Comparing Classifier Use in 1995 and 2010 Early Child Cantonese to Explore Social Change in Hong Kong

Hui Li

Eileen Chin Mei Wong

The University of Hong Kong

Acknowledgements

This work was partially supported by the GRF Fund (*RGC Ref No. 747109*), HKSAR. We are very grateful to Prof. K.K. Luke and Prof. C.F. Sun for their comments and suggestions.

Abstract

This study compared and analysed the classifier usage in two cohorts: the 492 Cantonese-speaking young children tested in 1995 (Tse & Li, 2011), and the 143 children tested in 2010 (Li &Tse, 2011). The results indicated that: (1) the 1995 and 2010 cohorts shared the same 8 most commonly used classifiers, 個 go3,隻 zek3, 架 gaa3,杯 bui1, 把 baa2, 嚿 gau6, 張 zoeng1,間 gaan1; (2) Two top 10 classifiers were found in the 2010 cohort: 支 zi1, and 部 bou6; (3) both cohorts overused the general classifier 個 *go3*; and (4) both cohorts overused 隻 zek3 and 嚿 gau6 to replace sortal classifiers and even to replace the general classifier 個 *go3*. The pragmatic features and the associated social changes were discussed to understand the minor changes in classifier use between 1995 and 2010 cohorts.

Comparing Classifier Use in 1995 and 2010 Early Child Cantonese to Explore Social Change in Hong Kong

Chinese is a classifier language that needs a morpheme to indicate the classification and quantification of lexicons (Mak, 1991; Tse, Li, & Leung, 2007; Tse & Li, 2011). There are four types of classifier languages such as numeral, concordial, predicate, and intra-locative ones (Allan, 1977), and Chinese (including Mandarin and Cantonese) is a typical numeral classifier language (Szeto, 1998). Appropriate use of Cantonese classifiers requires both semantic and syntactic knowledge to choose semantically correct classifier and to place it syntactically properly (Tse et al., 2007). Therefore, classifier usage in early childhood is a fascinating research area as it illuminates the mental processes of categorization in young learners as well as the social changes in a certain society (Tse & Li, 2011). Currently there is a considerable research into the developmental trends of Cantonese classifier use in young children (Mak, 1991; Tse et al., 2007; Tse & Li, 2011), but there are no studies on the historical changes in classifier use between young children of different cohorts. This study endeavors to address the research gap by comparing two Cantonese corpora: The 1995 Early Child Cantonese Corpus (N= 492) (Tse & Li, 2011) and The 2010 Early Child Cantonese Corpus (N= 143) (Li &Tse, 2011).

Cantonese classifier and its usage

Classifier (*Liangci*, 量詞) refers to the morphemes that denote salient perceived or imputed characteristics of the referents of associated nouns (Allan, 1977). Classifier is widely used in Chinese languages including Mandarin and Cantonese, to fulfill two major pragmatic functions: (1) classification; and (2) measurement. For example, some Chinese classifiers just classify the noun referents in accordance to their intrinsic features (a string, 一條繩), whereas other classifiers simply measure the

noun objects by their quantities (a pair of chopsticks, 一對筷子).

In Cantonese, classifier are obligatory in a noun phrase located after a demonstrative, a number, or some quantifiers (Mathews & Yip,1994), and some action verbs also require classifier to show the number of times an action has occurred (Mak, 1991). Therefore, there are two main types of classifiers in Cantonese, verb classifiers and noun classifiers (Tse, & Li, 2011). This study will primarily focus on the use of noun classifiers, which include mensural classifiers and sortal classifiers. Mensural classifiers are used to denote quantities of noun referents by using containers, such as 杯 *pui1* 'cup' in 一杯水 *jat1 bui1 seoi2* 'one cup of water', standard units of measurement, such as 分 *fan1* 'minute', or by indicating collections of items, such as 班 baan1 'class', and by indicating types, such as 類 *leoi6* 'a kind of'. Sortal classifiers classify nouns according to their intrinsic features, such as shapes or functions (Tse et al., 2007). Normally, shape classifiers categorize nouns according to the number of dimensions (1D, 2D, & 3D), and the flexibility or size of the items. Function classifiers assort with function of target nouns, for example, 把 *baa2* is used for objects with handles (Tse & Li, 2011).

The correspondence between noun and classifiers are neither fixed nor unchangeable. Same nouns could often be classified by more than one classifier (Matthews & Yip, 1994). Erbaugh (2002) found that both Mandarin-speaking and Cantonese-speaking adults could use various classifiers for the same object even though there was no perceivable change for the object. This implies that there may be variations in the use of classifiers within and across individuals. Among the classifiers, however, some are more frequently used than others by young children. In Szeto (1998)'s study, for example, 個 *go3*,隻 *zek3*,度 *dou6*, and 啲 *di1* were found to be the most frequently used ones by young children of 1 to 3 years. Tse, Li and Leung

(2007) analyzed the 1995 cohort of early child Cantonese and found that among the top 10 classifiers used by Cantonese-speaking young children, nine were sortal classifiers, and the only mensural classifier was 杯 *bui1* 'cup', which is itself a name of a container. The top three classifiers were 個 go3 ,隻 zek3, and 架 gaa3; while the other seven top classifiers were 杯 bui1, 把 baa2，條 tiu4， 間 gaan1，嚿 gau6，塊 faai3，and 張 zoeng1. This study will compare the 1995 and 2010 cohort to ascertain whether the top 10 classifiers are shared by the two cohorts.

