The Non-deictic Use of Demonstratives in Conversations and Interpreted Speeches in Contemporary Hong Kong Cantonese

Oi Yee Kwong
Department of Translation
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
oykwong@arts.cuhk.edu.hk

Abstract

This study aims at attesting and assessing the usage and functional shift of demonstratives in contemporary Hong Kong Cantonese. Preliminary results from a corpus-based comparison show that while demonstratives are still predominantly used for their deictic and anaphoric functions in face-to-face conversations, an obvious deviation is observed in public discussions and interpreted speeches, particularly for the distal demonstrative, which is used far more often in a way similar to pause fillers in linguistic planning and speech production. Potential factors underlying the phenomenon and implications on language teaching and interpreter training are discussed.

1 Introduction

Demonstratives in a language, typically "this" and "that" as in English, are supposed to be used for deictic purposes conventionally. For Hong Kong Cantonese, however, it is observed at least in some speech situations that demonstratives tend to exhibit more non-deictic usages, unlike the normal deictic and anaphoric functions specified in most Cantonese grammars. This study aims at attesting this phenomenon with a corpus-based approach and further investigating the functional changes of demonstratives with respect to different speech situations. In particular, we compare the use of demonstratives in three Hong Kong Cantonese corpora of spontaneous speech, for two types of conversations and interpreted speeches.

In Section 2, we review the use of Cantonese demonstratives as described in textbook grammars and observed in real life. Section 3 explains our research questions, and the approach and materials used for their investigation in this study. Section 4 presents the corpus analysis, further discussed in Section 5, followed by a conclusion with future directions in Section 6.

2 Demonstratives in Cantonese

The two demonstratives in Cantonese are the proximal 呢[1] *li1* and the distal 嗰 *go2*, generally corresponding to "this" and "that" in English, and 這 *zhè* and 那 *nà* in Mandarin, respectively.

2.1 Conventional Usages

Normally demonstratives in Cantonese are used for deictic functions, when combined with the appropriate classifiers and nouns, and with numerals as necessary to indicate singular or plural forms (Matthews and Yip, 2011), as in (1).

(1) a. 呢　份　　報告
     li1　fan6　bou3gou3
     this　CL　report
     'this report'

     b. 嗰　兩　　個　　重點
     go2　loeng5　go3　zung6dim2
     that　two　CL　main-point
     'those two main points'

Syntactically, demonstrative phrases may appear as sentential subjects or objects, as prenominal modifiers, in possessive constructions, relative clauses, and appositions, just as diversely as other noun phrases in general. Within a discourse, when

---

[1] The proximal demonstrative is often transcribed as 呢 or 哩, with pronunciation variants including *li1*, *lei1*, *ni1* and *nei1*, and occasionally as 依 *ji1*. In this paper, it is represented uniformly as 呢 *li1*, unless otherwise quoted from specific corpus data with a different transcription.

demonstratives are used, they are expected to carry deictic functions in one of the following types: situational, discourse deictic, anaphoric or recognitional (Himmelmann, 1996).

2.2 Other Observed Usages

Interestingly demonstratives are often optional in Cantonese such that a bare classifier structure could also have demonstrative force, although the definiteness of the noun phrase in the absence of demonstratives is context-dependent. As in (2), the bare classifier noun phrase 塊面 *faai3min6* 'CL face' could refer to the face of anyone or a particular person, depending on the context (e.g. whether the speaker was talking about beauty-loving ladies in general or his wife in particular).

(2) 成日   就   震   吓   塊   面
s*eng4jat6 zau6 zan3 haa5 faai3 min6*
　　always   ADV vibrate ASP CL   face
'apply (the vibrating facial massager) on one's face all the time'

Demonstrative noun phrases are always definite, in Mandarin and Cantonese alike (Sio and Song, 2015). Given that Chinese has no articles (like "a" and "the") to indicate definiteness, demonstratives in Mandarin were found to have developed some functions overlapping with those of definite articles in languages like English (anaphoric in a non-contrastive environment, in situation of shared general knowledge, etc.), although some specific properties of demonstratives still remain (Chen, 2004). However, in contemporary Cantonese, the referent for demonstrative phrases is not always identifiable from the discourse, as shown in (3).

