{34}$, as attested in (i):

(i) $ɦi^{23}$ $kã^{34}$ $və^{ʔ12}hɔ^{34}$ $ɦi^{23}kã^{34}$
     he   says   no   PRT
     'Unexpectedly, he says NO (rather than YES)!'

[3] The clause-medial use of $ɦi^{23}kã^{34}$ is indeed new, and probably emerged in the late 1900s. Unfortunately, the *Shanghai Spoken Corpus* (hereafter referred to as *SSC*) (Han et al. 2008–2012) is the only written material (in the authors' hands) appropriate for academic research, from which examples of such a topic-marking use of $ɦi^{23}kã^{34}$ are extracted. Further details and an earlier, untagged version of the *SSC* can be accessed via http://www.ualberta.ca/~johnnewm/SC/Shanghai/SSC.html.

[4] $ɦi^{23}kã^{34}$ in (4) not only occupies the conventional topic-marking position, the COUNTER-EXPECTATION meaning it triggers is also one of the five typical functions of topic markers in Shanghainese (see Han et al. 2013 for more details).

This paper is organized as follows. Section 2 examines the reportative $\hbar i^{23} k\tilde{a}^{34}$ and analyzes how it links the quotative usage of $\hbar i^{23} k\tilde{a}^{34}$ and the evaluative one. The function and motivation of the evaluative $\hbar i^{23} k\tilde{a}^{34}$ are also discussed. Section 3 provides a first approximation of the grammaticalization of $\hbar i^{23} k\tilde{a}^{34}$. Section 4 focuses on the more recent emergence of the topic-marking $\hbar i^{23} k\tilde{a}^{34}$ in modern Shanghainese, especially its syntactic position and pragmatic function. Section 5 follows up with a second approximation of the complete path of the grammaticalization of $\hbar i^{23} k\tilde{a}^{34}$ from quotative construction to topic marker. Finally, §6 concludes the paper.

2. Quotative, reportative, and evaluative $\hbar i^{23} k\tilde{a}^{34}$

2.1 Quotative versus reportative $\hbar i^{23} k\tilde{a}^{34}$

In many languages, verbs of saying often grammaticalize into quotation markers (with the function of hearsay evidentials); for example, Cahuilla *yáxqal* 'he says' in (6)[5] (see Seiler 1970: 1–2, 1977:187).

(6) *ʔáy péqi níyaxal ʔé péqi ʔetyáxik pen níyaxqal*   **yáxqal**   *táxat, múkat*
     'What I say is just what you would say'   **QOT**   said Múkat.
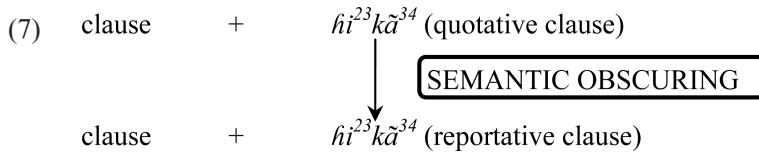
As we can see, *yáxqal* in (6) indicates the source of information by directly introducing the source *Múkat*. That is to say, the quotative clause itself (i.e. *yáxqal* 's/he says'), used as a quotation marker, contains no leads to a particular source of information. Although *yáxqal* in Cahuilla is bound to a definite source—in this case, *Múkat*—it is not uncommon, cross-linguistically, for a quotative clause to evolve into a reportative clause (with an obscure source) in the sentence-final position, as in the case of *mere* 's/he says' in Lepcha, which is employed "to indicate that the information did not come to the speaker's knowledge through direct information or through inference", and:

> When a speaker chooses to use the reportative clause, this highlights the fact that the information is based on something someone said, and there is an undertone that suggests that the speaker cannot necessarily vouch for the information or is unsure whether the information is necessarily accurate. (Plaisier 2007:37–38)

It may involve semantic obscuring of the corresponding pronoun as a definite source of information becoming indefinite (Han & Shi 2012), in other words, in the "he says" situation, "he" is obscured as "one" or "someone" to indicate the content reported as hearsay information (Heine &

---

[5] The grammaticalization process of a verb of saying into a quotative marker is seen, typologically, throughout many languages. Here are some other examples, among many others: *mĩ* and *tĩ* in Nama (Hagman 1977:137; Krönlein 1889:231, 309); *se* in Twi (Lord 1989:292ff.); *ye* in Kusasi (Kusal) (Lord 1993:198–199); *taá* in Saramaccan CE and *taki* in Sranan CE (Lord 1989:335); *sey* in West African PE (Lord 1989:332); *Áro* and *fó* in Vei (Koelle 1968:122, 134); *prepa* in Buru (Klamer 2000:76), etc.

Kuteva 2002:265; see also *luhuda* 'one says' in Lezgian Haspelmath 1993:232). Therefore, the use of 'he says' as a reportative clause with an indefinite source of information is to show the speaker's uncertainty about (and sometimes even refusal to accept) the information reported (rather than to change its truth value). It could be argued that, by being "dislocated," $\hbar i^{23} k \tilde{a}^{34}$ prepares itself to become a reportative clause, with an obscure source implying the speaker's doubt regarding the information, through semantic obscuring. Therefore, we can look at the sentence-final reportative $\hbar i^{23} k \tilde{a}^{34}$ as a marker with "quasi-evaluative" functions,[6] since the evaluation is rather subtle and implicit. The grammaticalization process can be described as per (7):

(7)   clause   + $\hbar i^{23} k \tilde{a}^{34}$ (quotative clause)

$$\downarrow \boxed{\text{SEMANTIC OBSCURING}}$$

       clause   + $\hbar i^{23} k \tilde{a}^{34}$ (reportative clause)

2.2 Evaluative $\hbar i^{23} k \tilde{a}^{34}$

Compared to the reportative $\hbar i^{23} k \tilde{a}^{34}$ (where the definite source of information is made obscure), there is no such source (be it a clear "he" or an obscure "someone"), nor is there an action of SAYING, found in the evaluative $\hbar i^{23} k \tilde{a}^{34}$ in (3), earlier. On the other hand, though the truth of the proposition remains unchanged, the speaker's evaluation on the counter-expectation proposition in (3) is unmistakably expressed.