Overgeneralization of the general and other classifiers

General classifier is the one that could be widely used with any nouns in many numeral classifier languages (Lyons, 1977). In Cantonese (and Mandarin), 個 *go3* (CL) is the most frequently used classifier and can be used to classify numerous nouns, so is widely regarded as the general classifier (Chao, 1968; Li & Thompson, 1981; Erbaugh, 1986; Tse et al., 2007; Tse & Li, 2011). For example, Tse, Li and Leung (2007) analyzed the 1995 cohort of early child Cantonese and found that *go3* was most frequently used to replace most of the subtypes of sortal classifiers such as 隻 *zek3* (animate), 條 *tiu4* (1D flexible), 支 *zi1* (1D rigid), 架 *gaa6* (function), 間 *gaan*1 (function), 張 *zoeng1* (2D shape),塊 *faai3* (2D shape), 棵 *po1*(3D shape), and 把 *baa2* (function). They noted that all of the overgeneralized applications of *go3* occurred in cases whereby a specific sortal classifier was needed to classify the noun. No cases of replacing mensural classifiers with go3 were found in the cohort. It seems that the children understood that the general classifier *go3* could only replace sortal but not mensural classifiers. This could be because the general *go*3 is a sortal rather than a mensural classifier, and the children have acquired the distinction between sortal and mensural classifiers. To verify this speculation, this study will compare two

cohorts in 1995 and 2010.

In addition to *go3*, other classifiers such as 隻 *zek3* (CL), 條 *tiu4* (CL) and 間 *gaan1* (CL), were also overused to replace sortal classifiers in the 1995 cohort (Tse et al., 2007). Since these are not general classifiers, overgeneralisation of their usages tend to be regarded as inappropriate (Tse et al., 2007). For example, 隻 *zek3* does not classify all nouns in Cantonese. It can be applied to animals (such as dogs), round objects (such as eggs), one out of a pair of objects (such as socks), and other various objects (such as ships) (Erbaugh, 2002). Szeto (1998) found that children aged 1 to 3 used 隻 *zek3* inappropriately on some occasions. It is however uncertain whether the use of 隻 *zek3* by young children would differ over time. The present study therefore seeks to fill in this research gap.

The context of this study

The two cohorts were collected in Hong Kong, a former British crown colony that was returned to China in 1997. The change in sovereignty, accompanied by the transition from a colonial to a postcolonial society, has brought up a series of transformations in the sociolinguistic background to the recognition of a local and global space for Hong Kong English (Bolton, 2000). A formal 'biliterate and trilingual' language policy was established in 1997. Since then, young children are supposed to learn English, Cantonese and Putonghua in the early years (Li & Rao, 2000). In light of the language policy, Cantonese has become more dominant in daily life and political speeches, and training courses in Mandarin are made available for Hong Kong civil servants and teachers. Some primary schools and almost all local kindergartens started to teach Mandarin to young children since 1997. With the two financial crises in 1998 and 2008, respectively, Hong Kong has opened the door to Mandarin-speaking tourists from Mainland China. That said, Cantonese, rather than

Putonghua, is still the dominant variety in Hong Kong (Bolton, 2000). Code-mixing between Cantonese and English (Tse & Li, 2011) alongside the frequent use of Mandarin in the city might have impacted on the language use, including the use of Cantonese classifier. Therefore, this study will explore the classifier use in two Cantonese corpora to explore the social changes in Chinese discourse. The 1995 cohort was collected before the return of Hong Kong, whereas the 2010 cohort was collected 13 years after the handover from British to Chinese rule. This cross-cohort comparison of classifier use will help to reveal the social changes and influences on Chinese language discourse. The following questions guided this research:

1. Will the top 10 Cantonese classifiers change over time? Are there any differences in the top 10 classifiers between the 1995 and 2010 cohorts?

2. Will the general classifier *go3* be consistently overused by the two cohorts of young children? Are there any differences?

3. Are there any similar overgeneralizations of the use of other Cantonese classifiers in the two cohorts? If yes, what are they?

Method

The Corpora

Two early child Chinese corpora, namely, the 1995 Early Child Cantonese Corpus (Tse & Li, 2011) and the 2010 Early Child Cantonese Corpus (ECMC) (Li & Tse, 2011), were examined in this investigation.

The 1995 Cohort. This corpus consisted of the utterances produced by 492 Cantonese-speaking preschoolers as they participated in pairs in a 30-minute play task. They were randomly sampled from 18 preschools in Hong Kong. This sample represented children from three age groups (ages 3;0, 4;0 and 5;0), with 30 boys and 30 girls in each age group. All the participants were native speakers of Cantonese, and

their parents and teachers also spoke Cantonese at home and at the preschools, respectively. For more information about this corpus, please refer to Tse and Li (2011).

The 2010 Cohort. This corpus consisted of the utterances produced by 143 Cantonese-speaking preschoolers randomly sampled from 8 preschools located in the four districts of Hong Kong. This sample represented children from three age groups (3;0, 4;0 and 5;0), with 24 boys and 24 girls in each age group (1 uncompleted case). All the participants were native speakers of Cantonese, and their parents and teachers also spoke Cantonese at home and at preschool respectively.