(3) 我   想   講講   呢   就係
ngo5 soeng2 gong2gong2 le1 zau6hai6
　　I   want   talk   PAR   be

　　嗰   個   少數   族裔   嘅
go2 go3 siu2sou3 zuk6jeoi6 ge3
　　that CL   minority   race   GE

　　福利   同埋   權利
fuk1lei6 tung4maai4 kyun4lei6
　　welfare   and   rights
'I'd like to talk about ethnic minority welfare and rights.'

The above example is actually more complicated than showing just the definiteness issue. First, it is structurally ambiguous, where the demonstrative 嗰個 *go2go3* 'that one' may attach to 'ethnic minority' or 'welfare and rights'. The former is not likely as there is no specific ethnic minority group mentioned before, hence a referent is in no way to be found. Even more unlikely with the latter, for not only a similar reason as above, but also the abstract nouns 福利 *fuk1lei6* 'welfare' and 權利 *kyun4lei6* 'rights' do not appropriately go with the classifier 個 *go3*, which is a general classifier often for concrete and countable objects.

More interesting still, with the perfectly grammatical structure, the idiosyncrasy of the utterance may be so subtle that it tends to go unnoticed by most hearers. While it is not possible to identify the referent of the demonstrative, no one will ever question "which ethnic minority" or "which welfare and rights" the speaker is talking about, as if the utterance is most natural. In the example, the use of the demonstrative resembles a definite article, but it seems that none of the situations observed for Mandarin applies here. In other words, the demonstrative in (3) must have other functions, probably non-deictic and more pragmatic ones, like how some other lexical items may have developed (e.g. Feng, 2008; Norrick, 2009; Yaguchi, 2001; Zhang and Gao, 2012).

Changes in the syntactic construction of demonstrative phrases in Cantonese have also been observed. Instead of being used with common nouns, proper nouns are found in demonstrative phrases, as in (4).

(4) 呢   隻   中交建   嚟   講   呢
li1 zek3 zung1gaau1gin3 lei4 gong2 le1
　　this CL   China-Com-Con come talk PAR
'as for China Communications Construction'

The demonstrative in the above example is non-contrastive and non-anaphoric. It only serves to emphasise the proper name that follows, or for more discourse interactional purposes like directing hearers' attention (Kirsner, 1979). In such cases, social attitude is also often reflected from the demonstrative used, where the distal demonstrative may tend to convey disapproval, although the proximal demonstrative is not necessarily positive (Tao, 1999). Example (5)

shows both an emphasis on the (small) amount and a slightly negative attitude associated with the speaker's comment in the rest of the utterance. In this example there may still be a minimal trace of the deictic function, as the speaker is referring to the two pages held and read by another person.

(5)　　嗰　兩　頁　紙　都　睇　咁　耐
　　　go2 loeng5 jip6 zi2 dou1 tai2 gam3 noi6
　　　that two CL paper also look so long
　　　'take so long to read just two pages'

Yet in other cases, the demonstrative actually seems quite unnecessary, if not unjustified. The use of 嗰個 *go2go3* in (6) may resemble the grammaticalization of the Mandarin demonstrative 那個 *nàge* 'that' as a definite determiner in spoken discourse as suggested by Huang (1999). But other than that, the demonstrative phrase is neither in an apposition nor a relative clause, and the demonstrative is more naturally replaced with the possessive marker 嘅 *ge3*.