In Shanghainese, as in many other languages, a *wh*-question such as $sa^{\Omega55} nin^{23} k \tilde{a}^{34} a^{\Omega55}$ (啥人講呢) 'who says' is used to question the source of quotation or report, as shown in (8a). The interesting thing is that $sa^{\Omega55} nin^{23} k \tilde{a}^{34} a^{\Omega55}$ in Shanghainese can also be used to question the non-quotative/non-reportative $\hbar i^{23} k \tilde{a}^{34}$, as interpreted in (8b).

(8) ---$\hbar o^{23} g \partial^{\Omega12}$ $li^{23} pa^{34}$ $\mathit{tc} y^{34} fa^{\Omega55}$ $\hbar i y^{23}$ $i \mathfrak{o}^{34}$ $k^h \mathfrak{o}^{34} s \mathit{1}^{34}(,)$ $\hbar i^{23}$ $k \tilde{a}^{34}$
          next   week   syntax   again   will   exam   he   say
     a.   '"There will be another exam on syntax next week," he says.'
     b.   'Unexpectedly, there will be another exam on syntax next week!'

     ---$sa^{\Omega55} nin^{23}$ $k \tilde{a}^{34}$ $a^{\Omega55}$
          who   say   PRT
     a.   'Who says that?'
     b.   'It's so unexpected!'

---

[6] A sentence-final reportative clause (with an obscure source of information) could be reanalyzed as a phrase of hearsay evidential (Itani 1991, 1994; Suzuki 2007, etc.), which adds counter-expectation meaning to the content reported (see Yap & Tamaji 2012), while a hearsay particle arguably indicates that the speaker shows less commitment to the truth of the proposition (Chafe 1986; Palmer 1986).

Obviously, $sa^{?55}\text{\textltailn}in^{23}k\tilde{a}^{34}a^{?55}$ in (8b) does not really question the source of the examination news; rather, it more or less echoes the speaker's counter-expectation on the proposition, that is, $\hbar i^{23}k\tilde{a}^{34}$, as used in (3) and (8b), expresses evaluative modality, and therefore is a modal particle with an evidential marking function (rather than a quotative or reportative clause).

Modal categories (specifically MOOD),[7] which are expressed in either verbal morphology (Palmer 1986:21) or independent words (verbs, auxiliaries, or *particles*) (Cinque 1999:78), mostly deal with the speaker's opinion or attitude toward the proposition (Lyons 1977:432). Cinque (1999:84) also points out that "across languages, such evaluative modalities … do not affect the truth of the proposition, but rather express the speaker's (positive, negative, or other) evaluation of the state of affairs described in it."

Before $\hbar i^{23}k\tilde{a}^{34}$, as Qian (1996) points out, there had been three traditional particles expressing counter-expectation modality in Shanghainese, namely $g\partial^{?}m\partial^{?}$ (個末), $ge^{?}\hbar\mathfrak{o}$ (個嚜), and $g\partial^{?}\gamma$ (個鴨).[8] Sometimes they are used together with $t\mathfrak{o}$ (倒), which is a typical topic marker in Shanghainese, to reinforce the speaker's evaluation of the proposition, as exemplified by (9)–(11):

(9) $no\eta^{23}$ $t\mathfrak{o}^{34}$ $m\varepsilon^{53}$ $\mathbf{ts}^{h}\mathfrak{o}^{53}\varsigma in^{53}$ $g\partial^{?}m\partial^{?}$
     you   TOP   very   concerned   MOD
     'You are really concerned!'

(10) $s\mathfrak{1}^{34}$ $k^{h}\varepsilon^{53}$ $g\partial^{?}\gamma$
      water   boil   MOD
      'It's boiling water (so don't touch it)!'

(11) $m\varepsilon^{53}$ $h\mathfrak{o}^{34}k^{h}\emptyset^{34}$ $ge^{?}\hbar\mathfrak{o}$
      very   pretty   MOD
      'It's really pretty!'

None of these three modal particles can be replaced by $\hbar i^{23}k\tilde{a}^{34}$, however, as can be seen in (12)–(14):

(12) *$no\eta^{23}$ $t\mathfrak{o}^{34}$ $m\varepsilon^{53}$ $\mathbf{ts}^{h}\mathfrak{o}^{53}\varsigma in^{53}$ $\hbar i^{23}k\tilde{a}^{34}$

(13) *$s\mathfrak{1}^{34}$ $k^{h}\varepsilon^{53}$ $\hbar i^{23}k\tilde{a}^{34}$

(14) *$m\varepsilon^{53}$ $h\mathfrak{o}^{34}k^{h}\emptyset^{34}$ $\hbar i^{23}k\tilde{a}^{34}$

In fact, although the three modal particles express counter-expectation, $g\partial^{?}\gamma$ suggests a warning to the hearer, while $g\partial^{?}m\partial^{?}$ and $ge^{?}\hbar\mathfrak{o}$ mostly confirm the proposition (though it is still unexpected).

---

[7] MOOD and MODALITY are treated as basically the same in this paper (see Cinque 1999 and Palmer 1986 for motivations).

[8] There might be alternative analyses of these SFPs. Other interpretations (if any exist), however, are not related to the current discussion. In addition, the transcriptions of $m\partial^{?}$, $\hbar\mathfrak{o}$ and γ follow Qian (1996).

