# A corpus-based study of the recurrent lexical bundle *ka li kong* 'let (me) tell you' in Taiwanese Southern Min conversations

Miao-Hsia Chang and Shu-Kai Hsieh
National Taiwan Normal University / National Taiwan University

This paper investigates the most frequent lexical bundle (LB) *ka li kong* (to-you-say) (KLK), in an 18.5-hour Taiwanese Southern Min conversation corpus. The analysis focuses on the discourse-pragmatic functions of KLK, the role it plays in the speaker's management of information in talk-in-interaction, and the collocations that are employed. The results show that the speaker utilizes KLK to imply epistemic authority regarding the veracity of the predication. Meanwhile, it expresses the speaker's stance or functions as a discourse organizer to initiate a narrative that is newsworthy. Prosodically, it is always processed as a holistic chunk with great phonological reduction. Along with the low transitivity of the verb *kong* demonstrated by the type of object it takes, we argue that KLK is developing into a discourse marker. Collocation of KLK with the marker *toh* further triggers the grammaticalization of the four-word bundle *toh ka li kong* (TKLK) to encode an extreme stance.

**Keywords:** recurrent lexical bundle, Taiwanese Southern Min, conversation, corpus, stance marking, discourse organizer

關鍵詞: 常用詞串，台灣閩南話，對話，語料庫，立場標記，語篇組織

## 1. Introduction

There has been a proliferation of research in the last three decades on multi-word expressions in spoken and written discourse (e.g. Pawley and Syder 1983; Altenberg 1998; Biber et al. 1999; Wray 2002; Biber et al. 2003; Biber et al. 2004; Carter and McCarthy 2006; Polio 2012; Hyland 2012), in language acquisition (e.g. Bannard and Matthews 2008; Bannard and Lieven 2009), in interlanguage pragmatics (e.g. Blum-Kulka et al. 1989; Bardovi-Harlig 1999), and in EFL writing

(e.g. Granger 1998; Wei 2007; Ma 2009). Related to these expressions are "lexical bundles" (LBs), "language chunks", "pre-fabricated units", or "prefabs", which are assumed to be "stored and retrieved whole from memory at the time of use" (Wray and Perkins 2000: 1). Studies of such word sequences are prompted by the increasing awareness of the fact that native-like language production is characterized by a significant portion of recurrent and ready sets of word sequences (Pawley and Syder 1983; Cowie 1998; Sinclair 1991; Ellis 2012) that serve various levels of communicative functions. One of the most influential findings along this line of thought is perhaps the "idiom-principle" proposed by Sinclair (1991), which argues that there is a "phraseological tendency" (Cheng et al. 2006: 411) for some words to be combined as units more frequently than others. Such units, called LBs hereafter, are extremely frequent in both spoken and written discourse (Sorhus 1977; Altenberg 1998; Biber et al. 1999; Erman and Warren 2000; Foster 2001; Taylor 2012) and allow language learners to achieve fluency in speaking and writing (Nattinger 1988: 77). Analysis of LBs in Chinese discourse has displayed converging results. In his seminal work on emergent grammar in Chinese conversation, Huang (2013) observes that certain intonation units (IUs) contain highly frequent co-occurring words in natural conversation, such as the evaluative expression *shizaishi* 'really', or the stance marker *wenti shi* 'the problem is'. Similar to English (Biber et al. 1999), these bundles straddle two grammatical units and are believed to be newly grammaticalized structures that emerge in discourse. They are "fused, stored, and accessed as units, although their combination shows no semantic or conceptual coherence" (Huang 2013: 60). This paper assumes a corpus-based approach to investigating the use of the recurrent LB *ka li kong* (to-you-say) (共你講) (KLK, hereafter) in conversations in Taiwanese Southern Min (TSM), a Chinese dialect originating in the Southern Min region of China and now spoken primarily in Taiwan and its outlying islets. The verb phrase *ka li kong* (to-you-say) consists of a preposition-like coverb *ka* 'to; give' and its object complement *li* 'you' followed by the main verb *kong* 'say'. This bundle is selected based on our preliminary analysis of recurrent LBs in a TSM conversation corpus in which *ka li kong* was the most frequent three-word bundle.[1] Analysis revealed that the use of this bundle enables speakers to deploy their epistemic resources and manage their knowledge and information vis-à-vis the interlocutors (Heritage and Raymond 2005; Stivers et al. 2011). Several research questions, therefore, are addressed in this study. First, what is the role that is played by KLK in the speaker's management of information in talk-in-interaction? Second, what are the discourse-pragmatic functions performed by KLK in TSM? Do the KLKs display any collocational

---

**1.** 共你講話 (to-you-say-word) 'talk to you', which has one occurrence in the corpus, is excluded from the analysis.

patterns? Third, what are the linguistic properties of KLK in conversation? Are they indicators of the development of KLK into a discourse marker? By investigating these issues, we aim to provide a systematic account of the functions of KLK in TSM conversations. The results will advance our understanding of recurrent LBs in TSM conversations, which are a much less research area in Chinese linguistics.

## 2.   Previous studies

The following two sections provide a review of previous studies in LBs and studies in corresponding forms of KLK in English and Mandarin Chinese. In particular, we focus on the structural and functional characteristics of lexical bundles discussed in the literature.

### 2.1  LBs in discourse

An LB is defined as a recurring and continuous sequence of three or more words (cf. Biber et al. 1999: 990). That is, the words in the sequence are "extended collocations" that "statistically co-occur in a register" (Cortes 2004: 400). For a chunk of words to be considered "recurrent", it needs to occur at least 10 times per million words in the corpus (Biber et al. 1999: 992) (See Section 3 for details). Another important property of LB is that they typically do not represent a complete structural unit but usually span across two structural units (Biber et al. 1999; Biber et al. 2003; Biber et al. 2004; Huang 2013).

LBs are found to be common across both written and spoken discourses (Biber et al. 1999; Carter and McCarthy 2006). The empirical finding from corpus data accords with the argument put forward by Sinclair (1991: 108), that every discourse comprises "a rich repertoire of multi-word patterns that make up a text … totally obscured by the procedures of conventional grammar". The forms and functions of LBs, however, differ slightly in written and in spoken discourse. In written texts, LBs are often grammatically incomplete (Altenberg 1998; Biber et al. 1999; Carter and McCarthy 2006) and more often take the form of noun phrases, or prepositional phrases, such as *the nature of the* preceding *physical world, soil, moral obligation*, or *as a result of* preceding *his work, this change, centuries of experience* (Biber et al. 1999: 991–992). These clusters mainly indicate relations of time and place (*in the middle of*), basic relations of prepositions (e.g. indication of agency: *by the*, indication of purpose: *for the*), or linking functions (*at the time*) (Carter and McCarthy 2006: 832–834). In academic writing, LBs are important for writers to organize their discourse (Sorhus 1977: 214; Biber et al. 1999: 991). It has been found that the LBs used by students in biology and history vary greatly from

those used by published authors in these disciplines (Cortes 2004: 421). By learning these patterns, students are equipped with "ready-made sets of works" in their writing. Additionally, they increase the fluency of the writing and enable the writers to interact with their peers, express their identity in a group, and "reveal the lexico-grammatical community-authorized ways of making-meanings" (Coxhead and Byrd 2007; Hyland 2012: 153).

The forms and functions of LBs in spoken discourse overlap only partially with those in written discourse. Structurally, like in written discourse, LBs also tend to be grammatically incomplete forms. For example, incomplete clauses rather than complete phrases are common, and full clauses are mainly limited to response forms such as *thank you very much* (Altenberg 1998: 104). Compared to those in written discourse, however, noun- or preposition-based forms mainly occur among two-word bundles, such as *kind of*, and *in the* (Carter and McCarthy 2006: 834). In other words, a majority of other multi-word bundles are parts of a clause, negation, or question (Altenberg 1998; Biber et al. 1999: 995; Carter and McCarthy 2006: 830). A commonly found pattern of these LBs is the combination of a pronominal subject and a succeeding VP (Altenberg 1998; Biber et al. 1999; Biber et al. 2004; Carter and McCarthy 2006; Nesi and Basturkmen 2006), some cases of which include the beginning of a complement clause, such as *I don't know why* (Biber et al. 1999: 991–992). Also common among LBs are utterance-ending vague category modifiers, such as *sort of thing*, linking words, such as *and I/but I mean*, and turn-taking signals, such as *don't you think* (Carter and McCarthy 2006: 835). In terms of their functions, LBs "can tell us about the way speakers compose sentences in discourse" (Altenberg 1998: 110). They can perform textual functions, serve interactional purposes, function as discourse organizers, or express speaker's stance. The textual relations signaled by LBs include logical relations such as apposition, contrast/concession, and result/inference (Butler 1997; Nesi and Basturkmen 2006). As discourse organizers, LBs mark the relations between different discourse topics (Cortes and Csomay 2007; Nesi and Basturkmen 2006). LBs can also be used to build interpersonal relations, to refer vaguely to things, as a turn-giving signal, or to invite response from listeners (Carter and McCarthy 2006: 835). Additionally, they also indicate epistemic and attitudinal stance (Biber et al. 1999; 2004; Cortes and Csomay 2007; Kim 2009), or as a rhetorical strategy to evade difficult questions (e.g. *I'm not aware of any*). Some of them even function like "orientational metaphors" (e.g. *as we move forward*) used by the speaker to move the issue under discussion in the right direction (Partington and Morley 2004: 186–7).

The discussion above reveals several common attributes of LBs. First, LBs tend to extend across structural units and are usually grammatically incomplete. Second, they perform a wide variety of functions, from textual building to interpersonal

functions. Third, many LBs involve verbs with low semantic content, such as *think, know, do,* and *get*. Such a tendency is consistent with the arguments made by Thompson and Hopper (2001), Scheibman (2002), and Huang (2013) that epistemic/evidential clauses or verbs of low transitivity account for most predicates in conversation. According to our initial observation, the characteristics of the LBs summarized above agree with the structural features of KLK, the most frequent LB in our TSM corpus and the focus of the present study. The functions of KLK will be discussed in Section 4, following a literature review of its near-synonyms in English and Mandarin Chinese and a description of the methodology.

## 2.2 *Let me tell you* and related expressions in English and Mandarin

There have been studies on the corresponding forms of KLK in English and Mandarin Chinese. For English, corpus studies have shown that the corresponding expressions *I'll tell you what/I'll tell you* are common in conversation. They serve as "utterance launchers" to present "a personal stance relative to the information in the following complement clause" (Biber et al. 1999: 1003). They can also be used as a focus or topic introducing marker (Conrad and Biber 2004: 66). The first person pronoun in such expressions in dialogic discourse achieves the effect of persuasion (Hrisonopulo 2008). In fact, early works in discourse analysis have already noted the topic introducing function and subjective meaning of *what*-associated expressions. In particular, Östman (1981: 52–54) notes that that children as young as four years old are able to employ *guess what/you know what* as an attention-getting, topic-changing or turn-taking device. In addition, expressions like *guess what* can be used to open narratives (Stubbs 1983: 23). The development of *what* from a pure interrogative to a signal of "interpersonal meaning of shared knowledge" can be regarded as a case of "pragmatic strengthening" whereby the speaker's attitude is increased via pragmatic inference (Traugott and Dasher 2002).