(6)　　中國移動　　嗰　個　業績
　　　zung1gwok3ji4dung6 go2 go3 jip6zik1
　　　China-Mobile that CL result
　　　'China Mobile's results'

The use of demonstratives in the absence of a referent in the discourse and the unconventional syntactic constructions may not exactly be new usages, but what do they mean to contemporary Hong Kong Cantonese? If the demonstratives are apparently redundant or inappropriate in some cases, and there are alternative ways to express the same message in a much more concise way, why would speakers still produce them? What might be the functional significance of such usages?

3　The Current Study

Based on the above observations, this study aims at attesting the non-deictic usage of demonstratives in contemporary Hong Kong Cantonese with a corpus-based approach. In particular, we analyse and compare the use of demonstratives in three Cantonese corpora, covering different speech types, to illuminate the relationship between the functions of demonstratives and speech contexts.

3.1　Research Questions

The current study thus aims at addressing the following questions.

First, apart from deictic usages, for what other functions are demonstratives used in contemporary Hong Kong Cantonese?

Second, do the usage patterns differ across speech situations? Previous studies, mostly on Mandarin, tend to focus on face-to-face conversations. In this study, we extend the scope and compare face-to-face conversations with other kinds of dialogues and monologues.

Third, if different patterns are found across speech types, how might interlocutory relation bear on the usage of demonstratives, to achieve the necessary discourse and pragmatic purposes?

3.2　Cantonese Corpora

Three corpora were used, covering different speech situations as described in the London-Lund Corpus of Spoken English (Svartvik, 1990)[2]. The first corpus consists of face-to-face conversations, the second corpus consists of public discussions, mainly interviews and interactive programmes on TV and radio, and the third corpus consists of a very specific type of spontaneous speech, namely simultaneously interpreted speeches from the Hong Kong Legislative Council meetings. Dialogues and monologues are thus covered. For dialogues, face-to-face conversations involve closer relations between interlocutors while public discussions, where the conversation is heard by third parties, would mean more distant relations between interlocutors. The monologues, that is the interpreted speeches, are also spontaneous, making them close to conversations in being unplanned, but at the same time incur quite different cognitive demand as the speakers have to be constrained by the speech in the source language while producing the interpreted speech.

Hong Kong Cantonese Corpus (HKCanCor)

According to Luke and Wong (2015), HKCanCor was intended to be an electronic repository of naturally occurring conversations among people in Hong Kong at the turn of the 21st century, with

---

[2] The six major speech situations are: face-to-face conversations, public conversations including debates and interviews, telephone conversations, radio broadcasts, spontaneous speeches and prepared speeches.

most of the recordings made during 1997 and 1998. The audio recordings for the corpus involve some 100 speakers, mostly in their 20s and 30s, balanced for gender. A supplementary set of recordings from radio chat shows was added at a later stage.

The data used in this study are 42 files of face-to-face conversations[3], with about 130K Chinese characters[4], noting that the proximal demonstrative is transcribed as 哩 *li1* in HKCanCor.

Corpus of Verbal Comments (HKVerCom)

According to Kwong (2015), the corpus contains transcribed spoken Cantonese data from television and radio programmes broadcasted in Hong Kong during late 2013 to early 2014. They cover various domains (politics and current affairs, economics and finance, and food and entertainment) presented in various ways (interviews, phone-in programmes, singing contests, and food/film critics).

The current study made use of the corpus data from TV interviews, radio phone-in programmes and financial commentaries by stock analysts, which contain about 249K Chinese characters.

Cantonese Interpreting Corpus (HKLECSIC)

This is a bilingual corpus containing transcribed speeches from the Hong Kong Legislative Council meetings. Although meetings are now mostly conducted in Cantonese, from time to time some members would speak in English, requiring interpreting into Cantonese[5]. The English speech by members during the year 2017-2018 and the corresponding Cantonese interpretation are transcribed and aligned to form the corpus.

The data used in this study contain about 44K Chinese characters, transcribed from the Cantonese interpreted speeches.