Significant similarities rather than differences were found in the two corpora, implying a satisfactory level of comparability. The two corpora were conducted by the same research team using the same research design, topic, communication task, protocol, and children`s toys. The only difference was that the ECMC had four age groups, with a younger age group (age 2;0) which was not included in the 1995 cohort. Therefore, we excluded 2-year-olds from the comparison and just focused on three age groups: age 3;0, 4;0 and 5;0, for a more valid and feasible comparison between the two corpora.

The communication task. The two corpora shared the same communication task. A toy play area furnished with the same set of toys, including cooking materials, food and fruits, furniture, electrical appliances, hospital materials and vehicles, were set up in the classroom of the participating kindergartens. Only one dyad of same-age participants at a time was allowed to play in the room for 30 minutes, and their conversations were audio-recorded using an unobtrusive recorder. The researchers observed the activities during the free play sessions but did not intervene.

All audio-taped conversations were accessed by two research assistants using

a digital audio player that allowed automatic rewinding for repeated playing at normal speed. The research assistants transcribed the conversations to a level of detail that captured all the words and word fragments (including overlapping speech) audible to the ear, non-lexical fillers (e.g. "*uh*"), as well as other vocalizations (e.g. laughter). Other researchers then independently proofread the transcripts against the taped audio clips to ensure accuracy of transcription.

The coding system. Classifiers (including wrongly-used ones) were identified in the types of noun phrases below in Cantonese (Szeto, 1998; Tse et al., 2007) by two research assistants. The coding results were then scrutinized by one author of this paper and all the authors carefully went through the listed classifiers to achieve a consensus of the Cantonese classifiers.

a. [Dem-Num-CL-(N)] 呢 一 隻 (狗)

nei1 jat1 zek3 (gau2)

this one CL dog (this dog)

b. [Dem-CL-(N)] 呢 隻 (狗)

nei1 zek3 (gau2)

this CL dog (this dog)

c. [Num-CL-(N)] 一 隻 (狗)

jat1 zek3 (gau2)

one CL dog (one dog)

d. [Q-CL-(N)] 每 隻 (狗)

mui5 zek3 (gau2)

each CL dog (each dog)

e. [Wh-CL-(N)] 邊 隻 (狗)

bin1 zek3 (gau2)

which CL dog (which dog)

f. [CL-N] 隻 狗

zek3 gau2

CL dog (the dog)

g. [CL-CL-(N)] 隻 隻 狗

zek3 zek3 (gau2)

CL CL dog (every dog)

h. [N1-CL-N2] 瑪莉 隻 狗

ma5lei6 zek3 gau2

Mary CL dog (Mary's dog)

The Statistical Analyses

First, descriptive analyses were conducted to calculate the total number of classifier types, the frequency, and percentages of every type found in the two corpora. This set of analyses attempted to identify the 10 most commonly used classifiers by the two cohorts, to address the first question of this study. Second, case analyses were conducted to explore the overgeneralization of the general classifier *go3* and other commonly used classifiers. This set of analyses aimed to address the second and third questions. The results are present in the following section.

Results

Cohort differences in early child Cantonese classifiers

We identified all the Cantonese classifiers from the two corpora and analyzed the 10 most commonly used classifiers ("Top-10") from the 1995 and 2010 cohorts. The results are presented in Table 1 and Figure 1. The results indicated that (1) the Top-10 classifiers in the 1995 cohort were: 個 go3,隻 zek3, 架 gaa3, 條 tiu4,杯 bui1, 把 baa2, 嚿 gau6, 張 zoeng1, 塊 faai3, and 間 gaan1; (2) the Top-10 classifiers in the 2010

cohort were: 個 go3,隻 zek3, 間 gaan1,張 zoeng1, 架 gaa3, 杯 bui1, 嚿 gau6, 支 zi1, 把 baa2, and 部 bou6; (3) The two cohorts shared eight of the Top-10 classifiers, 個 go3,隻 zek3, 架 gaa3,杯 bui1, 把 baa2, 嚿 gau6, 張 zoeng1, and 間 gaan1, indicating a high level of stability in classifier usage; (4) only two of the Top-10 classifiers in the 1995 cohort, i.e. 條 tiu4,塊 faai3, were not ranked Top-10 in the 2010 cohort; instead, the other two classifiers, 支 zi1 (15th in 1995) and 部 bou6 (28th in 1995), emerged as Top-10 in the 2010 cohort; (5) as shown in Figure 1, similar ranking and percentages were shared by the two cohorts, and it is very interesting to note that the ranking of 個 go3, 隻 zek3, and 嚿 gau6 remain the same over the 15-year span; (6) as shown in Figure 1, the overall percentage of Top-10 classifiers in 2010 is generally higher than that in 1995 (except for the dramatic decrease in tiu4 條). This indicates that the 2010 cohort of Hong Kong children tended to use them more frequently.