Studies in Mandarin Chinese have also focused on the pragmatic meanings of the functional equivalents of KLK. Different from TSM, which uses KLK exclusively to express the meaning of 'Let (me) tell you', Mandarin has a few corresponding forms, including *wo gen ni shuo* (I-with-you-say) 'I tell you' (Gan 2012), *wo gaosu ni* (I-tell-you) (Dong 2010), *wo gei ni shuo* (I-give-you-say) (Chen 2010), and *wo gen ni jiang* (I-with-you-talk) (I'm telling you) (Quan and Yu 2014) in Mandarin discourse. It has been found that the epistemic verb *shuo* co-occurs with an associated clause to emphasize a piece of information and may carry additional pragmatic force (Dong 2010; Chen 2010; Gan 2012; Quan and Yu 2014). These synonyms may occur as an idiomatic expression (Chen 2010), convey a sense of "epistemic primacy" (Quan and Yu 2014), may be semantically bleached, or function as a discourse marker (Dong 2010; Gan 2012) with procedural functions

(Chen 2010: 85). With these functions, the non-occurrence of this marker does not affect the propositional meaning of the corresponding utterance, such as *wo gaosu ni, buguan wo zou dao nali, wo buhui fangqi menghan* 'I'm telling you, wherever I go, I won't give up Menghan.' (Dong 2010: 279). Furthermore, the phrase can occur either with or without the first person subject overtly expressed (Chen 2010). Functionally, the phrase may provide or introduce an important piece of information to the addressee, such as *gen ni shuo, ni xianzai hen weixian* 'I'm telling you, you are very dangerous now.' It may also precede an associated clause that makes a request to the addressee, such as *wo gen ni shuo, nainai buhui zai yao ni ruzui le, zhejian shiqing guoqu le, ni kuai gen wo huijia ba!* 'I'm telling you this. Grandma will not ask you to marry and live with the wife's family anymore. Let's forget about this. Please come home with me now!'. Another usage of *wo gen ni shuo* is to act an attention getter to make sure that the hearer gets the information (Gan 2012: 149; Tao 2015: 343). Tao further shows that *gen ni shuo/jiang* (with-you-say) is among the top 50 bundles in his Mandarin spoken corpus (20th place). Another meaning, mainly found in the use of *wo gaosu ni*, is to issue a warning to the addressee, e.g. *wo gausu ni, ni zai fuzhou yaoshi gen le bieren, wo zai le ni!* 'I'm telling you this. If you follow someone else in Fuzhou, I'll kill you!' (Dong 2010: 281). Last, *wo gen ni shuo* may precede a clause reminding the addressee of an important fact since forgetting it has led to some misconduct by the addressee, for example, *wo gaosu ni, ni keneng shi ge daguan xianyao, danshi, wo bing bushi ni de buxia* 'Let me tell you this; you may be a high ranking official, but I'm not your subordinate!' (Dong 2010: 281).

The studies reviewed above shed important light on the use of the lexical phrase *wo gen ni shuo* and its semantic variants in Mandarin. However, none of these studies provide a systematic analysis of their functions in extended conversation. Nor do previous studies address the use of the corresponding form, KLK, in TSM. In this paper, we explore the functions of KLK by analyzing its frequency, structural properties, collocations, and sequential relevance. These interactional contingencies will be shown to provide strong support for the development of KLK into a discourse marker (For a definition of "discourse marker", see Section 3).


## 3. Data and terminology

This study examines the recurrent LBs retrieved from a spoken corpus in TSM. The corpus used is composed of 119 episodes of natural conversations comprising 18.5 hours. Altogether 209 different speakers (121 females and 88 males), of ages ranging from 22 to 82 years at the time of recording, participated in the

conversations. The conversation topics included daily routines, household chores, children, family relationships, work, neighbors, and childhood memories.

The data were entirely transcribed into Intonation Units (IUs) following the "narrow" convention developed by Du Bois et al. (1993), thus including notations of pause length, prosodic information, paralinguistic signals, and detailed transcribers' notes in the transcript (See Appendix I). Since a single, official Taiwanese orthographical system has yet to be created, the Chinese characters employed in the transcription primarily follow those issued by the Ministry of Education of the Republic of China (ROC) (http://twblg.dict.edu.tw/holodict_new/index.html), with certain modifications for ease of computation of the recurrent strings. Where a Taiwanese character was unavailable, the Church Romanization developed by Cheng and Cheng (1977) was adopted. The corpus yielded 251,419 words in 55,066 IUs with a mean length of 3.91 words per IU, a length similar to that of spoken Mandarin (Huang 2013) and English (Altenberg 1998: 103). Considering the challenges inherent in transcribing a spoken language devoid of a standard orthography and given the meticulous prosodic information the transcription required, accruing 18.5 hours of data was an arduous task. We did however build a corpus that is comparable to certain well-established corpora, such as the Santa Barbara Corpus of Spoken American English, which contains 249,000 words, and the corpus utilized by Thompson and Tao (2010), which comprises 268,928 words.[2] Our corpus is also half the size of the well-known London-Lund Corpus, which runs to 500,000 words in length (Greenbaum and Svartvik 1990). In short, we believe that this corpus is among the largest fully-transcribed spoken corpora of natural conversations in TSM including detailed prosodic information.

The transcription was segmented with the online system developed by Ungian Iunn (http://poj.likulaw.info/hanlo_hunsu.php) and refined based on the criteria proposed by Tseng (1997), who provides an extensive list of function words for the segmentation of Taiwanese words. The frequency list in our corpus was generated with the software R (http://www.r-project.org/), which scanned each utterance in the spoken corpus, extracted every two-word sequence, stored them, and generating an extensive frequency list. The same procedure was conducted to extract three-, four-, five, and six-word sequences.

The majority of the LBs fell on three-word sequences, with the most frequent LB being KLK.[3] Across 51 different episodes, 139 KLKs were used. Among them, 91 carried an overt subject *gua* 'I', with 60 of them immediately preceding KLK,

---

**2.** The Santa Barbara Corpus of Spoken American English: http://www.linguistics.ucsb.edu/research/santa-barbara-corpus#SBC015.

**3.** The top 10 LBs in the TSM corpus include those which also contain the saying verb *kong* as well as those which contain the knowing verb *tsai*.

forming the four-word bundle *gua ka li kong* (I-to-you-say) 'Let me tell you; (I'm) telling you'. The results are consistent with Biber et al. (1999) and Carter and McCarthy (2006) on the predominance of three-word bundles in conversations.

A recurrent bundle is defined as one that occurs with significant frequency in a corpus. The minimum frequency ranges from 10–40 times in various studies. Carter and McCarthy (2006: 829–31) adopted the frequency of 20+ times per 5 million words as the criterion; Biber et al. (1999), Altenberg (1998), Simpson-Vlach and Ellis (2010) and Partington and Morley (2004: 180) used 10 times per million words (PMW); Biber and Conrad (1999) and Hyland (2008) treated 20 PMWs as a recurrent bundle, while Biber et al. (2004) opted for 40 PMWs (see also Liu 2012: 26). Since KLK occurs 139 times in our corpus of quarter million words, we believe that it is of remarkable phraseological value and involves significant pragmatic functions.

As the discussion unfolds, it will become clear that KLK is developing in a discourse marker. There has been considerable research on discourse markers, which serve as connectors, topic-introducers, attitude signals, turn-takers, fillers, and discourse-organizers. As diverse as the functions may be, according to Brinton (1996) and Jucker and Ziv (1998), a discourse marker manifests a significant amount of the following features, although few markers display all these features:[4]

**Table 1.** Features of a discourse marker (Summary of Jucker & Ziv 1998: 3; Brinton 1996: 33)

| | |
|---|---|
| (a) Phonological reduction | (g) With little/no propositional meaning |
| (b) As a separate tone group | (h) Multifunctionality |
| (c) Not a traditional word class | (i) Typical of oral discourse/informality |
| (d) Sentence-initial position | (j) High frequency |
| (e) Outside matrix clause or with loose connection | (k) Being stylistically stigmatized |
| (f) Optionality | (l) Typical of female speech |

It is further noted that features (a-g) are crucial criteria for defining a discourse marker, whereas features (i-l) are more "descriptive" and "controversial" (Jucker and Ziv 1998: 4). In general, however, the more features a marker has, the more it can be taken as a prototypical discourse marker. In the following discussion, these criteria will be used for our account of the functions of KLK in TSM.

---

**4.** We are grateful to one anonymous reviewer for raising our attention to the definition of discourse markers.

## 4.   Functions of KLK in conversation

The 139 tokens of KLKs were analyzed and classified according to their discourse-pragmatic functions. All functions were discussed thoroughly between two researchers before a final classification was established. The analysis has shown that the majority of the KLKs (79%) are used to express speaker stance or to initiate a storytelling as the new topic of a discourse. The typical telling function is only found in about one fifth (21%) of its occurrences. Especially noteworthy is the fact that when KLK expresses an extreme attitude, it collocates with the emphatic marker *toh* (就) and is being grammaticalized into a discourse marker (See Section 4.2.3 for details). Table 2 presents the functional distribution of KLK in our corpus:

**Table 2.**  Functional distribution of KLK in the TSM corpus

|  | # | % |
| --- | --- | --- |
| Telling function | 29 | 21% |
| Stance marking | 89 | 64% |
| Discourse organizing | 21 | 15% |
| **Total** | **139** | **100%** |

Before we discuss the functions of KLK, there is an important caveat to acknowledge concerning the classification of KLK. As noted in previous research, one single occurrence of a discourse element can, in fact, have multiple functions. However, it is often the case that one function is usually prominent among other plausible interpretations (e.g. Lindemanna and Mauranen 2001; Biber et al. 2004: 383; Stenström 2004: 264; Nesi and Baskurkmen 2006: 293; Salazar 2014: 84). Therefore, instead of claiming that a given usage was KLK's sole function, we consider each function being discussed as primarily exemplifying the most important usage of this bundle (See also end of Section 4.2 for a discussion of the ambiguous cases).

### 4.1  Typical telling function

The typical telling use of KLK refers to the telling of a piece of information to the addressee. Most of them take a first person subject (18/29, 62%), and the remaining cases take a third person singular subject, either explicitly or implicitly, as shown in Table 3. As for the object of the verb *kong*, it can be an NP, a clause, or a zero anaphora, as shown in Table 4. Excerpt (1) illustrates this typical telling function of KLK.

**Table 3.** Subject type of KLK with the typical telling function

|   | 1p explicit | 1p implicit | 3p explicit | 3p implicit | Other | Total |
|---|---|---|---|---|---|---|
| # | 14 | 4 | 9 | 0 | 2 | 29 |
| % | 48% | 14% | 31% | 0% | 7% | 100% |

**Table 4.** Object type of KLK with the typical telling function

|   | NP | Clause | Zero | Total |
|---|---|---|---|---|
| # | 13 | 2 | 14 | 29 |
| % | 45% | 7% | 48% | 100% |

(1)  (The speakers are talking about which car to use for a trip during the holidays.)

```
67 AS:   toh i  meme             u    tsit tai tshia ma,\
         TOH 3S younger.sister have one  CL   car    PA
68 BR:   henn henn henn.\
         RT   RT   RT
69 AS:   m  tsai u      huatto bo?_
         NG know have method PA
70 BR:   enn,_
         PA
71 BR:   tsunkong m  si u    tsit tai kaulangtse       hit  tsiong siunnatshia?_
         PN         NG CP have one  CL  nine.people.seat that kind   van
72 AS:   a   tsai._
         how know
73 AS:   huinntit,_
         anyway
--> 74 AS: sian       ka li kong tsitle siausit anne      la honn.\
         previously to 2S say  this   news    this.way PA PA
75 BR:   o.\
         RT
```

AS (67): His sister has a car.
BR (68): Yeah yeah yeah.
AS (69): (I) don't know if (she) can give us a ride (during the New Year's holidays).
BR (70–71): Uh...doesn't Tsunkong also own a van that will seat nine people, the box-like car?
AS (72): (I) don't know.
AS (73–74): Anyway, *I'm telling you* this plan in advance.
BR (75): Um.