3.3 Analysis

Occurrences of the two demonstratives were extracted from the three corpora above. Examples were grouped into: (1) spatial or temporal deictic

---

[3] http://compling.hss.ntu.edu.sg/hkcancor/
[4] Although the HKCanCor data were segmented into words, the corpus size is reported in number of Chinese characters here for easier comparison with the other corpora used in this study. Punctuations and English words were excluded.
[5] Apparently interpreting into L2 has attracted more attention, as Mandarin-English and Cantonese-English interpreting corpora have been reported in previous research (e.g. Hu and Xie, 2010; Pan and Wong, 2018).

usages (including demonstrative with 度 *dou6* / 處 *cyu3* / 邊 *bin1* as spatial markers, and 次 *ci3* / 排 *paai4* / 陣 *zan6* / 時 *si4* / 陣時 *zan6si4* as temporal markers), (2) demonstrative with plural marker 啲 *di1*, with or without the number "one", (3) demonstrative with general classifier 個 *go3*, with or without the number "one", and (4) demonstrative with other numbers and classifiers. With exception of the third type, the others can be relatively safely assumed to retain at least some deictic functions and play a role in identifying the referents. Subsequent analysis thus focused on demonstratives used with the general classifier 個 *go3*. For both demonstratives, a random sample of 100 instances (for 呢個 *li1go3* and 嗰個 *go2go3*) were selected from each corpus. The instances were classified into deictic and non-deictic usages.

3.4 Hypotheses

Two specific null hypotheses were thus tested: One is the true proportion of demonstratives is the same in all speech situations. The other is the true proportion of deictic usages of demonstratives is the same in all speech situations. In the actual analysis we considered the two demonstratives (proximal and distal) together as well as separately.

4 Results

4.1 Frequency of Demonstratives

Table 1 shows the normalised frequency of the two demonstratives in the various corpora.

Corpus \ Demonstrative	呢/哩 *li1*	嗰 *go2*	Total
	(per million chars)		
HKCanCor	4,980	14,161	19,141
HKVerCom	9,128	7,604	16,732
HKLECSIC	15,264	5,201	20,465

Table 1: Normalised Frequency of Demonstratives

Based on the actual frequencies from the samples, a statistically significant difference was found for the proportion of demonstratives among the three corpora (χ^2=47.15, df=2, p<0.05). In particular, it is HKVerCom which shows a significantly lower proportion of demonstratives than the other two corpora, which is somewhat counter-intuitive. Despite the noticeable functional

changes observed with contemporary usage of demonstratives in speech found in the media, it turns out that demonstratives are not more frequently used in public discussions than in daily face-to-face conversations[6].

Even more interesting is the distribution of the two demonstratives in the three corpora, for which a statistically significant difference was found (χ^2=801.65, df=2, p<0.05). In HKCanCor, the distal demonstrative far outnumbers the proximal demonstrative. In HKVerCom, a somewhat opposite trend was observed. However, although there are more proximal demonstratives than distal demonstratives, the difference is not as huge as the even more unbalanced use of demonstratives found in interpreted speeches, where there are three times as many proximal demonstratives than distal demonstratives. Its total usage of demonstratives is also the highest among the three speech situations under investigation. This raises a question as to why, in the lack of a mutually familiar space of reference, unlike what is shared among interlocutors in conversations, there are even more demonstratives used in monologues, or interpreted speeches in this case. We will come back to discuss this point in Section 5.

4.2 Use of Proximal Demonstrative 呢 *li1*

Figure 1 shows the distribution of the various usages of the proximal demonstrative in terms of its collocations, across the three corpora.
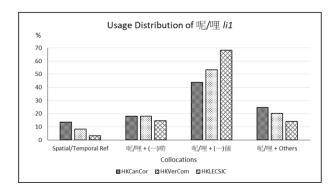

Figure 1: Collocations of Demonstrative *li1*

First, the proximal demonstrative is least used for spatial and temporal reference in all speech types, although it is still more used in this way in face-to-face conversations (13%) than public discussions (8%) and interpreting (3%). Second, considering all usages of 呢 *li1* except 呢(一)個 *li1(jat1)go3*, it means that at least 56%, 46.6%, and 31.8% of its occurrences retain some deictic function in face-to-face conversations, public discussions, and interpreted speeches respectively.