$ɦi^{23}kã^{34}$, on the other hand, focuses on how the proposition sounds unbelievable, and how the speaker tries to reject the truth. Therefore, it is only the $ge^{ʔ}ɦɔ$ in (15b), below, without confirmation of the proposition, rather than that in (15a),[9] that can be replaced by $ɦi^{23}kã^{34}$:

(15) $ɦi^{23}$ $tɔ^{34}$ $ɦuɛ^{23}$ $lɛ^{23}$ $ge^{ʔ}ɦɔ$
 he TOP still come MOD (confirmation)
 a. 'He really comes!' ($≠ɦi^{23}tɔ^{34}ɦuɛ^{23}lɛ^{23}$ $ɦi^{23}kã^{34}$)
 b. 'Unexpectedly, he comes!' ($=ɦi^{23}tɔ^{34}ɦuɛ^{23}lɛ^{23}ɦi^{23}kã^{34}$)

Intralinguistically, before $ɦi^{23}kã^{34}$ there were no such particles in Shanghainese used exclusively to express a counter-expectation mood involving denial, so $ɦi^{23}kã^{34}$ came about to fulfill these functional needs within the language (see also Tao & Li 2009), more or less in the same way that 'I think' or 'I guess' in English grammaticalized into particles of evidentiality (see Hopper & Traugott 2003:201–203).[10]

3. The grammaticalization of $ɦi^{23}kã^{34}$

The previous section implies that the evaluative $ɦi^{23}kã^{34}$ emerged after its reportative use. The question is what mechanism is involved in the re-analysis of the reportative clause into a modal particle.

Separately, both 'he' and 'say' are characteristic candidates to express MOOD in the sentence-final position. For example, $ɦi^{23}$ 'he' in Shanghainese carries a feature of [+Subjectivity] and is typically used to express mood of exaggeration, as illustrated in (16). It is also found that a pathway for the emergence of sentence-final particles in Chinese involves the grammaticalization of verbs of saying, as evidenced in the Taiwanese *kong* (Simpson & Wu 2002), Mandarin *shuo* (Wang et al. 2003), and Cantonese *waa* (Chui 1994; Leung 2006; Yeung 2006).

(16) $ɕiɪ^{ʔ55}ɦo^{23}lɛ^{23}$ $mə^{ʔ}$, $liɪ^{ʔ12}$ $liɪ^{ʔ12}$ $ɦi^{23}$
 rest TOP stand stand SFP [+Subjectivity]
 '(You should always) stand up (and relax yourself), when not occupied (with work)!'

In the case of the Shanghainese $ɦi^{23}kã^{34}$, it is proposed that an alternative pathway, which involves the integration of evaluative clauses as sentence-final mood particles of the preceding clause (see Wang & Yap 2009; Yap et al. 2009), is followed.[11] Yap et al. (2010) suggest that clausal

---

9 Indeed, which meaning is assigned is largely dependent on the context of situation.
10 Typologically, $ɦi^{23}kã^{34}$ belongs to Evidential Marking (Type II) particles (cf. Indirectivity Marking (Type I) particles; see Aikhenvald 2003 and 2004 for details), such as *ronki* in the Shipibo language (Valenzuela 2003:39). Additionally, it is also typical for an evidential-marking particle to express both evidentiality and mirativity (i.e. unexpected information), such as *lą̄ą́* in Western Apache (also see Aikhenvald 2003, 2004).
11 Some studies (e.g. Leung 2006; Yap & Ahn 2012; Yap et al. 2014), suggest that the Cantonese bare verb of saying *waa*⁶ in sentence-final position has also undergone clausal integration, however, in the different forms of *wo*³, *wo*⁵, etc.

integration is a simple and robust mechanism, productive in creating modal particles that capture subtle nuances of a speaker's mood or stance, which, therefore, gives rise to syntactic re-analysis. Such (re-)analysis, which is found cross-linguistically (e.g. in ancient and modern Chinese, Austronesian languages, etc.), indicates that the modal particle can function as an utterance tag following a propositional clause to convey meanings such as "is it [=the situation in the prior clause] thus?", or "is it [=the situation in the prior clause] true?" (Yap et al. 2010). (17) is an example of *bàle* (罷了) (from *XiYouJi* "西遊記") used as a sentence-final modal particle, while (18) describes the corresponding process of clausal integration.[12]

(17)　ràng　tā　mànman　zǒu　bàle
　　　let　he　slow　walk　MOD
　　　'**Just** let him walk slowly!'

(18)　clause$_1$　+　clause$_2$
　　　　　　　　　bà　*le*
　　　　　　　　　finish　ASP
　　　　　　　　　'That's it (I should give up)!'
　　　clause$_1$　+　MOD
　　　　　　　　　bàle
　　　　　　　　　'just'
　　　(adapted from Yap et al. 2010)

Now, let us postulate that *ɦi²³kã³⁴* first evolved from a quotative clause into a reportative clause in the sentence-final position, indicating the speaker's doubt of the proposition, through SEMANTIC OBSCURING (of the source of information). It is the reportative *ɦi²³kã³⁴* that is CLAUSALLY INTERGRATED to its preceding clause as an evaluative particle, conveying the speaker's mood of refusal and counter-expectation to the proposition. The grammaticalization path is summarized in (19):

(19)　Stage 1:　clause　+　*ɦi²³kã³⁴* (quotative clause)
　　　　　　　　　　　　　　　　　　↓　SEMANTIC OBSCURING
　　　Stage 2:　clause　+　*ɦi²³kã³⁴* (reportative clause)
　　　　　　　　　　　　　　　　　　↓　CLAUSAL INTEGRATION
　　　Stage 3:　clause　+　*ɦi²³kã³⁴* (evaluative particle)

---

[12] The use of *bàle* in sentence-final position is quite similar to *yiyi* (已矣) in old Chinese, while they are developed through the same clausal integration strategy. *Bàle* is a construction of a verb meaning 'stop' and a perfective particle (as *yi* 已 and *yi* 矣). Such an evaluative "terminal" clause is thus integrated into the preceding clause as a sentence-final particle, a process triggered by a strong desire for expressivity as well as cognitive pressures to routinize the evaluative expression as a pragmatic marker. (See Yap et al. 2010 for details.)