(2)  (The speaker is telling the addressee about how to decorate the house for better fengsui.)

```
-->184. DW:(0) kisit    ka li  kong si <M cankao      cankao  M> la.\
             actually to you say  CP    reference reference  PA
185. DW:…ah= toh=,_
         DM  TOH
186. DW:…bo besin       la,\
         NG superstition PA
```

DW: In fact, *I'm telling you this* simply for your reference. I'm not being superstitious.

In (1), KLK is used to close the conversation. In (2), along with the object NP, the clause serves as the sentential subject. Both cases of KLK carry their typical propositional meaning of telling.

## 4.2 Stance marking and grammaticalization of the discourse marker toh KLK

As a stance marker, KLK launches a spate of dialogue that encodes a strong evaluative tone within speaker meaning. This function accounts for over half of the occurrences of KLK (64%, 89/139) and is predominantly associated with a first person subject (Table 5). Different from the telling KLK, a stance marking KLK only takes a clausal object or a zero object (Table 6). Functionally, KLK may present a negative appraisal of a situation at hand, emphasize the obviousness of a situation that the addressee should have known earlier, and/or preface an utterance that performs an illocutionary speech act. The strong stance implied in the utilization of KLK is reinforced by a variety of collocating expressions that signal different degrees of negativity. Among them, the emphatic marker *toh* forms a four-word bundle with KLK, i.e. *toh ka li kong*, and is being grammaticalized into a discourse marker. The semanticition of *toh KLK* is motivated by invited inferencing through pragmatic strengthening of the contextual and intersubjectified meaning (Traugott and Dasher 2002: 35) as this bundle is recurrently used in TSM conversations. The different usages of stance-marking KLK are discussed respectively 4.2.1–4.2.4.

**Table 5.** Subject type of KLK with the stance marking function

|   | 1p explicit | 1p implicit | 3p explicit | 3p implicit | Other | Total |
|---|---|---|---|---|---|---|
| # | 62 | 21 | 5 | 0 | 1 | 89 |
| % | 70% | 24% | 6% | 0% | 1% | 100% |

**Table 6.** Object type of KLK with the stance marking function

|   | NP | Clause | Zero | Total |
|---|---|---|---|---|
| # | 0 | 83 | 6 | 89 |
| % | 0% | 93% | 7% | 100 |

### 4.2.1   *Presenting a negative appraisal*

As a stance marker, KLK can be employed to introduce a sequence of talk containing information to which the speaker has exclusive access. The bundle presents a critical evaluation of a situation at hand, wherein the subject of the verb, which is exclusively the first person singular *gua*, is always overtly expressed, forming the

bundle *gua ka li kong* 'I-to-you-say' (Let me tell you; I'm telling you). Excerpt (3) illustrates the use of KLK in a first assessment (Pomerantz 1984), prefacing an onslaught of the speaker's criticism of her son's untidy room:

(3)   (LY is complaining about the considerable time she spent cleaning her son's room before his wedding.)

```
--> 153 LY: gua ka li kong,_
              1S  to 2S say
   154 LY: kanna laibin honn,_
           only  inside PA
   155 LY: ukau       khua e  la.\
           extremely wide AP PA
   156 LY: kanna laibin toh piann a    be tshingkhi le kong,_
           only  inside TOH clean also NG clean     PA PA
   157 LY: piann tsap gua  kang ma   si iaukoh lasap ne!\
           clean ten  more day  also CP still   dirty PA
   158 LY: tsiok khoolian e.\
           very  pitiable AC
   159 LY: li long m  tsai gua anne—
           2S all  NG know 1S  this.way
   160 PJ: ah!\
           EX
   161 PJ: li na khi kau   guan hia    khuann li toh kamkak kong,_
           2S if go  reach 1PE there see     2S TOH feel   CM
   162 PJ: o,\
           RT
   163 PJ: lin tau  ukau       tshingkhi e  la.\
           2P  home extremely clean      AC PA
   164 LY: ho!\
           EX
   165 LY: tsintsiann ue      kah  that.that.      kah  tinn tinn tinn.\
           really     litter till stuffed.stuffed till full full full
   166 LY: long that a    that be khue anne.   khi   kah  anne.\
           al   kick also kick NG pass this.way angry till this.way
```

LY (153–156): *You know what?* The room is so large and (he) had littered the room to such an extent that I could never seem to complete the cleaning.
LY (157): I spent over 10 days cleaning it but there was still trash everywhere.
LY (158–159): You don't know how much I suffered.
PJ (160–1163): Ah! If you see my house, you will know how clean your house is.
LY (164–166): Alas! The trash was stacked up to the ceiling, and it was so crammed that there was almost no space to move around or kick things away. It really pissed me off!

The utterances with KLK in (3) include LY's aggravation on the ordeal she suffered because of the countless hours she spent cleaning the heaps of rubbish dumped in her son's junkyard-like room before his wedding. The appraisal is accompanied by the emphatic marker *toh* (Lin 1996) or adverbials that express exclusiveness (e.g. *kantan* 'only') or extremity (e.g. *ukau* and *tsiok* 'extremely'). Another interesting co-occurring element is the rhetorical frame *li m chai* 'you don't know', which

collocates with the adverb of totality *long* 'all' to enhance the distress the speaker experienced, hence increasing the critical effect of the propositional content.

KLK also occurs in a response, as a second assessment. With KLK, the speaker displays disalignment with the addressee:

(4)   (TR and RM are talking about their childhood. TR asks whether RM ever
       went swimming without his parents' knowledge.)

```
141 TR: ..li long bo,_
          2S all  NG
142 TR: ..tshutkhi—
          go.out
143 TR: ..thau       iuieng a=,_
          stealthily swim   PA
144 RM: ..be hiau a,_
          NG know PA
--> 145 RM: <F gua ka li kong [gua sehan guan tau  bo beh  gua F>,_
              1S  to 2S say   1S  young 1P   home NG want 1S
146 TR: [liah  hia]._
          catch fish
147 RM: ..liah  hia  ma   si ai   ham,_
          catch fish also CP need with
148 RM: situalang tshutkhi.\
          parents   go.out
```

TR (141–143): You never snuck off and swam alone (in the creek)?
RM (144–145): I didn't know (how to swim). *You know what?* When I was young, my family didn't allow me to (play alone).
TR (146): To catch fish or something.
RM (147–149): Even if I went fishing, I required adults' supervision.

In (4), RM defends himself by claiming that he did not learn swimming because he was not allowed to play in the water without adults' company. The utterance is surrounded by two negative expressions *behiau* 'not able to' (line 144) and *bo* 'not (want me to go out alone)' (line 145), about TR's inability to swim and TR's parents not allowing him to catch fish in the creek by himself. In particular, KLK along with its succeeding utterance is delivered in fast speech (indicated with <F F>), which implies RM's eagerness to refute TR's teasing. Furthermore, since the talk involves RM's personal experience, he has first-hand information and therefore is epistemically superior to TR (Heritage 2002; Heritage and Raymond 2005; Stivers et al. 2011; Hayano 2011; Quan and Yu 2014) in terms of the activity being discussed.

Another instance of KLK in the responsive position is given in (5), where a father and his wife engage in a tug-of-war using KLK to convince their son that their spouse has been addicted to TV drama watching.

(5)  (FA, MO and their son JB are chatting. FA invites his wife MO to go for a
    walk in the afternoon. However, MO refuses, and FA attributes her rejection
    to her desire to watch a Korean drama sequel that is airing that afternoon.
    Therefore, MO defends herself by accusing FA of his indulgence and his
    obesity instead.)

```
4 FA: epo         tshutkhi=--
      afternoon go.out
5 FA: …anne      kiann kiann e  honn,_
      this.way walk  walk  AC PA
6 FA: …(.9) tsingsin ke  khah ho.\
            spirit   add more good
7 MO: …(2.2) a  li tsiann pantuann e  la.\
             DM 2S very   lazy      AC PA


((11 IUs later))

19 FA: li he   long—
       2S that all
20 FA: …khi ho <M hanju M>      hai._
       go  PS   Korean.drama harm
21 MO: …henn,_
       RT
22 MO: …khuann <M hanju M>,_
       watch      Korean.drama
23 MO: …hmn,_
       RT
24 MO: a  bo lang   beh  pue      tongjian  ma  khuann <M hanju M>.\
       DM NG person want accompany of.course also watch    Korean.drama
25 JB: …(2.6) hit  tshut   o./
             that episode PA
26 FA: …(.9) tioh  a,_
             right PA


((4 IUs later.))

--> 30. MO: ..gua ka li kong o,\/ ((talking to her son)
            1S  to 2S say  PA
31 MO: lin papa   uitioh khuann <M hanju M>      o,\
       2S  father for    watch      Korean.drama PA
32 FA: …(1.2) pui sann  kongkin._
             fat three kilogram
33 MO: (0) pui sann  kongkin._
           fat three kilogram


((6 IUs later.))

40 MO: i  toh si [u      huatto] ka li siann./
       3S TOH CP  have tactic   to 2S entice
41 FA: [bo=],/
       NG
```

```
--> 42 FA: gua ka li kong--
           1S  to 2S say
43 FA: he   tso  nng tiamtsieng m  si tiongtiam._
       that play two hour        NG CP important.point
44 FA: …(1.1) tiongtiam      toh si kong,_
              important.point TOH CP say
45 FA: …i  tso hankok e  hitle,_
        3S do  Korea  NM that
46 FA: …(.8) hitlo <M lishi      ju M>._
             that     historical drama
47 FA: …(.8) a=,_
             DM
48 FA: …(1.7) toh e    ho li kamkak kong honn,_
              TOH will DP 2S feel   CM   PA
49 FA: ah--
       DM
50 FA: …anne=    ittit    khuann lokhi._
       this.way continue watch  DC
```

FA (4–6): Let's go for a walk this afternoon to refresh ourselves.
MO (7): You have been a couch potato.
FA (19–20): You've become addicted to this Korean drama sequel (so you won't go for a walk with me.)
MO (21–24): Yeah. I've been watching Korean dramas because I usually have no one for company.
JB (25): (Do you mean) that sequel?
FA (26): Yes.
MO (30–33): *You know what?* Your father has gained three kilograms because of Korean dramas. He always attaches himself to the chair, watching the sequel after lunch (without moving).
MO (40): Yes. They know the tactics to lure you.
FA (41–50): No. *Let me tell you why.* The point is not that they run for two hours. The point is that the dramas are historical stories that keep you glued to the couch (to know what happens next).

In lines 4–6 in (5), FA invites his wife MO to go for a walk. This reminds MO of how lazy FA has been recently because of his weakness for Korean dramas. MO claims Korean dramas turned FA into a couch potato, resulting in his increased weight caused by lack of exercise. The husband, however, disagrees with his wife's assessment, and again, with KLK, presents a second assessment with the negative marker *bo* 'No' (lines 41–42) to correct MO's claim regarding the reason for his obesity – the uninterrupted historical drams that kept FA on the couch and thus led to his weight gain.