The analysis of 100 random samples of 呢/哩個 *li1go3* shows a distribution of deictic and non-deictic usages as in Figure 2. It turned out that 60% to 70% deictic usages are found in this group, and no significant difference was found across the various corpora (χ^2=2.76, df=2, p>0.05). Combined with other deictic usages mentioned above, the proximal demonstrative tends to retain more deictic usages in face-to-face conversations and public discussions than in interpreted speeches.
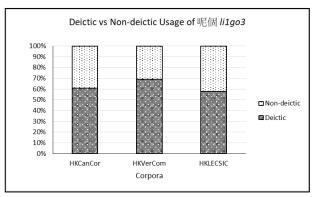

Figure 2: Deictic vs Non-deictic Usage of *li1go3*

Moreover, it should be noted that the proportion of 呢一個 *li1jat1go3* 'this one CL' occupies about 6%, 15%, and 31% of all occurrences of 呢(一)個 *li1(jat1)go3* in HKCanCor, HKVerCom, and HKLECSIC, respectively, and the presence of the number "one" is more often accompanied by other hesitation markers and self-repairs, as in (7).

(7)　呢　一　個　嘅　誒　呢　一　個
　　　li1　jat1　go3　ge3　e6　li1　jat1　go3
　　　this　one　CL　GE　PAR　this　one　CL

　　　免費　　諮詢　　服務
　　　min5fai3　zi1seon1　fuk6mou6
　　　free　consultation　service
　　　'this … uh … this free consultation service'

---

[6] Whether the use of demonstratives in the media has actually increased over these years has yet to be checked by comparing samples of media speech across the years, which will be part of our future work.

4.3   Use of Distal Demonstrative 嗰 *go2*

Figure 3 shows the distribution of the various usages of the distal demonstrative in terms of its collocations, across the three corpora.
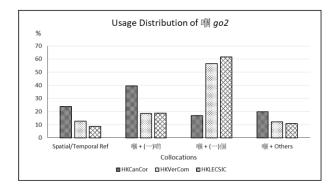

Figure 3: Collocations of Demonstrative *go2*

Compared with the proximal demonstrative, the distal demonstrative 嗰 *go2* is relatively more often used for spatial and temporal reference, especially in face-to-face conversations. Also in this speech situation, the use of 嗰啲 *go2di1* is about twice as much as in the other two corpora, and it is observed that they often appear in relative clauses, as in (8).

(8) 好   細   隻   好   多   毛   嗰啲
　　 hou2 sai3 zek3 hou2 do1 mou4 go2di1
　　 very small CL very much hair those
　　 'those which are very small and very hairy'

The abundance of the above usages in face-to-face conversations thus leaves a very big difference for 嗰個 *go2go3* between face-to-face conversations (17%) and public discussions (57%) as well as interpreted speeches (62%). This, in effect, even counter-balances the differences in the normalised frequencies of 嗰 *go3* across the three corpora. In other words, despite the much lower frequencies of 嗰 *go2* in public discussions and interpreted speeches, they actually use more 嗰個 *go2go3* than face-to-face conversations.

The analysis of 100 random samples of 嗰個 *go2go3* shows a distribution of deictic and non-deictic usages as in Figure 4. It turns out that while 嗰個 *go2go3* retains much deictic function in face-to-face conversations, it has lost most of this function in public discussions and interpreted

speeches.   The difference across the corpora is statistically significant (χ^2=55.41, df=2, *p*<0.05). The demonstrative has been so frequently used that sometimes they even become meaningless, if not misleading. As in (9), the context does not point to a particular election, but makes a general statement about the role of election in democracy.