As pointed out by a reviewer of this paper, the evolution of $ɦi^{23}kã^{34}$ is not only a process of grammaticalization but could also be a typical case of lexicalization, which is usually considered a necessary step for the word cluster before changing into its new status as a function word. Intuitively, it is possible that $ɦi^{23}kã^{34}$, starting from a matrix clause, first became a phrase, then lexicalized into a compound word—a process (to form words from "freely combined" syntactic elements; see Dong 2002) not uncommon in Chinese. Lexicalization takes place due to the fact that lexicalized items "can be more easily used as elements of sentences" (Quirk et al. 1985:1525; Sauer 2004). Unlike typical lexicalized phrases in Chinese, such as *xuewen* (學問) and *sikao* (思考) (Dong 2002), the Shanghainese $ɦi^{23}kã^{34}$ keeps very few "semantic traits" (see Cruse 1986:22) of its previous uses, in that the structure is morphologically transparent but lacks "semantic compositionality" (Bauer 1983:55–59), or is "semantically opaque" (Anttila 1989:151). In other words, the meaning of $ɦi^{23}kã^{34}$ is quite specific but unpredictable (according to its individual elements) (see also Lehmann 2002:14; Lipka 1992:7). The fact that $ɦi^{23}kã^{34}$ is not able to function independently and can only appear in the fixed syntactic positions as exemplified in this paper shows that it is indeed not as "easily used" as the others. Therefore, the lexicalization of $ɦi^{23}kã^{34}$ is rather transitional, just an intermediate stage of grammaticalization.

4. Topic marking $ɦi^{23}kã^{34}$

As shown in (4), $ɦi^{23}kã^{34}$ in Shanghainese, however, can also land in a post-topic position in the clause, rather than the sentence-initial or sentence-final position. Here are two further examples from the *SSC* (Han et al. 2008–2012):

(20) $mən^{23}kʰɤ^{34}dɤ^{23}gə^{ʔ12}$ $mo^{ʔ12}ʑiã^{23}$ $ɦi^{23}kã^{34}$, $ɦuɛ^{23}tə^{ʔ55}$ RAP
     outside   carpenter   ?   can   RAP
     '(It's unbelievable that) the carpenter outside can do RAP!'

(21) $tɕʰi^{34}$ $sɛ^{53}gə^{ʔ12}ɦɔ^{23}dɤ^{23}$ $ɦi^{23}kã^{34}$, $gə^{ʔ12}nən^{23}ɦiã^{23}tsɿ^{34}$ $ɦioŋ^{23}$ $tsʰɔ^{53}pʰiɔ^{34}$
     go   three months   ?   this way   spend   money
     '(It's unbelievable that) (all the) money has been used up in three months!'

This $ɦi^{23}kã^{34}$ also has a counter-expectation meaning (Han & Shi 2012), which demonstrates the speaker's reluctance to accept the proposition. However, it does not occupy the sentence-final position, nor is it in the so-called "SFP_2" position for modal particles (cf. Law 2002).[13]

---

[13] Law (2002) proposes a lower SFP_2 position (right nearside the topic in the CP domain) that also hosts particles of MOOD/FORCE, as in (i):

(i) SFP_1→TopP→SFP_2→FocP→TopP

$ɦi^{23}kã^{34}$ in (4), (20), and (21), however, does not fit in well with such an independent SFP_2 position. The reason, among many others, is that a $ɦi^{23}kã^{34}$ in SFP_2 always prohibits a topic marker (whether a comma or a lexical marker) in its previous TopP, which is well predicted by (i), as exemplified by (ii) and (iii):

Shanghainese, as a typical topic-prominent dialect (Han 2013; Liu 2001), is abundant with topic-marking devices, some of which are used to express counter-expectation meanings; for example, $t\mathfrak{o}$ (Qian 1996; Xu & Liu 2007), ne (呢), a (啊), $m\mathfrak{o}$ (末) (Han et al. 2013), etc. Unsurprisingly, the clause-medial $\hbar i^{23}k\tilde{a}^{34}$ can be replaced by the typical topic markers listed earlier. Take (4), for example:

(22) $\hbar i^{23}$ $t\mathfrak{o}$, $k\tilde{a}^{34}$ $v\mathfrak{o}^{\textit{ʔ}12}h\mathfrak{o}^{34}$
       he   TOP   says   no
       'Unexpectedly, he says NO!'

(23) $\hbar i^{23}$ $m\mathfrak{o}$, $k\tilde{a}^{34}$ $v\mathfrak{o}^{\textit{ʔ}12}h\mathfrak{o}^{34}$
       he   TOP   says   no
       'Unexpectedly, he says NO!'