### 4.2.2  *Repetition of an obvious fact*
A further assertive force of KLK utterances is to indicate an important and obvious fact that the protagonist should have known earlier. The sequential position of KLK in such usages can be considered as a responsive statement to an implicit

or explicit prior assessment. In particular, it involves a repetition of an assertion made in a previous interaction with a speech recipient to discourage the addressee from pursuing his or her line of thought (cf. Clift 2006: 583, on reported speech). The assertion is usually reinforced by the emphatic marker *toh* and *ikieng* 'already' to strengthen the obviousness (Lin 1996: 43) of the fact, thus portraying a negative stance, as exemplified in (6):

(6) (AR is a part-time housekeeper. She is telling her relatives about her employer. Before the exchange of talk shown below, AR told the hearers that her boss kept asking her when she could go back to work.)

```
--> 181 AR: <Q gua toh ikieng  ka li kong,_
               1S  TOH already to 2S say
    182 AR: gua na u  ti e,_
            1S  if AP in AC
    183 AR: gua toh itieng    e    lai Q>,_
            1S  TOH definitely will come
    184 AR: gua lohkhi  ekang,_
            1S  go.down south
    185 AR: ah   bo lohkhi ^kui    ^kang,_
            also NG go.down several day
    186 AR: ke  ke  e  lohkhi  nng kang la,_
            add add AC go.down two day  PA
    187 AR: huana,_
            babarian
    188 AR: he  ^ti  toh si ti._
            that pig TOH CP pig
```

AR (181–188): "*I have* already *told you* that if I am (in Taipei), I will come clean the house." I only spent a total of two days in the south, (but he kept bothering me with the same question). The barbarian! A pig is a pig!

The conversation presented in (6) contains a quoted material with KLK, in which AR reiterates the promise to her employer before the Chinese New Year, stating she (AR) would come back to work immediately after she returned from the south following the New Year's holidays. The continuous reminders by the employer enraged AR and triggered her belligerent comment that he was a *ti* 'pig'. Together with the emphatic marker *toh* and the aspect marker *ikieng*, KLK displays the speaker's extreme feelings. *Toh* collocates with KLK so frequently that it has almost developed into a discourse marker expressing extreme stance. Section 4.2.3 elaborates on this.

### 4.2.3   *Grammaticalization of the emphatic discourse marker toh KLK*
There are 24 cases of KLK that collocate with the emphatic marker *toh*, nineteen tokens of which immediately precede KLK, forming a four-word bundle *tohkalikong* (TKLK hereafter). TKLK typically occurs in a second assessment, with the

emphatic function of *toh* and the epistemic superiority implied by KLK creating an even stronger assertive force. It is worth noting that the subject of KLK is only weakly referential. A total of 15 out of the 19 cases of TKLKs have unexpressed subjects, and even if the subject is expressed, the primary intention of the utterance is to convey the speaker's stance rather than to indicate who utters it. Furthermore, *ka* and *li* are usually reduced to *kai*, hence *tohkaikong*. The conversation in (7) illustrates an instance of TKLK that conveys the speaker's firm, negative stance:

(7)   (PY, a relative of TF, is visiting TF's house and asking for the prescription of a herbal powder for his wound.)

```
 9 PY: bo a  gua=,_
          NG PA 1S
10 PY: toh% toh% iong,_
          TOH  TOH  use
11 PY: iong wan    khi a—
          use  finish go  PA
12 PY: ah na u,_
          DM if have
13 PY: …gua khuann aujit       kaki,_
          1S   see     another.day self
14 PY: tshau tshau le gua kaki lai ^kap    la.\
          copy  copy  PA 1S  self come match PA
--> 15 TF:(0)<L toh ka li kong li bo= te     kap   li kanna
                  TOH to 2S say  2S NG  place match 2S insistently
           beh  tshau= tshau L>._
           want copy    copy
16 TF: …(4.5) toh khiam hit  bi._
              TOH lack  that ingredient
17 TF: …khaim hit  bi         toh bo de    be  li kanna        bueh._
          lack  that ingredient TOH NG place buy 2S insistently want
18 TF: …(.8) iahbo     gua ah   behiau    beh  tshau ho   li._
              otherwise 1S  also not.know want copy  give 2S
```

PY (9–14): No. I have run out of the powder. You can give me (the ingredient list). I can (go to a Chinese medicine store) and buy the powder myself.
TF (15–18): *I have (already) told you* that no place makes (this powder). Why do you keep saying that you want to copy (the prescription)? One ingredient is unavailable and it's not being sold anywhere. Why do you keep (saying you want to do it yourself)?. (If the ingredient was available), why wouldn't I just give you the prescription?

The extract in (7) occurs when PY, a nephew of TF, visits TF to ask for a herbal powder that cures one's chronic wounds. The powder was developed by TF and a friend, since they had studied Traditional Chinese Medicine. PY visited TF earlier and was given some of the powder. At that time, TF only had a small amount left. However, since the powder is no longer available while PY continuously asks TF to collect the ingredients, annoying TF who responds impatiently that one ingredient is unavailable. As TF informed PY about this during his previous visit, line

14 in (7) indexes the obviousness of the information and the addressee's lack of knowledge regarding the availability of the ingredient. Here, the main proposition that conveys TF's annoyance is *li bote kap li kanna beh tshau tshau* 'No place makes (this powder). "Why do you keep saying you want to copy" (the prescription)?'.

   The following exchange presents another case of TKLK that marks the repetition of a previous statement used to convey the speaker's impatience. As the speaker HG has exclusive access to the "completeness" of the story (Stivers et al. 2011: 18), her speech expresses a sense of epistemic primacy. Note that the subject is omitted, suggesting the subject is relegated to the background. What is foregrounded is the speaker's strong stance.

(8)   (HG is telling the addressee AL about an Englishman who sued McDonald's
      because his daughter was accidentally burned while he was eating
      McDonald's apple pie.)

```
458 HG: iengkok tsit ui=,_
        Britain one  CL
459 HG: <M lamtsu M>honn,\
           male      PA
460 HG: tsit ui== lamsu     piengyu._
        one  CL   gentleman friend
461 AL: (0) hm.\
            RT
462 HG: (0) i  ti tsit e <M ^mai^tanglao M> tsiah mikiann e  sitsun ne,_
            3S at this CL    McDonald's      eat   thing  AC time   PA
463 HG: ..(H) unn=,_
             DM
464 HG: khi tsiah tse <M pingguapai M>.\
        go  eat   this  apple.pie
```

((118 IUs here, where HF continues the story that the apple sauce splattered as the
father bit it and hurt his daughter sitting next to him. The burn was so serious
that he sued MacDonald's for his daughter's loss. AL then wonders why the burn
was so serious.))

```
583 AL: hit  e  hit te <M pingguopai M> si gua tua te anne     nng e
        that CL that CL   apple.pie     CP how big CL this.way two CL
        lang?_
        people
--> 584 HG: a  toh ka li kong anne     ka   lokhi [tsut 1] tsite,\
            DM TOH to 2S say  this.way bite down   splash   AP
585 AL: [tsut   tsite honn= 1],\
         splash AP    PA
586 AL: hio a [hio a 2]._
        RT  PA RT  PA
587 HG: [tsoann 2] tsut   khi a  i=,_
          thus      splash go  DM 3S
588 HG: ..<M xiaonuer M>     na   tua      tse ti i e pinna e [sitsun 3] a,_
            little.daughter seem just.now sit at 3S AC side  AC time      PA
```

```
589 AL: [henn henn 3].\
         RT   RT
590 HG: (H) tioh  bo?\
             right PA
591 AL: (0) henn.\
            RT
592 HG: a  tsu     tsite tsoann%,_
        DM splatter AP    thus
593 HG: tsut    ui i  e [bin  khi a  toh 4]--
        splatter to 3S AC face go  DM TOH


HG (456–460): Uh, the story goes like this. There was an Englishman.
AL (461): Yeah.
HG (462–464): He was eating an apple pie at McDonald's.
AL (583): Was the apple pie big enough to injure two people?
HG (584): I've already told you that the moment he bit (the apple pie), the sauce
splattered.
AL (585–586): It splattered. I see.
HG (587–588): Then it splattered onto his little daughter's face. She was sitting
beside him.
AL (589): Hm hm.
HG (590): Okay?
AL (591): Hm hm.
HG (592–593): And the girl's face was covered with (the sauce).
```

The preceding discussion focuses on the emergence of the four-word bundle TKLK as a fixed expression to demonstrate a strong speaker stance toward a given situation. As our corpus contains a significant number (19 tokens) of the four-word bundle TKLK (19/0.25 million, equivalent to 76 time per million words), we argue that that *toh ka li kong* is developing into a discourse marker. The features of TKLK discussed above correspond to those of a discourse marker reviewed in Section 3 except the features (k) and (l). In fact, most of the characteristics that TKLKs exhibit are also shared by KLKs (cf. Section 4.5) except those with a typical telling function. However, TKLK deserves special attention because of the strong association between *toh* and KLK. In short, TKLK is stored as part of the speaker's linguistic repertoire (Langacker 1987: 57; De Smet 2017: 75) and processed automatically as a holistic chunk to indicate an extreme negative stance regarding the information over which the speaker has epistemic authority. The occurrence of TKLK also typifies the pragmatic strengthening (Traugott and Dasher 2002) of meaning through its recurrence in discourse.

### 4.2.4   *Prefacing a speech act*

The stance-marking KLK also provides a resource for speakers to perform an illocutionary speech act of threat, request, advice, accusation, or reminding. In this use, the bundle always takes an overt first person subject *gua* 'I'. Sequentially, KLK appears as a second assessment of a state of affairs discussed in the immediately

preceding discourse. The appraisal conveys a strong sense of epistemic superiority. That is, there seems to exist an asymmetrical power relation regarding the access to the knowledge imparted by the current speaker. As for its position in an IU, KLK may occur in the utterance-initial or final position, the latter of which characterizes KLK with a higher degree of face threat to the addressee. In the final position, KLK marks the end of a conclusive remark about the speaker's extended admonishment to the addressee. Three cases of stance-taking KLK are found to occur at such a sequential position. Excerpts (9) and (10) respectively illustrate KLK in the utterance-initial and final position:

(9)   (NE and GB are watching TV while chatting. NE just killed a mosquito that was sitting on her hand.)

```
147 NE: iausiu,\
        darn.it
148 NE: kong tsiah kah angkiki.\
        DM   eat   to extremely.red
149 NE:..tsiah kah tsia tse  hueh.\
        eat   to so    much blood
--> 150 GB: gua chimmi   ka li kong li toh bo ai    thiann._
            1S  just.now to 2S say  2S TOH NG love listen
151 NE: gua tsaiiann li anne      pi    li pi    sann!\
        1S  know     2S this.way point 2S point what

NE (147–149): Darn it!. (The mosquito's stomach) is full of (my) blood!
GB (150): I've already told you, but you wouldn't listen.
NE (151): I knew you were pointing (at something), but I didn't know what you meant.
```