(9) 嗰   個   選舉   係   一   個   手段
　　 go2 go3 syun2geoi2 hai6 jat1 go3 sau2dyun6
　　 that CL election be one CL means
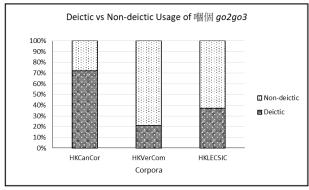　　 'Election is a means (to achieve something).'

Figure 4: Deictic vs Non-deictic Usage of *go2go3*

Another piece of evidence indicating that demonstratives are not actually intended for deictic and referential purposes comes from the mixed use of them in a single utterance, as in (10) and (11).

(10) 呢   個   土地   同埋   房屋
　　 li1 go3 tou2dei6 tung4maai4 fong4uk1
　　 this CL land   and   housing

　　 嗰   個   問題
　　 go2 go3 man6tai4
　　 that CL   problem
　　 'the land and housing issue'

(11) 嗰   個   驗毒   助   康復   呢   個
　　 go2 go3 jim6duk6 zo6 hong1fuk6 li1 go3
　　 this CL test-drug help recover this CL

　　 計劃   其實   我   好   支持   㗎
　　 gai3waak6 kei4sat6 ngo5 hou2 zi1ci4 gaa3
　　 scheme   actually I   very support SFP
　　 'Actually I support the RESCUE Drug Testing Scheme very much.'

The utterances contain both demonstratives "this" and "that", very close to each other, but the speaker is probably only referring to "the" issue and "the" scheme, respectively. The mixed and confusing usages provide strong evidence for other pragmatic functions of demonstratives, which are obviously not deictic.

5 Discussion

As Kirsner (1979) suggested, demonstratives may be more instrumental than simply contrasting spatial distances. They may be used for directing hearers' attention, serving more discourse interactional purposes. As Tao (1999) observed from Mandarin conversational data, concrete spatial uses of the demonstratives are least frequent. The findings in this study tend to echo these previous views, but how do we account for the even more severe deviation from deictic usages in speech situations other than conversations?

While face-to-face conversations, public discussions, and interpreted speeches can all be considered spontaneous, they differ in at least the following respects: interlocutor relationship, expected audience, speaker freedom, and cognitive processes, which may partially account for the observed differences in the use of demonstratives.

Among the three speech situations, interlocutors in face-to-face conversations enjoy the closest relationship. On the one hand, participants in the conversation are both speakers and hearers who are personally involved in the communication. On the other hand, they are often friends discussing mutually familiar topics. The communication barrier is therefore low, and the space of reference is well known by both sides. The exchange can be as spontaneous as possible, and demonstratives can be more naturally used deictically, or otherwise used as an emphasis marker, as in (12), where the demonstrative before the proper noun serves to highlight the proper noun which is more than enough for identifying the unique referent.

(12)　我　　鄉下　　　喺　呢　個　海豐
　　　ngo5 hoeng1haa2 hai2 li1 go3 hoi2fung1
　　　my　hometown　at　this　CL　Haifeng
　　　'My hometown is (at) Haifeng.'

In public discussions, however, one might need to be more mindful of their language and try to sound official or diplomatic as appropriate. In TV interviews, for instance, often the interviewee may represent the government or some organisation. Even when the speaker only speaks for himself, the views they express in response to the interviewer's questions need to be more carefully phrased as the speech may concern the speaker's or an organisation's image, expertise and stance. The speaker would want to be as accurate as possible, and more speech planning is often required. The interviewer, on the other hand, is more indifferent and often inquisitive. The topic under discussion is often more serious. Moreover, the discussion is not just between the interlocutors, but is expected to be heard by others. Thus not only is the relationship between interlocutors less close, the personal involvement and social distance could be quite different from those in face-to-face conversations. Speakers need more time for speech planning, and starting a potential noun phrase with a demonstrative would leave most room for expanding the phrase, possibly at the expense of the naturalness and comprehensibility of the speech, as in (13). In this short utterance there are already three demonstratives. First, none of them is definite, if we look carefully at the speaker's meaning from the context. Second, the first part of the utterance (the subject noun phrase) could well be simplified to 堆填區每日嘅垃圾量, and all the demonstratives serve no syntactic or semantic purpose. They are more like pause fillers for speech planning in the pragmatic sense. But why did the speaker not use non-lexical fillers?