Compared to $\hbar i^{23}k\tilde{a}^{34}$, however, $t\mathfrak{o}$ usually pairs with another modal particle (even with $\hbar i^{23}k\tilde{a}^{34}$) in the sentence-final position (cf. (9)) to "magnify" the mood of COUNTER-EXPECTATION. On the other hand, ne, a, and $m\mathfrak{o}$ are topic markers of multi-functions with counter-expectation meaning as one of their "peripheral" uses, so that the mood, as shown in (23), can never be as strong or straightforward as with $\hbar i^{23}k\tilde{a}^{34}$ in the TOP position. Therefore, just as with the emergence of $\hbar i^{23}k\tilde{a}^{34}$ as a modal particle in the sentence-final position, it is because there was no such particle functioning independently as a topic marker of COUNTER-EXPECTATION (while there was a need for such a particle in the language) that $\hbar i^{23}k\tilde{a}^{34}$ evolved from a modal particle into a topic-marking particle,[14] by shifting its position into the topic phrase.[15]

(ii)   *$\hbar i^{23}$, $\hbar i^{23}k\tilde{a}^{34}$ $k\tilde{a}^{34}$ $v\mathfrak{o}^{\textit{ʔ}12}h\mathfrak{o}^{34}$
       he   TOP SFP_2 says   no
       'Unexpectedly, he says NO!'

(iii) *$\hbar i^{23}$ $m\mathfrak{o}$, $\hbar i^{23}k\tilde{a}^{34}$ $k\tilde{a}^{34}$ $v\mathfrak{o}^{\textit{ʔ}12}h\mathfrak{o}^{34}$
       he   TOP SFP_2 says   no
       'Unexpectedly, he says NO!'

   The truth is, there is no particle in Shanghainese that occupies an "SFP_2" and allows for a topic marker (in its left TopP), which, however, should be always inserted freely after a topic in Shanghainese as a typical topic-prominent dialect (see Han 2010, 2013; Xu & Liu 2007).

[14] In fact, typologically, a few languages do have verbs of saying used in the CP domain (e.g. Egyptian (Gardiner 1957:173); Koranko (Kastenholz 1987:265–336); Vai (Koelle 1968:123); Bemba (Givón 1980:365–366); Ewe (Lord 1989:307–308); Chamling (Ebert 1991:79–80); Buru (Klamer 2000:78); Lezgian (Haspelmath 1993:367); and Negerhollands CD (Stolz 1986:229), to name just a few). See also Cinque (1999) for examples of evaluative (as well as epistemic) modalities appearing in the complementizer "space." That such a particle, originating from a clause, finally lands in the topic phrase is not often seen, and is even unique within the "topic-prominent" Chinese language.
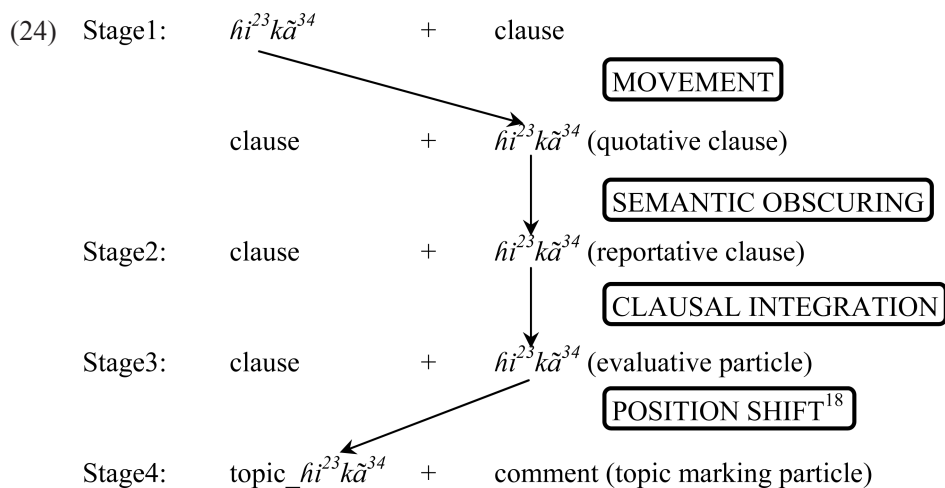
[15] It should be noted that there is no evidence yet to show that the topic-marking $\hbar i^{23}k\tilde{a}^{34}$ has evidential readings.

5. The whole picture

Grammaticalization refers to "the development from lexical to grammatical forms (or functional categories), and from grammatical to even more grammatical forms" (Heine 2003:163). Therefore, it is a complicated process that involves a series of related mechanisms, such as "semantic bleaching" and "syntactic reanalysis" (see Wischer 2010). We have shown that $\hbar i^{23} k \tilde{a}^{34}$ has undergone a universal path evolving from less grammatical to more grammatical elements, and that the stages of its grammaticalization are supported by a chain of cross-linguistic/typological evidence.

As the PRINCIPLE OF DECATEGORIZATION suggests, grammaticalization always involves a loss of categoriality (Hopper 1991, 1996).[16] It is true since, in the later stages, $\hbar i^{23} k \tilde{a}^{34}$ involves neither a speaker nor an action of SAYING; rather, the fresh topic-marking particle is a specialized grammatical marker with specialized functions.[17] It is the PRINCIPLE OF SPECIALIZATION, however, that particularly governs the grammaticalization of $\hbar i^{23} k \tilde{a}^{34}$. As grammaticalization proceeds, the number of possible choices is gradually narrowed down, and one specific item is specialized for this particular function (Wischer 2010). Therefore, SPECIALIZATION is a central aspect of grammaticalization (Hopper 1996). The modal particle in the sentence-final position is thus the result of such functional specialization, since, intralinguistically, it is needed for a specialized particle evolved to express the particular mood of evaluation. Moreover, it is out of the same motivation that such a modal particle evolved into a topic-marking particle (specialized) of the COUNTER-EXPECTATION meaning.