(10)   (BY is talking to AU about her gum disease and knee pain.)

```
46 BY: gua tsitma long e     e       huatiam anne.\
       1S  now      all  will will inflamed this.way
47 AU: ah li khan  khi ho lang   khuann la.\
       DM 2S hurry go  DP people look   PA
48 AU: he   kuie  long. kui   pai <M yachi hui  diao guang M> ne.\
       that whole all   whole row      tooth will fall empty     PA
49 AU: he   na giamtiong <M zheng pai yachi diao guang M>,\
       that if serious        whole row tooth fall empty
50 AU: ah li tsitma iau    etang tiliau khan  khi ho lang   khuann la.\
       DM 2S now     still can   cure  hurry go DP people look   PA
51 AU: he   li toh tsiahlau li toh. aupai li anne     kanna. tsiah muai
       that 2S TOH old       2S TOH later 2S this.way only   eat   congee
       toh ho   a  o?\
       TOH good PA PA
--> 52 AU: ah li he    iengiong  na bo kau    li toh hai le gua ka li kong.\
           DM 2S that nutrition if NG enough 2S TOH bad AP 1S  to 2S say
53 AU: ah li kha  kut  ah ahkoh iau   be ho      o?\
       DM 2S foot bone DM still still NG recover PA

((8 IUs omitted.))
```

```
62 AU: ah li toh ai  itieng   ai  ai  kiann ai  tsau la.\
       DM 2S TOH need definitely need need walk  need run  PA
```

```
BY (46): Now I tend to have inflammation in the gums.
AU (47–52): You should go to the dentist as soon as possible. (If you do not deal with
it soon,) you will lose a whole row of teeth and your problem will get too severe.
Please go to the dentist for treatment now before it is too late. Otherwise, when
you get old, you can eat nothing but congee. Then you are doomed due to malnutrition,
I'm telling you.
AU (53): And do your knees feel better now?
AU (62): But you still have to walk and run more.
```

In (9), the exchange of talk occurs when NE kills a mosquito that was biting her, and she protests about the blood that has been extracted. This triggers GB's accusation that NE did not follow her advice to watch the mosquito that was hovering around. The accusation, preceded by KLK, focuses on the obviousness of the fact and NE's ignorance of it. As the accusation is a face threat to the addressee, NE defends herself by saying that she did not notice it. As for the discussion on BY's health issues in (10), the speaker AU acts as a knowledgeable authority and experienced patient warning BY about losing her teeth and suffering from severe knee pain if she does not visit the dentist soon and regularly exercise. A strong sense of epistemic superiority is connoted in AU's speech. Collocating with the warning are expressions that denote the negative consequences of overlooking one's oral health, including *yachi hui tiau kuang* 'losing all teeth' (in Mandarin), *tsiahlau kanna tsia mue* '(can) only eat congee when old (because of lack of teeth)', *iengiong bokau* 'malnutrition', and *hai le* 'bad (teeth)'. Briefly speaking, KLK may convey the speaker's epistemic authority while conveying an additional speech act of warning, threat, accusation, and advice. All of these utterances possess the overtly expressed subject *gua* 'I'.

One point needs to be noted here regarding the KLK function assigned to (9). As stated in the introduction to Section 4, a certain bundle may possess multiple functions in a given context. The discussion in this section is reminiscent of the use of KLK in Extract (6), where KLK marks an obvious fact. It seems that KLK in (6) can be also treated as a speech act marker, particularly, an accusation. While this indicates the polyfunctionality of KLK, (6) differs from (9) in that KLK in (6) is preceded by the emphatic marker *toh*, forming the bundle TKLK that always marks an obvious fact that the addressee should have known earlier. Therefore, KLK in (6) is not classified as a speech act introducer while KLK in (9) is.

## 4.3 Discourse organizer: Facilitation of storytelling

Another function of KLK is to facilitate storytelling about a present or absent protagonist. A total of 21 cases of KLK (15%) are used to introduce an extended story. The majority of these KLKs take a 1st person singular *gua* as the subject (18/21,

85.7%) (Table 7), and the object type of the storytelling KLK is exclusively a clause (Table 8). As a discourse organizer, KLK prepares the ground for the speaker to move from a previous discourse topic to a new one with an anecdote that intrigues the addressee. Sequentially, the discourse-organizing KLK always occurs in a turn-initial or first position (Whalen and Zimmerman 1998; Heritage 2013). In terms of the information deployed, this LB involves direct knowledge of the truth of the story possessed by the speaker. By initiating the turn with an exclusive story, the speaker claims the right to the turn (Heritage and Raymond 2005; Clift 2006: 575) and dominates the floor with an upcoming speech.

**Table 7.** Subject type of KLK with the storytelling function

|   | 1p explicit | 1p implicit | 3p explicit | 3p implicit | Other | Total |
|---|---|---|---|---|---|---|
| # | 15 | 3 | 2 | 0 | 1 | 21 |
| % | 71% | 14% | 10% | 0% | 5% | 100% |

**Table 8.** Object type of KLK with the storytelling function

|   | NP | Clause | Zero | Total |
|---|---|---|---|---|
| # | 0 | 21 | 0 | 21 |
| % | 0% | 100% | 0% | 100% |

The storytelling KLK is usually accompanied by a time expression that sets the temporal frame of the narrative, e.g. *hit tsuna honn* 'at that time' or *tsuikin* 'recently'. Consider (11):

(11)   (SA and SB are sisters. They are talking about their sister SC, who has been suffering serious health problems. However, instead of guarding her health, SC has been worrying about other issues, such as house renting and property disputes.)

```
676 SA: tak   hang long beh  huanlo toh tioh  la.\
        every item all  want worry  TOH right PA
677 SB: ah!\
        EX
678 SB: ah toh kong hittsuna  beh  he   la._
        DM TOH say  that.time want that PA
679 SB: anne      la.\
        this.way PA
680 SB: ah gua si te ka mng kong,_
        DM 1S  CP AP to ask CM
681 SB: <Q ah li binhiong hia,_
            DM 2S PN       there
682 SB: kam  u,_
        dare have
683 SB: u=   hittsuna  kong kiosi beh  thi  tshu,_
        have that.time say  think want iron house
```

```
684 SB: ta     luei tsoo lang Q>.\
        build down rent people
```

((7 IUs omitted, in which the house renting topic continues.))

```
692 SB: henn la.\
        RT   PA
693 SB: ah tansi tsitma lang   beh  tsoo uan  kieng lang   tsoo la.\
        DM but   now    people want rent also pick  people rent PA
694 SB: [anne]    la.\
         this.way PA
695 SA: [hum].\
         RA
696 SA: hum hum.\
        RT   RT
--> 697 SB: atshun *gua ka li kong* honn,_
            PN     1S to 2S say   PA
698 SB: i  hittsuna  honn,_
        3S that.time PA
699 SA: hm.\
        RT
700 SB: hite sioko          honn,_
        that sue.each.other PA
701 SB: tshinke ittit        m  ho iong a.\
        in.law  continuously NG DP do  PA
702 SA: hm.\
        RT
```

((25 IUs omitted here, in which the story continues.))

```
728 SB: ah kong,_
        DM say
729 SB: ka   ma    kah  tsiok phainn thiann.\
        give scold till very  bad    hear
```

```
SA (676): (She) has been worrying about everything.
SB (677-684): Alas! She told me that (she had bought a land in Binhiong). So I asked
her, "Did you build a metal house for rent?")
SB (692-694): Right. However, now it is not easy to find a tenant.
SA (695-696): Yeah. yeah, yeah.
SB (697-698): Atshun, *you know what?* At the time (when she just bought the land),
SA (699): Hm.
SB (700-701): her in-laws wouldn't let her deal with the land dispute.
SA (702): Hm.
SB (728-729): And (the in-laws) spoke so ill of her.
```

This conversation occurs between SA and SB about their sister. Line 697 marks a departure from a previous topic about house renting to the report of a land dispute between SC and her in-laws. The utterance-final particle *honn*, a "negotiation marker", terminates KLK, signaling a "disruption in discourse coherence" and

facilitating a new turn of talk (Li 1999: 82) about a story which only SB knows. The vocative *tshun*, that is, SA's name, serves as a metalinguistic alerter (Biq 1991; Blum Kulka et al. 1989: 17, 276) engaging the addressee with the report of another issue that has been troubling SC. The sequence of interaction introduced by KLK extends over 38 IUs, from IU 697 to IU729, in which SB is the dominant speaker while SA only responds with minimal tokens.

An elaborate version of storytelling is illustrated by (12), with KLK introducing a narrative that extends over 100 IUs.

(12)   (XL is talking to YC about her alcoholism due to her husband's business
        failure. The following IUs are contributed by the same speaker, XL.)

```
762 XL: lan tongjian  e!
        1PI certainly VOC
763 XL:…lim [tsiu    honn],\
        drink alcohol PA

((8     IUs omitted here, about one's indulgence in drinking when under great
pressure.))

772 XL: ..ah tansi,_
          DM but
--> 773 XL: gua ka li kong,_
            1S  to 2S say
774 XL: ..i  tsuekin,_
         3S recently
775 XL: i  tsuekin,_
        3S recently
776 XL: annua li tsai o?\
        how    2S know PA
777 YC: …(1.4) heinn.\
               RT
778 XL: in phootongsia e  si   putsi lim   e kitiong tsit e  piengiu la.\
        3P usual.time  AC time often drink AC among   one CL friend  PA

 ((7 IUs omitted, about this drinking buddy's background.))

786 XL: …li tsit e  lang   honn,\
          2S this AC person PA
--> 787 XL: gua ka li kong,_
            1S  to 2S say
788 XL: i  tsitma toh si tshiamthau./
        3S now    TOH CP head.of.lottery.game
789 XL: …tsit e tsiok    <L ai   lim   e L>.\
         this AC extremely   love drink AC

((28 IUs omitted here, in which XL said that this drinking buddy is her husband's
college friend.))
```

```
818 XL: honn,\
        PA
819 XL: ..ah i  toh si te kau       tsit e  tshita._
           DM 3S TOH CP AP associate one  CL girlfriend


((20 IUs omitted here.))

839 XL: guan nng e  tsiok siak      e  m._ ((YC laughs at the same time.))
        1PE  two CL very   familiar AC PA
840 XL: …i  toh kong honn,\
         3S TOH say   PA
841 XL: toh khahtianue,_
        TOH call
842 XL: e!\
        VOC
843 XL: <Q<M shuchen shuchen M>,_
            PN       PN
844 XL: li kuelai    la,\
        2S come.over PA
845 XL: kinle.\
        hurry
846 XL: o!\
        EX
847 XL: lua  tshenntshau le la Q>.\ ((YC laughs at the same time.))
        very abundant     AP PA
848 XL: …tsio   gua lim._
         invite 1S  drink


((25 IUs omitted, with the story continued.))

874 XL: gua toh khia kue .\
        1S  TOH ride pass
875 XL: …ah kiatko      lim,_
         DM as.a.result drink
876 XL: toh ka   i  khaisi nng e  toh khaisi,_
        TOH with 3S begin  two NM TOH begin
877 XL: honn,\
        PA
878 XL: kanna guan nng e  ti hia ,_
        only  1S   two NM at there
```

XL (762–763): Sure, speaking of drinking,
XL (772–778): But, *you know what*? Recently, one of (my husband's friends) who
frequently went drinking with him,
XL (786–789): This person, who is the bill collector of a gambling agency, loves
drinking.
XL (818–819): Okay. Now he is having an affair.
XL(839–848): We know each other very well. He would sometimes call me and say, "Hey,
*shuchen shuchen,* come on! Hurry! Join us for dinner. There is plenty to eat!" (He) was
inviting me to go drinking with him (and his mistress).
XL (874–878): I would then ride (my scooter to the restaurant) and we started to drink.
Okay? Only the two of us were there.

Line 773 in (12) exemplifies a shift in topic from XL's alcoholism because of business failure to a story about drinking with a friend. This friend was having an affair and liked to invite XL for a drink with his mistress. Two occurrences of *gua*-KLK can be found in this spate of conversation providing impetus to the storyline. Both of them introduce information and encounters the speaker has experienced. The first acts as a prelude to the main story of the affair, which intrigues the hearer, by introducing the protagonist with the rhetorical frame (*annua li tsai o* 'You know what?') (line 776) (see also Schiffrin 1987; Tao 2003). The second re-occasions the storytelling, especially the characterization of the protagonist, based on the primary content regarding XL's interaction with this friend – going on a binge with his mistress and him.