Insert Table 1 and Figure 1 about here

Overgeneralization of general classifier go3

The general classifier 個 *go3* was found the most commonly used one in both cohorts, and the majority of the 1995 (90%) and 2010 (93%) cohorts used *go3* in their peer communication. A minor increase in the frequency of using *go3* was found between the 1995 and 2010 cohorts (see Figure 1). The increase was obvious at least in one of the three age groups, i.e. age 3 (86% vs 90%), age 4 (95% vs 92%) and age 5 (**89% vs 98%**). This tiny increase indicated that relatively more 5-year-olds in the 2010 cohort than their counterparts in the 1995 cohort used 個 *go3* in their speech.

The overextended use of *go3* (CL) occurred most frequently in the two cohorts to

refer to everyday objects which have their own specific classifiers, such as eggs *go3 gai1-daan6* (CL egg; 個雞蛋) and *jat1 go3 gai-daan6* (one (CL) egg; 一個雞蛋), and round objects as *go3 si6-do1-be1-lei2* (one (CL) strawberry; 個士多啤梨) and *loeng5 go3 tai4-zi2* (two (CL) grapes; 兩個提子). As shown in Table 2, the results indicated that: (1) 個 *go3* (CL) was used by the 1995 cohort to replace the subtypes of sortal classifiers such as *隻 zek3* (animate), *條 tiu4* (1D flexible) , 支 *zi1* (1D rigid), 張 *zoeng1* (2D shape), 架 *gaa6* (function), and 間 *gaan1* (function); (2) the 2010 cohort, however, used *go3* to replace 隻 *zek3* (animate), 張 *zoeng1* (2D shape), 塊 *faai3 (*2D shape*)*, 粒 *nap1* (3D shape), 朵 *do2* (function), 架 *gaa6* (function), 把 *baa2* (function) and 包 *baau1* (function); (3) both cohorts used 個 *go3* to replace *隻 zek3* (animate), 張 *zoeng1* (2D shape), and 架 *gaa6* (function); (4) the 2010 cohort tended to overuse *go3* to replace more sortal classifiers such as 塊 *faai3* (2D shape), 粒 *nap1* (3D shape), 朵 *do2* (function), 把 *baa2* (function) and 部 *baau1* (function). It implies that the overgeneralized use of *go3* by Hong Kong young children was extending (in scope) and increasing (in frequency) between 1995 and 2010. It is very important to note that all the overgeneralized use of *go3* only occurred in cases whereby a specific sortal classifier was needed to classify the noun. Tse et al. (2007) found that there were no cases of replacing mensural classifiers with *go3* in the cohort 1995, and they hypothesized that the general classifier *go3* was used to replace sortal classifiers only, but mensural classifier. This study lent support to Tse et al. (2007) as it showed no cases of replacing mensural classifiers with *go3* in the two corpora.

Insert Table 2 about here

Overgeneralization of other classifiers

The general classifier *go3*, however, was not the only overgeneralized classifier used by the two cohorts of children. As shown in Table 3, the comparison results indicated that: (1) the 1995 cohort used 隻 *zek3* (CL), 條 *tiu4*(CL), and 間 *gaan1* (CL) to replace other specific classifiers; (2) the 2010 cohort 'borrowed' 隻 *zek3* (CL), 嚿 *gau6* (CL), and 把 *baa2* (CL) to replace other specific classifiers; (3) it is very interesting to find the overgeneralized use of 隻 *zek3* (CL) in the two cohorts, and it was used to replace 粒 *nap1* (3D shape), 條 *tiu4* (1D flexible) , 間 *gaa6* (function), and 個 *go3* (general CL); (4) another noteworthy finding is that the 2010 cohort tended to overuse 嚿 *gau6 (CL)* to replace 架 *gaa6* (function), 間 *gaan1* (function), 件 *gin6* (function), and 個 *go3* (general CL).

Insert Table 3 about here

Discussion

As the first corpus-based comparison of classifier usage in the early years to date, the research set out to examine a rich set of data in an attempt to ascertain the changes in early child Cantonese pragmatics. The two cohorts of corpora were established in 1995 and 2010 and have provided substantially sized samples and abundant evidence for investigating classifier usage associated with social changes in Hong Kong. This section will discuss all the changes found in this study and explore their implications.

Changes of top 10 Cantonese classifiers

The present study found that eight of the top 10 Cantonese classifiers were

shared by the 1995 and 2010 cohorts: 個 go3,隻 zek3, 架 gaa3,杯 bui1, 把 baa2, 嚿 gau6, 張 zoeng1, and 間 gaan1, indicating a high level of stability in classifier usage. Their ranking in frequency was very similar, especially for 個 go3,隻 zek3, and 嚿 gau6 which ranked the same after 15 years. This finding implies that the classifier usage is very stable among Hong Kong young children over the past decades. Apart from the ranking, generally, the percentage of classifier usage across ages increased (See Figure 1) over time, it indicates that the Chinese (classifier) proficiency has increased among Cantonese-speaking young children between 1995 and 2010. The focus of this study, however, is on the social changes, we will focus on the two new top-10 classifiers emerged in the 2010 cohort, namely, 支 zi1 (15th in 1995) and 部 bou6 (28th in 1995). Below are examples of children's use of the two classifiers.