Thus, we have traced the complete path of the grammaticalization of $\hbar i^{23} k \tilde{a}^{34}$, and the stages are summarized in (24):

(24) Stage1: $\hbar i^{23} k \tilde{a}^{34}$ +   clause

                                                    MOVEMENT

              clause   + $\hbar i^{23} k \tilde{a}^{34}$ (quotative clause)

                                                    SEMANTIC OBSCURING

      Stage2:   clause   + $\hbar i^{23} k \tilde{a}^{34}$ (reportative clause)

                                                    CLAUSAL INTEGRATION

      Stage3:   clause   + $\hbar i^{23} k \tilde{a}^{34}$ (evaluative particle)

                                            POSITION SHIFT[18]

      Stage4:   topic_$\hbar i^{23} k \tilde{a}^{34}$ +   comment (topic marking particle)

---

[16] Hopper (1991, 1996) proposes five principles of grammaticalization, namely the principle of layering, the principle of divergence, the principle of specialization, the principle of decategorization, and the principle of persistence. It is proposed that these five principles are observed in a cross-linguistic manner.

[17] It should be noted that "specialization" is a tendency. A certain lexical/grammatical element being specialized with special functions does not mean the "old" functions give way to "new" ones all of a sudden.

[18] Qiang (2010), when discussing the derivation of the Mandarin topic marker *ma* 嘛, proposes that movement from the position of SFPs to that of the post-topic is the typical way to produce new topic markers in Modern

As we can see from (24), the new uses do not necessarily replace the earlier ones; rather "the old layers … remain to co-exist with and interact with the newer layers" (Hopper 1996) (the PRINCIPLE OF DIVERGENCE and the PRINCIPLE OF LAYERING).[19] In the case of $ɦi^{23}kã^{34}$, however, the only thing that remains unchanged is the truth of the proposition.

6. Conclusion

This paper draws the grammaticalization path of $ɦi^{23}kã^{34}$ 'he says' in Shanghainese. We propose that, from the original quotative use, a reportative use was developed in the sentence-final position with an obscured source of information indicative of the speaker's doubt about the proposition (or the speaker's "confirmation-seeking" or "counter-expectation-marking" intentions, in Yap & Ahn's 2012 terms). The reportative clause was then integrated with its preceding propositional clause and was re-analyzed as a modal particle expressing the counter-expectation meaning, based on which $ɦi^{23}kã^{34}$ was further grammaticalized into the new topic-marking particle of counter-expectation. The emergence of $ɦi^{23}kã^{34}$ as a topic particle, as well as a modal particle, was shown to be driven by intralinguistic needs for functional variety and/or specialization. Such a path is claimed to be elegant and practical since not only do the stages involved gracefully follow the principles of grammaticalization described by Hopper (1991, 1996), but the novel evolution of $ɦi^{23}kã^{34}$ as a topic-marking particle is also helpful for the analysis of similar patterns across other languages.

References

Ahn, Mikyung, & Foong Ha Yap. 2012. On the extended uses of evidential markers in Korean: diachronic and discourse perspectives. Paper presented at the Conference on the Nature of Evidentiality, June 13–14, 2012. Leiden: Leiden University.

Aikhenvald, Alexandra Y. 2003. Evidentiality in typological perspective. *Studies in Evidentiality*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 1–31. Amsterdam & Philadelphia: John Benjamins.

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford & New York: Oxford University Press.

Chinese, which Qiang calls "position lowering". In a personal communication between the authors of this paper and Dylan Tsai, it was suggested by the latter that the "lowering" movement could be "raising" in nature. However, one thing we agree upon here is it is typical (though not exclusive) that topic markers in Chinese can be derived from an SFP shifted to the post-topic position. The manner of how it moves (raising or lowering), however, will be dealt with in a follow-up paper focusing on the syntactic nature of SFPs and topic markers.

[19] The final stage of the grammaticalization of $ɦi^{23}kã^{34}$ is also supervised by the Wackernagel rule (Wackernagel 1892)—namely that a (clitic) grammatical element is inclined to occupy the clausal second position—a rule that is found to be applicable not only to Indo-European languages (Harris & Campbell 1995:24; Sweet 1900:43), but also to other language families (Anderson 1981, 1993; Harris & Campbell 1995:216; Kaisse 1985; Nevis & Joseph 1992).

Anderson, Stephen R. 1981. Topicalization in Breton. *Berkeley Linguistics Society* 7:27–39. Berkeley: Berkeley Linguistics Society.

Anderson, Stephen R. 1993. Wackernagel's revenge: clitics, morphology, and the syntax of second position. *Language* 69.1:68–98.

Anttila, Raimo. 1989. *Historical and Comparative Linguistics*. Amsterdam & Philadelphia: John Benjamins.

Bauer, Laurie. 1983. *English Word-formation*. Cambridge & New York: Cambridge University Press.

Chafe, Wallace L. 1986. Evidentiality in English conversation and academic writing. *Evidentiality: The Linguistic Coding of Epistemology*, ed. by Wallace L. Chafe & Johanna Nichols, 261–272. Norwood: Ablex.

Chui, Kawai. 1994. Grammaticalization of the saying verb *wa* in Cantonese. *Santa Barbara Papers in Linguistics* 5:1–12.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York: Oxford University Press.

Cruse, D. Alan. 1986. *Lexical Semantics*. Cambridge & New York: Cambridge University Press.

Dong, Xiufang. 2002. *Cihuihua: Hanyu Shuangyinci de Yansheng he Fazhan* [*Lexicalization: The Origin and Evolution of Chinese Disyllabic Words*]. Chengdu: Sichuan Nationalities Publishing House.

Ebert, Karen H. 1991. Vom Verbum dicendi zur Konjunktion: Ein Kapitel universaler Grammtikentwicklung. *Von Europa bis Ozeanien: von der Antonymie zum Relativsatz*, ed. by Walter Bisang & Peter Rinderknecht, 77–95. Zürich: Seminar für Allgemeine Sprachwissenchaft der Universität Zürich.

Gardiner, Alan H. 1957. *Egyptian Grammar*. Oxford: Griffith Institute.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4.3:333–377.