(13)　嗰 個　　堆填區　　嘅 嗰 個
　　　go2 go3 deoi1tin4keoi1 ge3 go2 go3
　　　that CL　　landfill　　GE that CL

　　　每日　　嗰啲　　　垃圾量　　　呢
　　　mui5jat6 go2di1 laap6saap3loeng6 le1
　　　daily　those　garbage-quantity PAR

　　　減　　百分之四十　　　嘅
　　　gaam2 baak3fan6zi1sei4sap6 ge3
　　　minus　40-percent　　　SFP
　　　'The daily quantity of garbage disposed at the landfills will reduce by 40%.'

As for interpreted speeches, it is essentially a monologue produced by the interpreter. Although it is still a type of spontaneous speech, unlike face-

to-face conversations and public discussions where the speakers have absolute freedom in producing the speech, interpreters are somehow constrained by the source speech. Among the three speech situations, interpreting is the most cognitive demanding as the speaker has to go through bilingual processing in a very short time, from comprehending the source speech to organising the message to producing it in the target language. The speakers thus also very much need the extra time for speech planning, and starting with a demonstrative allows easier expansion of a noun phrase. Although it is not uncommon for the utterances to be accompanied by incomplete structures and self-repairs, the speech errors may not be as serious a problem as compared to other speech situations since the hearers in interpreted speeches are non-participants. But again, why not use non-lexical fillers if it is solely for speech planning? Moreover, if interpreting is so cognitively demanding, does it not make more sense to save some time from uttering such unnecessary and meaningless words for paying closer attention to what is to be interpreted?

Alternatively, could the use of demonstratives have been part of the translation prompted by the English source? Looking at the data, however, this is not exactly the case. The original English speech often contains no demonstratives, while they are somehow introduced by the interpreters in the Cantonese speech. As in Example (3) given earlier, the original speech is "I'm going to talk about the ethnic minority rights and welfare" while 嗰個 *go2go3* 'that' is used in the interpreted speech; and as in (14) below, the original is "As a practitioner of the legal profession" while 呢個 *li1go3* 'this' is introduced by the interpreter.

(14) 因為　　我　　係　呢　個
　　　jan1wai6　ngo5　hai6　li1　go3
　　　because　　I　　be　this　CL

　　　法律界　　　　嘅　一　　員
　　　faat3leot6gaai3　ge3　jat1　jyun4
　　　legal-profession　GE　one　member
　　　'because I am a member of the legal
　　　profession'

So if it is not prompted by the source speech, the use of demonstratives in the interpreted speech in those examples is even more unjustified. Without

a shared and mutually familiar space of reference with the hearers, the use of demonstratives by speakers, especially in unconventional syntactic distributions, serves no indispensable purpose. Using demonstratives instead of non-lexical pause fillers for speech planning can only be understood as the speaker's intention to keep the hearer's attention by appearing to be talking continuously, regardless of the actual content.

But this is not advisable. The demonstratives are absolutely unnecessary in those cases. Including them may increase the talking speed, reduce the information content, and introduce distractors. They may lead to more garden path sentences, aggravating the cognitive load in comprehension. Despite a possible hidden agenda to cover up uncertainty while appearing to present a huge amount of information, overusing demonstratives is disadvantageous to both speakers and hearers, and should be consciously avoided to ensure effective communication. One should stay alert, especially in language teaching and interpreter training, so as not to compromise the clarity of speech.