In this section, we have argued that KLK can be employed in the first position to imply that the speaker has epistemic authority over the truth of the forthcoming story. Most of them have an overt subject *gua* 'I' preceding them, forming the bundle *gua*-KLK. It serves as a "lead-in" to preface the speaker's move from a previous topic to the next one with a narrative (cf. House 1996: 234; Conrad and Biber 2004: 66, about *I'll tell you what*). The collocation of various discourse strategies – temporal expressions, vocatives, and rhetorical questions – further aids the storytelling process. KLK differs from other topic-introducing discourse markers in Chinese (e.g. *na* in Huang 1999: 83) in that it indicates the speaker's epistemic primacy regarding the information provided, while the addressee is primarily a listener with minimal responses.

## 4.4 Summary

The above discussion has illustrated that KLK acts as a recurrent discourse bundle employed by the speaker to imply epistemic "authority" (Heritage and Raymond 2005; Stivers et al. 2011; Quan and Wu 2014) regarding the veracity of the predication at a first position or a responsive position, to express speaker's stance or to initiate a newsworthy narrative. The epistemic primacy characterizing KLK approximates the "privileged access" discussed in Labov and Fanshel (1977: 62) about A-events, events which are known to the speaker but not necessarily to the listener. Behaviors representing such events are emotion expressing and information giving (p. 61), which accord with the stance expressing and storytelling functions of KLK.

Concerning the collocations of KLK, in the stance-marking use, a variety of expressions co-occur with KLK in the same or at an adjacent IU to increase the specificity of the evaluative meaning. These collocations include the negative markers (*m, bo,* and *be*) used to challenge the previous speaker's comment, the adverbial *toh* for emphasis (Lin 1996), the contrastive marker *kisit* 'in fact' to correct

the addressee's assumption (Chang 2008), the marker of exclusiveness *kanna* 'only', the aspect marker *ikieng* 'already', the extreme degree modifiers *ukau* 'very, extremely', or words conveying distaste or criticism *pui* 'obese, *ti* 'pig', and *hai le* 'bad (teeth)'. As for the story-introducing KLK, the main collocations are temporal adverbials which set the time frame for the upcoming talk.

A few similarities and differences can be observed between KLK and its corresponding forms in Mandarin Chinese and in English. Compared to Mandarin, KLK invariably occurs in different contexts to perform a constellation of meanings characterized by its implication of epistemic authority: stance marking, speech act marking, and news reporting. Although Quan and Yu (2014) have noted the "epistemic primacy" (Hayano 2011) of *wo gen ni shuo*, and the speech act functions are addressed in Dong (2010), the current study has provided a systematic analysis with corpus evidence showing its frequency and illustrated how different sequential positions and turn types correlate with the meanings of KLK, how the functions are related to epistemic primacy, and what collocations partake in the making of the meanings of this bundle. Comparison with the English equivalents such as *Let me tell you, tell you what*, etc. shows that KLK shares the structural property of straddling two grammatical units. However, KLK enjoys much higher frequency than its English counterparts, ranking top among the LBs in our TSM corpus.[5] In addition, we have provided a more detailed discussion of the sequentiality and discourse-pragmatic functions of KLK.

In short, KLK performs several functions that cannot be summarized by simply treating the bundle as a V+say sequence. Rather than contributing much semantic information to the predicate as a whole (See also Scheibmann 2002: 48, on *say* in English), KLK serves as a recurrent phraseological unit of epistemic stance for speakers to manage information to which they have direct access. We argue that KLK is developing into a discourse marker via pragmatic strengthening (Traugott and Dasher 2002: 35) to fulfill interactional functions. Further evidence in support of this argument can be garnered from the phonological, prosodic and syntactic properties of KLK, which demonstrate the features of a discourse marker (see also Section 3). We turn to these features in the following.

## 4.5   Fixedness and automaticity: Emergence of KLK as a discourse marker

In terms of the fixedness of the bundle, we discovered that the sequence KLK is always processed as a holistic chunk. In other words, whenever *ka, li*, and *kong* co-occur with each other, they are always immediately adjacent to each other without

---

**5.** Although *I'll tell you what/I'll tell you* is a frequent LB in English (Biber et al. (1999: 1002), there is no corpus study showing that these two LBs rank top among LBs in spoken discourse.

any intervening lexeme/modifier. In addition, there is a great phonological reduction of the object pronoun *li* 'you' and the attrition of the vowel [a] in the coverb *ka*, yielding the frequent phonetic realization *keikong*. These prosodic features suggest that *kalikong* is "entrenched" in the memory and processed automatically by the speaker (Langacker 1987; Van Lancker 1987; Wray 2002; Schmid 2017; Langacker 2017) as a prefabricated unit. Along with its high frequency, we argue that *ka li* and *kong* are "glued together" (Hyland 2012: 15) as an emerging structural unit used by speakers to avoid disfluencies and gain time for their speech during real-time language processing (Nattinger 1988: 76) when speakers wish to express his stance or mark news while showing their epistemic primacy with regard to the information conveyed.

Apart from the cognitive dimension of the use of KLK, there are other syntactic and prosodic properties that suggest that KLK is developing into a discourse marker. First, a significant number of KLKs occur as a solitary IU (22/139, 16%). If we include a lone KLK followed by an utterance final particle, the number is almost twice as frequent (42/139). The cognitive grounding of this LB along with its textual and interactional functions discussed above are associated with the characteristics of a "regulatory" IU such as *well, mhm, let me see*, etc, which are used by speakers to "[regulate] interaction or information flow" (Chafe 1994: 63). These features also coincide largely with the discourse markers discussed above, i.e., they often constitute an independent IU. Second, there is a high frequency of utterance-final particles immediately after KLK (36/139, 25%), e.g. *honn* and *la*. The use of a particle intervening the bundle and the following NP/clause manifests the weak relation between the verb *kong* and its object complement and hence the low transitivity of the verb. A third and even stronger fact that supports the role of KLK as a discourse marker concerns the structural type of the object complement. That is, a large majority the object complements of KLK are clausal propositions (106/139, 76%, including those following KLK with an IU-final particle). An overt post-*kong* NP object which is typical of verb of higher transitivity constitutes only a small number (13/139, 9%) and is only found with the typical telling function (cf. Table 4). In fact, the literature is replete with evidence showing the low transitivity of saying verbs in conversation, e.g. the arguments by Altenberg (1998) and Biber et al. (1999) that say usually carries low semantic content and the discussions by Hopper (1991), Thompson and Hopper (2001), Scheibman (2002), and Huang (2013) that epistemic/evidential clauses or verbs of low transitivity account for the majority of predicates in conversation. The information packaging of using a shorter, low-content matrix clause (KLK) followed by a clause which carries the main semantic weight is also in line with the "end-weight principle" (Quirk et al. 1985: 323) or "processibility principle" (Leech 1983: 64–6, Kaltenböck 2015: 120). Such "pleonastic matrix clauses" are important resources for grammaticalization

of discourse markers (Kaltenböck 2015: 125). Table 9 presents the features possessed by KLK in terms of the defining qualities of a discourse marker. Since each feature may only reflect a prevailing rather than an exclusive quality (Jucker and Ziv 1998: 4), a star (*) is given when a feature is only predominant in a certain functional type, while two stars (**) indicate an omnipresent feature with a given use:

**Table 9.** Prototypicality of KLK as a Discourse Marker

| Functional type<br>Feature | Typical telling | Stance marking | Discourse organizing | TKLK |
|---|---|---|---|---|
| (a) Phonological reduction | * | * | * | * |
| (b) As a separate tone group | ** | ** | ** | ** |
| (c) Not a traditional word class | ** | ** | ** | ** |
| (d) Sentence-initial position | * | * | ** | * |
| (e) Outside main clause/with loose connection | ** | ** | ** | ** |
| (f) Optionality | * | ** | * | ** |
| (g) With little/no propositional meaning | ** | ** | ** | ** |
| (h) Multifunctionality | * | * | * | * |
| (i) Typical of oral discourse/informality | ** | ** | ** | ** |
| (j) High frequency | ** | ** | ** | ** |
| (k) Being stylistically stigmatized | | | | |
| (l) Typical of female speech | | | | |

As clearly shown in Table 9, except for KLKs with the typical telling function, the other uses of KLK display most of the features of a discourse marker, especially (a-g), which are crucial tests for a discourse marker. However, since not all the features are present in these non-typical uses, we argue that such KLKs are still in the process of grammaticalizing into a discourse marker.

## 5.   Discussion and Conclusion

When a person makes a statement, "the speaker opens himself up to the possibility that the intended recipients" will not recognize and validate his talk (Goffman 1967: 37, cited in Heritage and Raymond 2005: 14). Therefore, one of the primary concerns for speakers in the conversations is for them to manage the "rights and responsibilities related to knowledge and information" so that they are granted the right to make assessments, narrations, and news announcements (Heritage and Raymond 2005: 15). A language may provide a variety of resources for speakers to

accomplish such goals (Clift 2006). In TSM, the LB *kalikong* 'to-you-say' figures as the most frequent LB, affording the speakers with the means to express their epistemic authority over the recipients about knowledge or experiences to which they have exclusive access. The discussion has shown that it behaves more like a discourse marker for expressing stance and for organizing discourse. Collocation of KLK with the emphatic marker *toh* further triggers the grammaticalization of the four-word LB TKLK to encode an extreme stance. These changes are effected by their frequent utilization in natural conversation (Langacker 1987; Bybee 2007; Schmid 2017) and phonological reduction along with the demotion of the transitivity of the saying verb and of the referential status of the subject and the object. The collapse of the syntactic boundary among the components of KLK, apart from its frequent co-occurrence with an utterance final particle, demonstrates that a majority of KLKs should be treated as a holistic prosodic unit whose componential meaning is relegated to the background.

Compared to Mandarin, which has several expressions for the meaning of 'Let me tell you', KLK is the sole LB in TSM with this meaning. While different Mandarin equivalents slightly differ in their main interactional functions (see Section 2.2), KLK encompasses meanings that are distributed in its various synonyms in Mandarin. Sequential positions, turn types, and collocations, therefore, are important for the interpretation of KLK in conversations. The all-inclusive nature of this LB may also be the contributing factor for its high frequency in TSM discourse.

As for the distributions of KLK and its English counterparts (cf. Section 4.4), speaker relationships might play a role. As we have discussed in Section 4, KLK frequently carries an authoritative tone, which is congruent with the general tone implied in advice-giving, a social action regarded by the Chinese as "a form of support provision" (Feng 2015: 1143), especially among closely related speakers (Feng and Magen 2016). In the case of stance-marking KLKs, among those collocating with a piece of advice, 57% occur in conversation with close relatives and 38% with friends. Likewise, in a cross-cultural study of advice giving by Russians and European Americans (Chentsova-Dutton and Vaughn 2012), although collectivism-based Russians are more inclined to giving advice than autonomy-valuing European Americans in most categories of relationships (e.g. colleagues, supervisors, etc.), both groups demonstrate similar preference for giving advice to friends (pp. 693–694). These results suggest that distance and social dominance rather than language per se may also play a role.[6] While most of the speakers in our corpus are friends or acquaintances, since KLK also serves other purposes than advice-giving, no claim is made here about the direct relation between KLK

---

**6.** We thank one of the reviewers for pointing out this issue.

frequency and this speech act. However, the intriguing relationship among illocutionary speech acts, speaker relationships, and forms of support building in different cultures is worth further investigation.