(1) 支 *zi1*:

畀 支 筆 我 吖

bei3 zi1 bat1 ngo5 aa1
give CL pen me SFP

'Give me a pen'

In this case, the classifier 支 *zi1* was used to classify and quantify the target subject, pen, which is the most familiar object to young children among all the possible referents of *zi1* such as pen, candle, rifle, and song. The emergence of this new top-10 indicated that children in the 2010 cohort talked more about 'pen', implying that writing and drawing are very important learning activities in their early years. This is in line with our previous studies (Li, Rao, & Tse, 2011; Li & Rao, 2000), which found that Hong Kong young children tended to have more copying exercises and writing activities in classroom and at home than before. And 'pen' is the most important learning tool to them.

(2)部 *bou6*:

我 買 到 兩 部 車

ngo5 maai5 dou2 loeng5 bou6 ce1

I buy PRT two CL car

'I have bought two cars'

In this case, the classifier 部 bou6 was used to classify and quantify the target subject, car, which is the most familiar object to young children among all the possible referents of 部 *bou6* such as car, movie, novel, washing machine, and masterpiece. The emergence of this new top-10 indicated that children in the the 2010 cohort talked more about car, which is an important mode of transport to Hong Kong young children. The child in this case might be repeating their parents' words, which implies that car has become a hot topic in their family. In addition, 部 bou6 is frequently used to quantify mobile phone, which has become a household necessity like children's toys. It is thus understandable that 部 bou6 has become a top-10 classifier in the 2010 cohort.

The emergence of 支 zi1 and 部 bou6 in the 2010 cohort demonstrates a noticeable evidence of the association between social and linguistic changes. If some changes are noted in young children’s language, the similar changes might have already occurred in the speech of their older 'models' such as parents and teachers. And this kind of 'model-induced’ changes in classifier use might be caused by the major sociolinguistic change in Hong Kong, that is, the increased use of Putonghua (Mandarin) in school, at home, on the street, and in TV and radio programs. For example, the competing classifier for car in Cantonese is 架 gaa3, which is colloquial Cantonese; whereas 部 bou6 is shared by Putonghua in which 辆 liàng is also

applicable, but not so in colloquial Hong Kong Cantonese. Thus, the following hypothesis has received some empirical support from this finding. Paired with 車 ce1 (car), 部 bou6 is gaining ground in Hong Kong Cantonese under the growing influence of the Putonghua style of speech among the current subjects' older siblings, who have mostly been exposed to Putonghua language classes. More generally, in cases where young Cantonese speakers have the option of two or more classifiers and only one of which is shared by Putonghua, the emerging trend is to favor the shared variant. Meanwhile, we cannot rule out the vast influence of using mobile phones in children's daily life, thus 部 bou6 has become a popular classifier. Nevertheless, more comparisons between the two cohorts in other lexical categories such as nouns and verbs are needed to further verify this hypothesis.

Changes in overgeneralization of go3

The present study found that the two cohorts tended to use *go3* (CL) to replace only the sortal classifiers (not the mensural classifiers) including: (1) animal sortal classifiers; (2) 1D, 2D, and 3D shape sortal classifiers; and (3) function sortal classifiers. These findings are consistent with those of Matthews and Yip (1994) and the findings of Fang (1985), Stokes and So (1997), Szeto (1998) , Wong (2000) and our previous studies (Tse et al., 2007; Tse & Li, 2011). In addition, relatively more children in the 2010 cohort used *go3* in their speech and 'borrowed' it to replace more sortal classifiers such as *faai3 塊 (*2D shape*), nap1* 粒(3D shape), *do2* 朵(function), *baa2* 把 (function) and *baau1* 部 (function). It implies that the overgeneralized use of *go3* in Hong Kong young children has extended (in scope) and increased (in frequency) between 1995 and 2010.

There are cognitive, linguistic and contextual influences that may have shaped the overuse of general classifiers by the 2010 cohort. First, using general classifiers is

a cognitively accessible and economical strategy (Tse, Li, & Leung, 2007), as children can ease the burden on cognitive processing and shorten response time by using the general classifier.

Second, discourse factors are also important. It was noted in our previous study (Tse et al., 2007) that the first mention of an object often featured a specific classifier, but that later mention very often was simplified to the general classifier. Similar finding was also reported in this study. Accordingly, in both cohorts, other classifiers such as 隻 *zek3* (CL), and 嚿 *gau6 (CL)* were also employed by the children to replace specific classifiers. The next subsection will discuss this finding.

Third, the language context of today's young children in Hong Kong might also have some influences on the overuse of general classifiers like *go3*. For example, the overgeneralization of go3 has also been observed in Cantonese-speaking adults' daily communication (Erbaugh, 2002). This overgeneralized use of go3 by parents and teachers, as a language input, will be imitated and acquired by young children gradually and naturally. Accordingly, this kind of 'model-induced' overgeneralization could be also observed in young children's utterance in a play context.

Changes in overgeneralization of other classifiers

The present study found that both 隻 *zek3* (CL) and 嚿 *gau6* (CL) were frequently used to replace other classifiers (and even the general classifier 個 *go3*) in the 1995 and 2010 cohorts. This implies that Cantonese-speaking children in Hong Kong tend to be more tolerant to the overgeneralized use of 隻 *zek3 (CL)* and especially 嚿 *gau6* (CL). Are 隻 *zek3* (CL) and 嚿 *gau6* (CL) another general classifiers? This question is really challenging and requires further studies and theoretical exploration. Conventionally, this overgeneralization should be regarded as

an `inappropriate' usage.