Hagman, Roy S. 1977. *Nama Hottentot Grammar*. Bloomington: Indiana University Press.

Han, Weifeng. 2010. *A Typological Study of the Syntactic Properties of Topic and Topic Markers*. Shanghai: Shanghai International Studies University dissertation.

Han, Weifeng. 2013. *Word Order Typology: Topic and Topic Marker Structure*. Shanghai: Shanghai Foreign Language Education Press.

Han, Weifeng, Antti Arppe, & John Newman. 2013. Topic marking in a Shanghainese corpus: from observation to prediction. *Corpus Linguistics and Linguistic Theory* June: 1–29. DOI: 10.1515/cllt-2013-0014.

Han, Weifeng, John Newman, & Jiali Mao. 2008–2012. Shanghai Spoken Corpus (SSC) project. A collection of samples of the Chinese dialect spoken in the Municipality of Shanghai, China. China Institute/Department of Linguistics, University of Alberta. Available at http://www.ualberta.ca/~johnnewm/SC/Shanghai/Download.html.

Han, Weifeng, & Dingxu Shi. 2012. The syntactic position and function of $\hbar i^{23} k\tilde{a}^{34}$ in Shanghainese. Paper presented at the 16th Annual Conference of the Chinese Linguistic Society, August 21–23, 2012. Kunming: Yunnan University.

Harris, Alice C., & Lyle Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge & New York: Cambridge University Press.

Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Berlin & New York: Mouton de Gruyter.

Heine, Bernd. 2003. On degrammaticalization. *Historical Linguistics 2001*, ed. by Barry J. Blake, Kate Burridge & Jo Taylor, 163–179. Amsterdam & Philadelphia: John Benjamins.

Heine, Bernd, & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge & New York: Cambridge University Press.

Hopper, Paul J. 1991. On some principles of grammaticalization. *Approaches to Grammaticalization*, Vol. 1: *Focus on Theoretical and Methodological Issues*, ed. by Elizabeth Closs Traugott & Bernd Heine, 17–35. Amsterdam & Philadelphia: John Benjamins.

Hopper, Paul J. 1996. Some recent trends in grammaticalization. *Annual Review of Anthropology* 25:217–236.

Hopper, Paul J., & Elizabeth Closs Traugott. 2003. *Grammaticalization* (2nd edition). Cambridge & New York: Cambridge University Press.

Itani, Reiko. 1991. Japanese sentence-final particle *tte*: a relevance-based analysis. *International Journal of Pragmatics* 2:66–75.

Itani, Reiko. 1994. A relevance-based analysis of hearsay particles: Japanese utterance-final *tte*. *University College London Working Papers in Linguistics* 6:379–402.

Kaisse, Ellen M. 1985. *Connected Speech: the Interaction of Syntax and Phonology*. Orlando: Academic Press.

Kastenholz, Raimund. 1987. *Das Koranko: ein Beitrag zur Erforschung der Nord-Mande-Sprachen*. Köln: University of Cologne dissertation.

Klamer, Marian. 2000. How report verbs become quote markers and complementisers. *Lingua* 110.2:69–98.

Koelle, Sigismund Wilhelm. 1968. *Outlines of a Grammar of the Vei Language*. Farnborough: Gregg International Publishers.

Krönlein, J. G. 1889. *Wortshatz der Koi-Khoin*. Berlin: Deutsche Kolonialgesellschaft.

Law, Ann. 2002. Cantonese sentence-final particles and the CP domain. *UCL Working Papers in Linguistics* 14:375–398.

Lehmann, Christian. 2002. New reflections on grammaticalization and lexicalization. *New Reflections on Grammaticalization*, ed. by Ilse Wischer & Gabriele Diewald, 1–18. Amsterdam & Philadelphia: John Benjamins.

Leung, Wai-mun. 2006. *On the Synchrony and Diachrony of Sentence Final Particles: The Case of wo in Cantonese*. Hong Kong: University of Hong Kong dissertation.

Lipka, Leonhard. 1992. Lexicalization and institutionalization in English and German. *Linguistica Pragensia* 1:1–13.

Liu, Danqing. 2001. Wuyu de jufa leixing tedian [Two major typological features of Wu dialects]. *Fangyan* [*Dialect*] 2001.4:332–343.

Lord, Carol Diane. 1989. *Syntactic Reanalysis in the Historical Development of Serial Verb Constructions in Languages of West Africa*. Los Angeles: University of California dissertation.

Lord, Carol Diane. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam & Philadelphia: John Benjamins.

Lyons, John. 1977. *Semantics*. Cambridge & New York: Cambridge University Press.

Nevis, Joel A., & Brian D. Joseph. 1992. Wackernagel affixes: evidence from Balto-Slavic. Manuscript. Columbus: Ohio State University.

Palmer, Frank R. 1986. *Mood and Modality*. Cambridge & New York: Cambridge University Press.

Plaisier, Heleen. 2007. *A Grammar of Lepcha*. Leiden: Brill.

Qian, Nairong. 1996. Shanghai fangyan de yuqi zhuci [Modal particles in Shanghai dialect]. *Yuyan Yanjiu* [*Studies in Language and Linguistics*] 1996.1:34–47.

Qiang, Xingna. 2010. Huati biaoji 'ma' yu yuqici 'ma' [On 'ma': from indicative mood particle to topic marker]. *Hanyu Xuexi* [*Chinese Language Learning*] 2010.4:59–64.

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

Sauer, Hans. 2004. Lexicalization and demotivation. *Morphology: An International Handbook on Inflection and Word-formation*, Vol. 2, ed. by Geert E. Booij, Christian Lehmann, Joachim Mugdan, Wolfgang Kesselheim & Stavros Skopeteas, 1625–1636. Berlin: Walter de Gruyter.