To conclude, in this paper, we have provided a systematic analysis of the cognitive grounding, the discourse-pragmatic functions, and the structural and prosodic qualities of the recurrent LB KLK in TSM conversations. In future studies, it would be interesting for linguists to investigate how speakers employ other linguistic resources to manage their information and knowledge during conversation and how cultural patterns influence the distributions and usages of LBs in conversation.

## Acknowledgements

## References

Altenberg, Bengt. 1998. "On the Phraseology of Spoken English: Recurrent Word-Combinations." In *Phraseology: Theory, Analysis, and Applications*, ed. by Anthony Paul Cowie, 101–122. Oxford: Clarendon Press.

Bannard, Colin, and Elena Lieven. 2009. "Repetition and Reuse in Child Language Learning." In *Formulaic language: Vol. 2. Acquisition, Loss, Psychological Reality, Functional Explanations*, ed. by R. Corrigan, E. Moravcsik, H. Ouali, and K. Wheatley, 297–321. Amsterdam: John Benjamins.  doi: 10.1075/tsl.83.04ban

Bannard, Colin, and Danielle Matthews. 2008. "Stored Word Sequences in Language Learning: The Effect of Familiarity on Children's Repetition of Four-word Combinations." *Psychological Science* 19: 241–248.  doi: 10.1111/j.1467-9280.2008.02075.x

Bardovi-Harlig, Kathleen. 1999. "Exploring the Interlanguage of Interlanguage Pragmatics: A Research Agenda for Acquisitional Pragmatics." *Language Learning* 49: 677–713.  doi: 10.1111/0023-8333.00105

Biber, Douglas, and Susan Conrad. 1999. "Lexical Bundles in Conversation and Academic Prose." In *Out of Corpora*, ed. by Hilde Hasselgård, and Signe Oksefjell, 181–190. Amsterdam: Rodopi.

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

Biber, Douglas, Susan Conrad, and Viviana Cortes. 2003. "Lexical Bundles in Speech and Writing: An Initial Taxonomy." In *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*, ed. by Andrew Wilson, Paul Rayson, and Tony McEnery, 71–92. Frankfurt/Main: Peter Lang.

Biber, Douglas, Susan Conrad, and Viviana Cortes. 2004. "If You Look at …: Lexical Bundles in University Teaching and Textbooks." *Applied Linguistics* 25 (3): 371–405.
doi: 10.1093/applin/25.3.371

Biq, Yung-O. 1991. "The Multiple Uses of the Second Person Singular Pronoun *Ni* in Conversational Mandarin." *Journal of Pragmatics* 16 (4): 307–321.
doi: 10.1016/03782166(91)90084B

Blum-Kulka, Shoshana, Juliane House, and Gabriele Kasper. 1989. *Cross-cultural Pragmatics: Requests and Apologies*. New Jersey: Ablex Publishing Co.

Brinton, Laurel J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/New York: Mouton de Gruyter.   doi: 10.1515/9783110907582

Butler, Christopher S. 1997. "Repeated Word Combinations in Spoken and Written Text: Some Implications for Functional Grammar." In *A Fund of Ideas: Recent Development in Functional Grammar*, ed. by Christopher S. Butler, John H. Connolly, Richard A. Gatward, and Roel M. Vismans, 60–77. Amsterdam: IFOTT, University of Amsterdam.

Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.   doi: 10.1093/acprof:oso/9780195301571.001.0001

Carter, Ronald, and Michael McCarthy. 2006. *Cambridge Grammar of English*. Cambridge: Cambridge University Press.

Chafe, Wallace. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

Chang, Miao-Hsia. 2008. "Discourse and Grammaticalization of Contrastive Markers in Taiwanese Southern Min: A Corpus-based Study." *Journal of Pragmatics* 40: 2114–2149.
doi: 10.1016/j.pragma.2008.04.009

Chen, Li-jun. 2010. "Huayu Biaoji "Wo Gei Ni Shuo" de Yanbian Guocheng [Grammaticalization of the Discourse Marker "Wo Gei Ni Shuo" in Chinese]." *Zhejiang Shifan Daxue Xuebao (Shehui Kexue)* [Journal of Zhejiang Normal University (Social Sciences)] 171: 84–87.

Cheng, Robert L., and Susie S. Cheng. 1977. *Taiwan Fujian Hua de Yuyin Jiegou Ji Biaoyin Fa* [Phonological Structure and Romanization of Taiwanese Hokkian]. Taipei: Student Book Co.

Cheng, Winnie, Chris Greaves, and Martin Warren. 2006. "From N-gram to Skipgram to Congram." *International Journal of Corpus Linguistics* 11 (4): 411–433.
doi: 10.1075/ijcl.11.4.04che

Chentsova-Dutton, Yulia E., and Alexandra Vaughn. 2012. "Let Me Tell You What to Do: Cultural Differences in Advice-Giving." *Journal of Cross-Cultural Psychology* 43 (5): 687–705.   doi: 10.1177/0022022111402343

Clift, Rebecca. 2006. "Indexing Stance Reported Speech as an Interactional Evidential." *Journal of Sociolinguistics* 10 (5): 569–595.   doi: 10.1111/j.1467-9841.2006.00296.x

Conrad, Susan, and Douglas Biber. 2004. "The Frequency and Use of Lexical Bundles in Conversation and Academic Prose." In *The Corpus Approach to Lexicography*, ed. by W. Teubert, and M. Mahlberg (Eds. of Thematic Part), Lexicographica: International Annual for Lexicography 20: 56–71.

Cortes, Viviana. 2004. "Lexical Bundles in Published and Student Disciplinary Writing: Examples from History and Biology." *English for Specific Purposes* 23: 397–423. doi: 10.1016/j.esp.2003.12.001

Cortes, Viviana, and Eniko Csomay. 2007. "Positioning Lexical Bundles in University Lectures." In *Spoken Corpora in Applied Linguistics*, ed. by Mari Carmen Campoy, and Maria Jose Luzon, 57–76. Frankfurt am Main: Peter Lang.

Cowie, Anthony Paul. 1998. "Introduction." In *Phraseology: Theory, Analysis, and Applications*, ed. by Anthony Paul Cowie, 1–20. Oxford: Clarendon Press.

Coxhead, Averil, and Pat Byrd. 2007. "Preparing Writing Teachers to Teach the Vocabulary and Grammar of Academic Prose." *Journal of Second Language Writing* 16: 129–147. doi: 10.1016/j.jslw.2007.07.002

De Smet, Hendrik. 2017. "Entrenchment Effects in Language Change." In *Entrenchment and the Psychology of Language Learning: How we Reorganize and Adapt Linguistic Knowledge*, ed. by Hans-Jörg Schmid, 75–99. Washington, D.C.: American Psychological Association; Berlin: Walter de Gruyter GmbH.

Dong, Xiufang. 2010. "Laiyuan yu Wanzheng Xiaoju de Huayu Biaoji "Wo Gaosu Ni" [A Discourse Marker Derived from Clausal Form: Wo Gao Su Ni]." *Yuyan Kexue* [Linguistic Sciences] 46: 279–286.

Du Bois, John, Stephan Schuetze-Coburn, Susanna Cumming, and Danae Paolino. 1993. "Outline of Discourse Transcription." In *Talking Data: Transcription and Coding in Discourse Research*, ed. by Jane Anne Edwards, and Martin D. Lampert, 45–89. Hillsdale, NJ: Lawrence Erlbaum.

Ellis, Nick C. 2012. "Formulaic Language and Second Language Acquisition: Zipf and the Phrasal Teddy Bear." *Annual Review of Applied Linguistics* 32: 17–44. doi: 10.1017/S0267190512000025

Erman, Britt, and Beatrice Warren. 2000. "The Idiom Principle and the Open-choice Principle." *Text* 20: 29–62.

Feng, Bo, and Eran Magen. 2016. "Relationship Closeness Predicts Unsolicited Advice Giving in Supportive Interactions." *Journal of Social and Personal Relationships* 33 (6): 751–767.

Feng, Hairong. 2015. "Understanding Cultural Variations in Giving Advice among Americans and Chinese." *Communication Research* 42: 1143–1167. doi: 10.1177/0093650213486668

Foster, Pauline. 2001. "Rules and Routines: A Consideration of Their Role in the Task-based Language Production of Native and Non-native Speakers." In *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, ed. by Martin Bygate, Peter Skehan, and Merrill Swain, 75–93. Harlow: Longman.

Gan, Min. 2012. "*Zuo Wei Huayu Biaoji de "Wo Gen Ni Shuo"* [*Wo Gen Ni Shuo* as a Discourse Marker]." *Wenxue Jiaoyu* [Literature Education] 2012 (3): 148–50.

Goffman, Erving. 1967. *Interaction Ritual: Essays in Face to Face Behavior*. Garden City, NY: Doubleday.

Granger, Slviane. 1998. "Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae." In *Phraseology: Theory, Analysis, and Applications*, ed. by Anthony Paul Cowie, 145–160. Oxford: Clarendon Press.

Greenbaum, Sidney, and Jan Svartvik. 1990. "The London-Lund Corpus." In *The London-Lund Corpus of Spoken English: Description and Research* (Lund Studies in English 82), ed. by Jan Svartvik, 11–45. Lund: Lund University Press

Hayano, Kaoru. 2011. "Claiming Epistemic Primacy: Y-marked Assessments in Japanese." In *The Morality of Knowledge in Conversation*, ed. by Tanya Stivers, Lorenza Monda, and Jakob

Steensig, 58–81. Cambridge, UK: Cambridge University Press.
doi: 10.1017/CBO9780511921674.004

Heritage, John. 2002. "*Oh*-prefaced Responses to Assessments: A Method of Modifying Agreement/Disagreement." In *The Language of Turn and Sequence*, ed. by Cecilia Ford, Barbara Fox, and Sandra A. Thompson, 196–224. Oxford: Oxford University Press.

Heritage, John. 2013. "Turn-initial Position and Some of Its Occupants." *Journal of Pragmatics* 57: 331–337. doi: 10.1016/j.pragma.2013.08.025

Heritage, John, and Geoffrey Raymond. 2005. "The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction." *Social Psychology Quarterly* 68 (1): 15–38. doi: 10.1177/019027250506800103

Hopper, Paul. 1991. "Dispersed Verbal Predicates in Vernacular Written Narrative." *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Grammar of Event Structure*, 402–413. Berkeley: University of California. doi: 10.1075/slcs.36.04hop

House, Juliane. 1996. "Developing Pragmatic Fluency in English as a Foreign Language: Routines and Metapragmatic Awareness." *Studies in Second Language Acquisition* 18 (2): 225–252. doi: 10.1017/S0272263100014893

Hrisonopulo, Katherine. 2008. "Beyond Reference and Designation: On Interactive Implications of the Pronoun *I* in English." *Lodz Papers in Pragmatics* 4 (2): 277–292. doi: 10.2478/v10016-008-0006-2

Huang, Shuanfan, 1999. "The Emergence of a Grammatical Category *Definite Article* in Spoken Chinese." *Journal of Pragmatics* 31 (1): 77–94. doi: 10.1016/S0378-2166(98)00052-6

Huang, Shuanfan. 2013. *Chinese Grammar at Work*. Amsterdam: John Benjamins. doi: 10.1075/scld.1

Hyland, Ken. 2008. "As Can Be Seen: Lexical Bundles and Disciplinary Variation." *English for Specific Purposes* 27: 4–21. doi: 10.1016/j.esp.2007.06.001

Hyland, Ken. 2012. "Bundles in Academic Discourse." *Annual Review of Applied Linguistics* 32: 150–169. doi: 10.1017/S0267190512000037

Jucker, Andreas H., and Yael Ziv. 1998. "Discourse Markers: Introduction." In *Discourse Markers: Description and Theory*, ed. by Andreas H. Jucker, and Yael Ziv, 1–12. Amsterdam: John Benjamins. doi: 10.1075/pbns.57.03juc

Kaltenböck, Gunther. 2015. "Processibility." In *Corpus Pragmatics: A Handbook*, ed. by Karin Aijmer, and Christoph Rühlemann, 117–142. Cambridge, UK: Cambridge University Press. doi: 10.1016/j.pragma.2016.06.001

Kim, YouJin. 2009. "Korean Lexical Bundles in Conversation and Academic Texts." *Corpora* 4 (2): 135–165. doi: 10.3366/E1749503209000288

Labov, William, and David Fanshel. 1977. *Therapeutic Discourse: Psychotherapy as Conversation*. New York: Academic Press.