Why did the young children overgeneralize the use of these two high frequency classifiers? First, the overusing of 隻 *zek3* (CL) might be attributed to the fact that this classifier is highly salient in Cantonese speakers (Erbaugh, 2002), as it can be applied to animals (such as dogs), round objects (such as eggs), one out of a pair of objects (such as socks), and other various objects (such as ships and bananas). In this connection, it was the second most frequently used classifier in the present study. Second, the overuse of 嚿 *gau6* (CL) is of special interest to the writers, as there have been no similar reports in other studies.

In Cantonese, 嚿 *gau6* (CL) refers to a whole item without any specific shape and function. Therefore, in their discourse, the children tended to use this classifier to refer to any tangible objects to emphasize its wholeness and substance. Possibly, this case might reflect the fact that the children had not yet reached adult-level proficiency in classifier use. In addition, this overuse might also be associated with the setting of the communication task. In the task, the children had a 30-minute free play at the well-furnished toy area, which was equipped with many Lego or plastic toy items to simulate their daily life. Those toys were only miniatures of real objects, thus the children tended to use gau6 to describe them (with an emphasis on their wholeness and substance). For example, in the phrase "賣嚿汽車 (maai6 gau6 hei3-ce1, selling a car)", the child just used gau6 to replace "架"(gaa3) (See Table 2). This overuse of 嚿 gau6 made their peer conversation more vivid and interesting. But it might also be due to the fact that the tiny toy car in the role play was made of a lump of plastic.

Literally, 架 *gaa*3 means 'framework' , the chassis and body for cars and airplanes, but not for ships, boats, yachts, etc. However, when the expected

framework is not ‘salient’ (in evidence), the ‘lump perspective’ seems to determine the classifier chosen by some of these children. This kind of ‘perspectival’ or ‘perceptual shift’ is also attested among mature speakers, as clearly noted by Matthews & Yip (1994:105-107). Nevertheless, this finding shows that the young children tended to use 嚿 *gau6* to refer to lumpy, round-shaped, 3-dimensional, and very tiny objects. It might be possible that when using real objects in the communication task, the young children would try to use correct classifiers rather than 嚿 gau6. This hypothesis, however, needs further studies.

The present study has limitations. First, the two cohorts had different sample size, N_{1995}=492, N_{2010}=143. But this does not necessarily mean that they are not comparable, as we are comparing the most commonly used classifiers in two cohorts that were stratified randomly sampled in Hong Kong. There is no doubt about the representativeness of the data. Second, the time-gap between 1995 and 2010 might be too short (*T* gap = 15 years) to reveal any statistically significant linguistic change in the groups studied. But the results indicated that eight of top 10 classifiers were shared by the two cohorts whereas two ‘new’ classifiers emerged in the 2010 cohort. This reflects substantial language stability and a tiny but noticeable language change. Nevertheless, following a corpus-based study with a longer time gap could be conducted to track the changes in classifier use.

As it stands, the study offers a descriptive but accurate account of the changes in Cantonese classifier usage. The discovery that young Cantonese-speaking children use the same eight Top-10 classifiers at least since the 1990s, and that they tend to overgeneralize the use of several commonly used classifiers should be noted by early childhood educators. The two new Top-10 classifiers emerged in 2010, 支 *zi1* and 部 *bou6*, reflecting the social changes in young children's school and home life. Such

changes should also be recognized and children should be provided with structured learning to utilize and extend their linguistic repertoire. The researchers were struck by the considerable variability and social changes in children's language and believe that there is clearly scope for further study of Cantonese classifiers.

References

Allan, K. (1977). Classifiers. *Language, 53*, 285–311.

Bolton, K. (2000). The sociolinguistics of Hong Kong and the space for Hong Kong English. *World Englishes, 19*, 265—85.

Carpenter, K. L. (1991). Later than sooner: Extralinguistic categories in the acquisition of Thai classifiers. *Journal of Child Language, 18*, 99–113.

Chao, Y. R. (1968). A grammar of spoken Chinese. Berkeley : University of California Press.

Erbaugh, M. S. (1986). Taking stock : The development of Chinese noun classifiers : historically and in young children. In C. G. Craig (ed.), *Typological studies in language : Noun classes and categorization, Vol. 7*, 399–436. Amsterdam/Philadelphia : John Benjamins.

Erbaugh, M. S. (2002). Classifiers are for specification : Complementary functions for sortal and general classifiers in Cantonese and Mandarin. *Cahiers de Linguistique – Asie Orientale, 31* (1), 33–69.

Fang, F. X. (1985). An experiment on the use of classifiers by 4 to 6 year olds. *Acta Psychological Sinica, 17*, 384–92.

Gandour, J., Petty, S. H., Dardarananda, R., Dechongkit, S. & Mukngoen, S. (1984). The Acquisition of numeral classifiers in Thai. *Linguistics, 22*, 455–79.

Li, C. N. & Thompson, S. A. (1981). A reference grammar of Mandarin Chinese. Berkeley: University of California.

Li, H., & Rao, N. (2000). Parental influences on Chinese literacy development: A comparison of preschoolers in Beijing, Hong Kong and Singapore. *International Journal of Behavioural Development, 24*(1), 82-90.