Seiler, Hansjakob. 1970. *Cahuilla Texts with an Introduction*. The Hague: Mouton.

Seiler, Hansjakob. 1977. *Cahuilla Grammar*. Banning: Malki Museum Press.

Simpson, Andrew, & Zoe Wu. 2002. IP-raising, tone sandhi and the creation of S-final particles: evidence for cyclic spell-out. *Journal of East Asian Linguistics* 11.1:67–99.

Stolz, Thomas. 1986. *Gibt es das kreolische Sprachwandelmodell? Vergleichende Grammatik des Negerholländischen*. Frankfurt am Main & New York: Peter Lang.

Suzuki, Ryoko. 2007. (Inter)subjectification in the quotative *tte* in Japanese conversation: local change, utterance-ness and verb-ness. *Journal of Historical Pragmatics* 8.2:207–237.

Sweet, Henry. 1900. *The History of Language*. London: J. M. Dent.

Tamaji, Mizuho, & Foong Ha Yap. 2013. Where does pragmatic marker *tte* come from? Convergence of to *ifute*, to *itte* and to *te* constructions in the diachronic development of 'say' constructions in Japanese. Paper presented at the 15th Annual International Conference of the Japanese Society for Language Sciences (JSLS 2013), June 28–30, 2013. Nagasaki, Japan.

Tao, Huan, & Jialiang Li. 2009. Fangyan yu xiuci de yanjiu jiemian: jian lun Shanghaihua 'yi jiang' de xiuci dongyin [The interface between dialect and rhetoric: the rhetoric motivations for using s/he says in Shanghai dialect]. *Xiuci Xuexi* [*Rhetoric Learning*] 2009.3:1–8.

Valenzuela, Pilar M. 2003. Evidentiality in Shipibo-Konibo: with a comparative overview of the category in Panoan. *Studies in Evidentiality*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 33–61. Amsterdam & Philadelphia: John Benjamins.

Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1:333–436.

Wang, Jiao, & Foong Ha Yap. 2009. A study of negator *bu* as interrogative sentence final particle in Chinese. Paper presented at the 42nd International Conference on Sino-Tibetan Languages and Linguistics, November 2–4, 2009. Chiang Mai: Payap University.

Wang, Yu-Fang, Aya Katz, & Chih-Hua Chen. 2003. Thinking as saying: *shuo* ('say') in Taiwan Mandarin conversation and BBS talk. *Language Sciences* 25.5:457–488.

Wischer, Ilse. 2010. Aspects of grammaticalization: current resources and future prospects. Invited talk given at Aoyama Gakuin University, Tokyo. Retrieved on August 1, 2012 at http://ir.lib. hiroshima-u.ac.jp/metadb/up/kiyo/AN0002064X/HiroshimaStud-EnglLangLit_55_1.pdf.

Xu, Liejiong, & Danqing Liu. 2007. *Huati de Jiegou yu Gongneng* [*Topic: Structural and Functional Analysis*] (extended edition). Shanghai: Shanghai Education Press.

Yap, Foong Ha, & Mikyung Ahn. 2012. On the development of evidentiality in Cantonese and Korean. Manuscript. Hong Kong: Hong Kong Polytechnic University; Seoul: Hankuk University of Foreign Studies.

Yap, Foong Ha, & Mizuho Tamaji. 2012. On the development of 'say' constructions as evidential and counterexpectation markers in Japanese: a unified analysis of *tte*, *datte* and *ndatte* constructions. Paper presented at the Workshop on Epistemicity, Evidentiality and Attitude in Asian Languages: Typological, Diachronic and Discourse Perspectives, September 3–5, 2012. Hong Kong: Hong Kong Polytechnic University.

Yap, Foong Ha, Charles Tsz-kwan Lam, & Jiao Wang. 2009. Multiple pathways in the emergence of sentence final particles in Chinese: a comparison of *eryiyi* and *yeyiyi*. Online Proceedings of the Second International Conference on East Asian Linguistics, Simon Fraser University, November 7–9, 2008. Vancouver, Canada.

Yap, Foong Ha, Jiao Wang, & Charles Tsz-Kwan Lam. 2010. Clausal integration and the emergence of mitigative and adhortative sentence-final particles in Chinese. *Taiwan Journal of Linguistics* 8.2:63–86.

Yap, Foong Ha, Ying Yang, & Tak-Sum Wong. 2014. On the development of sentence final particles (and utterance tags) in Chinese. *Asymmetrical Functions at the Left and Right Periphery*, ed. by Kate Beeching & Ulrich Detges. Leiden: Brill.

Yeung, Kawai. 2006. On the status of the complementizer *waa6* in Cantonese. *Taiwan Journal of Linguistics* 4.1:1–48.

Weifeng Han (corresponding author)
Department of Chinese & Bilingual Studies
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
shhwf@163.com

Dingxu Shi
Department of Chinese & Bilingual Studies
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
ctdshi@polyu.edu.hk

上海話「伊講」的語法化

韓巍峰　　石定栩

香港理工大學

　　本文對上海話「伊講」多種用法之間的語法化路徑進行了分析描述。我們認為，在直引小句的基礎上，「伊講」通過信息來源的模糊化，在句末的位置發展出了報告性的用法，並最終與之前的小句相結合從而演變為一個句末助詞，用以表達說話者的主觀評價。在其最近的發展中，「伊講」更提升至主題標記的位置並具有表達「反預期」意義的功能。「伊講」語法化的整個過程可作如下描述：直引小句→報告小句→句末助詞→主題標記。該語法化路徑不僅與 Hopper (1991, 1996) 所預測的語法化原則相一致，更有助於對跨語言間相似現象的分析，因而具有理論上的優越性和實用性。

關鍵詞：「伊講」，語法化，上海話，句末助詞，主題標記