Langacker, Ronald W. 1987. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.

Langacker, Ronald W. 2017. "Entrenchment in Cognitive Grammar." In *Entrenchment and the Psychology of Language Learning: How we Reorganize and Adapt Linguistic Knowledge*, ed. by Hans-Jörg Schmid, 39–56. Washington, D.C.: American Psychological Association; Berlin: Walter de Gruyter GmbH.

Leech, Geoffrey. 1983. *Principles of Pragmatics*. London: Longman.

Li, Ing C., 1999. *Utterance-Final Particles in Taiwanese: A Discourse-Pragmatic Analysis*. Taipei: The Crane Publishing Co.

Lin, Min-ching. 1996. *Discourse Functions of TOH and CHIAH in Taiwanese*. Unpublished M. A. Thesis. National Taiwan Normal University.

Lindemann, Stephanie, and Anna Mauranen. 2001. "It's Just Real Messy: The Occurrence and Function of *Just* in a Corpus of Academic Speech." *English for Specific Purposes* 20 (Supplement): 459–475.  doi: 10.1016/S0889-4906(01)00026-6

Liu, Dilin. 2012. "The Most Frequently-used Multi-word Constructions in Academic Written English: A Multi-corpus Study." *English for Specific Purposes* 31: 25–35.  doi: 10.1016/j.esp.2011.07.002

Ma, Guanghui. 2009. "Yingyu Zhuanye Xuesheng Eryu Xianshi Xiezuo Zhong de Cikuai Yanjiu [Lexical Bundles in L2 Timed Writing of English Majors.]" *Waiyu Jiaoxue yu Yanjiu* [Foreign Language Teaching and Research] 41 (1): 54–60.

Nattinger, James. 1988. "Some Current Trends in Vocabulary Teaching." In *Vocabulary and Language Teaching*, ed. by Ronald Carter, and Michael McCarthy, 62–82. London and New York: Longman.

Nesi, Hilary, and Helen Basturkmen. 2006. "Lexical Bundles and Discourse Signalling in Academic Lectures." *International Journal of Corpus Linguistics* 11 (3): 283–304.  doi: 10.1075/bct.17.03nes

Östman, Jan-Ola. 1981. '*You Know*': A Discourse-Functional Study*. Amsterdam: John Benjamins.  doi: 10.1075/pb.ii.7

Partington, Alan, and John Morley. 2004. "From Frequency to Ideology: Investigating Word and Cluster/bundle Frequency in Political Debate." In *Practical Applications in Language and Computers – PALC 2003*, ed. by Barbara Lewandowska-Tomaszczyk, 179–192. Frankfurt am Main: Peter Lang.

Pawley, Andrew, and Frances Hodgetts Syder. 1983. "Two Puzzles for Linguistic Theory: Native-like Selection and Native-like Fluency." In *Language and Communication*, ed. by Jack C. Richards, and Richard W. Schmidt, 191–226. London: Longman.

Polio, Charlene. 2012. "Editor's Introduction." *Annual Review of Applied Linguistics* 32: vi–vii.  doi: 10.1017/S0267190512000116

Pomerantz, Anita. 1984. "Agreeing and Disagreeing with Assessments: Some Features of Preferred/dispreferred Turn Shapes." In *Structures of Social Action: Studies in Conversation Analysis*, ed. by J. Maxwell Atkinson, and John Heritage, 57–101. Cambridge, UK: Cambridge University Press

Quan, Tong, and Yu Guo-dong. 2014. "Zhong Ri "Zhishi Youxian" Pingjia Bijiao Yanjiu – Yi Huayu Biaji "Wo Gen Ni Jiang" han "Yo" Wei Li" [Comparative Study on Epistemic Primacy" in the Chinese and Japanese Language – Case Study of "I'm telling you" and "you"]." *Studies in the Philosophy of Science and Technology* 31(3): 38–44.

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

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Salazar, Danica. 2014. *Lexical Bundles in Native and Non-native Scientific Writing: Applying a Corpus-based Study to Language Teaching* (Vol. 65). Amsterdam/Philadelphia: John Benjamins.  doi: 10.1075/scl.65

Scheibman, Joanne. 2002. *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation*. Amsterdam/Philadelphia: John Benjamins.  doi: 10.1075/sidag.11

Schiffrin, Deborah. 1987. *Discourse Markers*. New York: Cambridge University Press.
doi: 10.1017/CBO9780511611841

Schmid, Hans-Jörg. 2017. "A Framework for Understanding Linguistic Entrenchment and Its Psychological Foundations." In *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge*, ed. by Hans-Jörg Schmid, 9–38. Washington, D.C.: American Psychological Association; Berlin: Walter de Gruyter GmbH.

Simpson-Vlach, Rita, and Nick C. Ellis. 2010. "An Academic Formulas List: New Methods in Phraseology Research." *Applied Linguistics* 31: 487–512. doi: 10.1093/applin/amp058

Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sorhus, Helen B. 1977. "To Hear Ourselves – Implications for Teaching English as a Second Language." *English Language Teaching Journal* 31 (3): 211–221. doi: 10.1093/elt/XXXI.3.211

Stenström, Anna-Brita. 2004. "What Is Going on between Speakers." In *Corpora and Discourse*, ed. by Alan Partingon, John Morley, and Louann Haarman, 259–283. Bern: Peter Lang.

Stivers, Tanya, Lorenza Monda, and Jakob Steensig. 2011. "Knowledge, Morality and Affiliation in Social Interaction." In *The Morality of Knowledge in Conversation*, ed. by Tanya Stivers, Lorenza Monda, and Jakob Steensig, 3–24. Cambridge, UK: Cambridge University Press.
doi: 10.1017/CBO9780511921674.002

Stubbs, Michael. 1983. *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. Chicago: The University of Chicago Press.

Tao, Hongyin. 2003. "Cong Yuyin Yufa han Huayu Tezheng Kan "Zhidao" Geshi Zai Tanhua Zhong de Yanhua [Phonological, Grammatical, and Discourse Evidence for the Emergence of *Zhidao* Constructions]." *Zhongguo Yuwen* [Chinese Language] 295: 291–302.

Tao, Hongyin. 2015. "Profiling the Mandarin Spoken Vocabulary Based on Corpora." In *Oxford Handbook of Chinese Linguistics*, ed. by Chaofen Sun, and William S-Y Wang, 336–347. Oxford: Oxford University Press.

Taylor, John R. 2012. *The Mental Corpus: How Language Is Represented in the Mind*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199290802.001.0001

Thompson, Sandra A., and Paul Hopper. 2001. "Transitivity, Clause Structure, and Argument Structure: Evidence from Conversation." In *Frequency and the Emergence of Linguistic Structure*, ed. by Joan Bybee, and Paul Hopper, 27–56. Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/tsl.45.03tho

Thompson, Sandra A., and Hongyin Tao. 2010. "Conversation, Grammar, and Fixedness: Adjectives in Mandarin Revisited." *Chinese Language and Discourse* 1 (1): 25–52.
doi: 10.1075/cld.1.1.01tho

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

Tseng, Chin-chin. 1997. "Taiyu Duanci Yuanze Taolun [Discussion of Taiwanese Word Segmentation Principles]." In *Taiyu Wenxue Chubanwu Shouji, Mulu, Xuandu, Bianji Jihua Jiean Baogao* [Project Report of the Collecting, Cataloging, and Editing of Taiwanese Literature Publications], 45–72. Taipei: Council for Cultural Affairs.

Van Lancker, D. 1987. "Nonpropositional Speech: Neurolinguistic Studies." In *Progress in the Psychology of Language*, Vol III, ed. by Andrew W. Ellis, 49–118, Hillsdale, NJ: Lawrence Erlbaum.

Wei, Naixing. 2007. "Zhongguo Xuesheng Yingyu Kouyu de Duanyuxue Tezheng Yanjiu – COLSEC Yuliaoku de Cikuai Zhengju Fenxi [Phraseological Characteristics of Chinese Learners' Spoken English: Evidence of Lexical Chunks from COLSEC.]" *Xiandai Waiyu* [Modern Foreign Languages] 30 (3): 281–291.

Whalen, Jack, and Don H. Zimmerman. 1998. "Observations on the Display and Management of Emotion in Naturally Occurring Activities: The Case of "Hysteria" in Calls to 9-1-1." *Social Psychology Quarterly* 61 (2): 141–159.  doi: 10.2307/2787066

Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge, UK: Cambridge University Press.  doi: 10.1017/CBO9780511519772

Wray, A., and M. R. Perkins. 2000. "The Functions of Formulaic Language: An Integrated Model." *Language and Communication* 20: 1–28.  doi: 10.1016/S0271-5309(99)00015-4

## Appendix I.  Transcription notations

| | | | | |
|---|---|---|---|---|
| – | truncated intonation unit | … | medium pause (0.3–0.6 second) |
| [] | speech overlap | .. | short pause (<0.3 second) |
| . | final | @ | laughter |
| , | continuing | % | glottal stop |
| _ | level | <E   E> | code-switching to English |
| ? | appeal | <M   M> | code-switching to Mandarin |
| ! | exclamation | <L   L> | loudness |
| ^ | primary accent | <J   J> | code-switching to Japanese |
| = | lengthening | <F   F> | fast tempo |
| …(N) | long pause (>0.7 second) | <P   P> | piano: soft |
| (0) | latching | <Q   Q> | quotation quality |
| (()) | transcriber's notes | | |

## Appendix II.  Abbreviations

| | | | |
|---|---|---|---|
| AC | associative marker | NM | nominalizer |
| AP | aspect marker | PA | (utterance final) particle |
| CL | classifier | PS | passive marker |
| CM | complementizer | PN | proper name |
| CP | copula | TOH | the adverbial particle toh |
| CV | coverb | VOC | vocative |
| DC | directional complement | 1s | 1st person singular |
| DM | discourse marker | 1PI | 1st person plural inclusive |
| DP | disposal marker | 1PE | 1st person plural exclusive |
| EX | exclamatory marker | 2s | 2nd person singular |
| NG | negative marker | 3s | 3rd person singular |

*Corresponding author's address*

Miao-Hsia Chang
National Taiwan Normal University
Department of English
162, Sec. 1
Heping E. Road
Da-an District
Taipei 106
Taiwan

mhchang@ntnu.edu.tw