Li, H., Rao, N., & Tse, S.K. (2011). Bridging the Gap: A Longitudinal Study of the

Relationship between Pedagogical Continuity and Early Chinese Literacy Acquisition. *Early Years, 31*(1).

Li, H., & Tse, S.K. (2011). *The Early Child Mandarin (Cantonese) Corpus*. Hong Kong: The University of Hong Kong.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: a Functional Reference Grammar.* The University of California Press.

Loke, K. K. (1991). A semantic analysis of young children's use of Mandarin shape classifiers. In A. Kwan-Terry (ed.), *Child language development in Singapore and Malaysia*,98–116. Singapore: Singapore University Press.

Lyons, J. (1977). *Semantics.* Cambridge: Cambridge University Press.

Mak, D. L. W. (1991). *The acquisition of classifiers in Cantonese.* Unpublished doctoral dissertation, University of Reading.

Matthews, S. & Yip, V. (1994). *Cantonese : A comprehensive grammar.* London: Routledge.

Stokes, S. & So, L. K. H. (1997). Classifier use of language disordered and age-matched Cantonese speaking children. *Asia Pacific Journal of Speech, Language and Hearing, 2,* 83–101.

Szeto, K. K. S. (1998). *The acquisition of Cantonese classifiers.* Unpublished PhD dissertation, University of Hong Kong.

Tse, S. K., & Li, H. (Eds.). (2011). *Early Child Cantonese: Facts and Implications.* Berlin: De Gruyter Mouton.

Tse, S.K., Li, H., & Leung, S.O. (2007). The Acquisition of Cantonese Classifiers by Preschool Children in Hong Kong. *Journal of Child Language, 34*(2), 495-517.

Wong, C. S. P. (2000). How Cantonese-speaking two-year-olds fend for themselves through the ticket of classifier. In E. V. Clark (ed.), *Proceedings of the thirtieth*

annual child language research forum, 149–58. Stanford: Centre for the Study of Language and Information, Stanford University.

Table 1
Comparison of Top 10 Classifiers identified in 1995 and 2010 Early Childhood Cantonese Corpus

Total (Age 3-5) (N=492)	*Age 5 (N=164)*	*Age 4 (N=164)*	*Age 3 (N=164)*	1995 Ranking	*1995 Classifier*	*2010 Classifier*	*2010 Ranking*	Ranking in 1993	*Ag (N=*
442	146	155	141	1	個 go3	個 go3	1	1	4
148	56	52	40	2	隻 zek3	隻 zek3	2	2	1
67	20	13	34	3	架 gaa3	間 gaan1	3	10	9
57	28	17	12	4	條 tiu4	張 zoeng1	4	8	1
55	22	17	16	5	杯 bui1	架 gaa3	5	3	5
39	8	15	16	6	把 baa2	杯 bui1	6	5	5
38	16	13	9	7	嚿 gau6	嚿 gau6	7	7	5
37	13	16	8	8	張 zoeng1	支 zi1	8	15	6
34	15	10	9	9	塊 faai3	把 baa2	9	6	6
33	12	10	11	10	間 gaan1	部 bou6	10	28	2
21	11	5	5	14	樣 joeng6	樣 joeng6	11	14	3
19	8	9	2	15	支 zi1	條 tiu4	12	4	1
15	5	5	5	18	次 ci3	塊 faai3	13	9	1
6	2	2	2	28	部 bou6	次 ci3	-	18	6

Table 2

*Summary of the Overgeneralizations of **go3** in 1995 Cohort (N=492) and 2010 Cohort (N=143)*

Type of Overgeneralization	**Cases in 1995 Cohort (N=492)**	
Replace ‘zek3 隻’ (Sortal, Animate)	攞個杯 *Lo2 **go3**bui1* *take CL cup* *"Take a cup"*	從前我有一個青蛙 *Cung4-cin4 wo3 jau5 jat1 **go3** cing1-waa1* *in-past-days I have one CL frog* *"Once upon a time, there is a frog"*
	一個雞蛋 *jat1 **go3** gaai1-dan2* *one CL egg* *"An egg"*	有個熊人見到好多蟻 *Jau5 **go3** hung4-jan2 gin3 dou2 ho2 do1 ngai5* *have CL bear see PRT very many ant* *"There was a (toy)bear which saw many ants"*
Replace ‘*tiu4* 條’ (Sortal, Shape:1D flexible)	呢個魚嚟？ *Li1 **go3** jyu2 lei1* *this CL fish SFP* *"Is this a fish?"*	整個虫虫，整個鼻仔呀 *Zing2 **go3** cong4-cong2, zing2 **go3** bei6-zai2 a3* *make CL caterpillar, make CL nose small SFP* *"Make a CL caterpillar, make a CL small nose"* 呢個橋嚟□！ *Li1 **go3** kiu4 lai4 gaa3* *this CL bridge SFP SFP* *"This is a bridge!"*
Replace ‘*zi1* 枝’ (Sortal, Shape: 1D rigid)	做個火箭 *Zou6 **go3** fo2 zin3* *make CL rocket* *"Make a rocket"*	畀個益力多我吖 *Bei2 **go3** jik1lik6dou1 ngo3 a3* *give CL Yakult (a drink) me SFP* *"Give the Yakult to me"*
Replace "張 zoeng1" (Sortal, Shape: 2D)		砌個 BB 床 *Cai3 **go3** bi4 bi1 cong4* *build CL baby bed* *"Build a baby bed"*
Replace ‘*gaa6* 架’ (Sortal, Function)	整爛個火車 *Zing2 lan6 **go3** fo2-ce1* *make break CL train* *"Broke the CL train"*	嘩！你有坐呢個纜車 *Wai3! Lei5 jau6 cuo3 li1 **go3** laam6 ce1* *wow you have sit this CL cable car* *"Wow! you again take this CL cable car"*
	嘩！呢個巴士 *Wai3! Lai4 **go3** baa-si2* *wow this CL bus* *"Wow! give me the bus"*	依度有個影相機呀 *Ji1-dou6 jau5 **go3** jing2 seong2 gei1 aa3* *here have CL camera SFP* *"Here is a camera"*