6 Conclusion and Future Directions

In this corpus-based study, we have attested the diverse usage of demonstratives in contemporary Hong Kong Cantonese. Apart from deictic functions, demonstratives were found to be used for other pragmatic functions, including speech planning devices and emphasis markers. It was also observed that these non-deictic functions are more frequent with the distal demonstrative in public discussions and interpreted speeches, where the relationship between interlocutors is often less close, as compared to face-to-face conversations. Being heard by non-participants, speakers may tend to be more mindful of their speech, and starting with a demonstrative allows the most room for expanding a potential noun phrase, thus allowing them to buy time to organise their message. This extra time is also much needed for interpreters to process the source speech, but it is not clear why they would do that at the risk of increasing the cognitive demand on both speakers and hearers. Future work includes more in-depth analysis of corpus examples in relation to the cognitive aspects of language processing in the various speech situations.

Acknowledgments

References

Ping Chen. 2004. Identifiability and Definiteness in Chinese. *Linguistics, 42(6)*:1129-1184.

Guangwu Feng. 2008. Pragmatic markers in Chinese. *Journal of Pragmatics, 40(10)*:1687-1718.

Nikolaus P. Himmelmann. 1996. Demonstratives in Narrative Discourse: A Taxonomy of Universal Uses. In Barbara Fox (ed.), *Studies in Anaphora*, 205–254. Amsterdam: John Benjamins.

Kai-bao Hu and Li-xin Xie. 2010. The Use of English Relative Clauses in Chinese-English Conference Interpreting. *Shandong Foreign Language Teaching Journal, 4*:8-16.

Shuanfan Huang. 1999. The emergence of a grammatical category definite article in spoken Chinese. *Journal of Pragmatics, 31(1)*:77-94.

Robert S. Kirsner. 1979. Deixis in discourse: an exploratory quantitative study of the modern Dutch demonstrative adjectives. In Talmy Givón (ed.), *Syntax and Semantics, Volume 12: Discourse and Syntax*, 355-375. New York: Academic Press.

Oi Yee Kwong. 2015. Toward a Corpus of Cantonese Verbal Comments and their Classification by Multi-dimensional Analysis. In *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation: Posters*, Shanghai, pp.10-18.

Kang Kwong Luke and May L.Y. Wong. 2015. The Hong Kong Cantonese Corpus: Design and Uses. In B.K. Tsou and O.Y. Kwong (eds.), *Linguistic Corpus and Corpus Linguistics in the Chinese Context*. Journal of Chinese Linguistics Monograph Series 25, pp.312-333.

Stephen Matthews and Virginia Yip. 2011. *Cantonese: A Comprehensive Grammar*. Routledge, London.

Neal R. Norrick. 2009. Interjections as pragmatic markers. *Journal of Pragmatics, 41(5)*:866-891.

Jun Pan and Billy Tak Ming Wong. 2018. A Corpus-driven Study of Contrastive Markers in Cantonese-English Political Interpreting. *BRAIN – Broad Research in Artificial Intelligence and Neuroscience, 9(2)*:168-176.

Joanna Ut-Seong Sio and Sanghoun Song. 2015. Divergence in Expressing Definiteness between Mandarin and Cantonese. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar*, Nanyang Technological University (NTU), Singapore (pp. 177–194). Stanford, CA: CSLI Publications.

Jan Svartvik (ed.). 1990. *The London Corpus of Spoken English: Description and Research*. Lund Studies in English 82. Lund, Sweden: Lund University Press.

Hongyin Tao. 1999. The Grammar of Demonstratives in Mandarin Conversation Discourse: A Case Study. *Journal of Chinese Linguistics, 27(1)*:69-103.

Michiko Yaguchi. 2001. The function of the non-deictic *that* in English. *Journal of Pragmatics, 33*:1125-1155.

Wei Zhang and Hua Gao. 2012. The functions of "zikhai (即係)" in Cantonese conversation and its grammaticalization. *Linguistic Sciences, 11(4)*:387-395.