Replace '*gaan1* 間' (Sortal, Function)	有一個郵局喺度 *Jau5 jat1* ***go3*** *jau4-gok2 hai2 dou6* *have one CL post office at here* *"There is a post office"*	有一個店係買船既店 *Jau5 jat1* ***go3*** *dim3 hai3 maai3 syun4 ge3 dim3* *have one CL shop be sell ship LP shop* *"There is a shop for buying ship"*
	Animate (horse) "匹 pei1"	我有呢個馬，呢度喎 *Ngo3 jau5 ne1* ***go3*** *ma3, nei11 dou6 wo3* *I have this CL horse here SFP* *"I have this horse, here"*
	Shape: 2D "塊 faai3"	兩塊膠布 *Leong3* ***go3*** *gau1-bou3* *two CL plaster* *"Two plasters"*
	Shape: 3D "棵 po1"	砌曬成個聖誕樹 *Cai3 sai3 sing4* ***go3*** *sing3-daan3-syu6* *build ASP whole CL Christmas tree* *"Build the whole Christmas tree"*
	Function "把 ba2"	聽到一個聲音 *Tang1 dou2 jat1* ***go3*** *sing1-jum1* *hear PRT one CL voice* *"I heard a voice"*

Type of Overgeneralization	**Cases in 2010 Cohort (N=143)**	
Replace 'zek3 隻 (Sortal, Animate)	呢度仲有個匙羹 *nei1-dou6 zung6jau5* ***go3*** *ci4gang1Here still-have CL spoon* *"One more spoon is here"*	一個雞蛋飛在這裏 *jat1* ***go3*** *gai1daan6 fei1 zoi6 ze5-leoi5* *One CL egg fly at here* *"One egg flew here"*
Replace 'zoeng1 張" (Sortal, Shape:2D)	有兩個梳化呀 *jau5 loeng5* ***go3*** *so1-faa3 aa1* *have two CL sofa SFP* *"There are two sofas"*	
Replace 'faai3 塊' (Sortal, Shape:2D)	有個鏡 *jau5* ***go3*** *geng3* *have CL mirror* *"one mirror"*	哪個床？ *naa5* ***go3*** *cong4* *which CL bed* *"Which bed?"*
Replace 'nap1 粒" (Sortal, Shape:3D)	拎埋個士多啤梨 *ling1-maai4* ***go3*** *si6-do1-be1-lei2* *Take-PRT CL strawberry* *"Took the strawberry"*	仲有一個積木 *zung6-jau5 jat1* ***go3*** *zik1-muk6* *still-have one CL brick* *"still have one brick"*
Replace other Sortal Classifiers	Function, 'do2 朵"	擺個花落去煮先 *baai2* ***go3*** *faa1 lok6-heoi3 yu2 sin1put* *CL flower down-go cook first* *"Put the flower and cook"*
	Function, 'gaa3 架'	買個大車 *maai5* ***go3*** *daai6 ce1* *buy CL big car*

	"buy a big car"
Function, 'baa2 把'	呢個較剪攞住啦
	nei1 ***go3*** *gaau3-zin2 lo2-zyu6 laa1*
	This CL scissors take-ASP SFP
	"Take this scissors"
Function, ' baau1 包'	唔記得煮埋個薯條呀
	m4 gei3-dak1 zyu2-maai4 ***go3*** *syu4-tiu4 aa1*
	not remember cook PRT CL French fries SFP
	"I have forgotten to cook the French fries"

Note: LP-Linking Particle; ASP-Aspect Marker; PRT-Particle; SFP-Sentence Final Particle.

Table 3
Summary of the Overgeneralizations of Other Classifiers in 1995 Cohort (N=492) and 2010 Cohort (N=143)

Types		*Cases*
Overgeneralization of *zek3 隻*	*Replace "nap1 粒"*	四隻花生 *Sei3 **zek3** faa1 saang1* *four CL peanut* *"Four peanuts"*
	Replace "tiu4 條"	我見到隻蛇喺度行緊 *Ngo3 gin3 dou2 **zek3** se4 hai2 dou6 hang4gan2* *I see PRT CL snake at here walk ASP* *"I saw the snake walking"*
	Replace "gaa3 架"	點知突然有隻機械人 *Dim2 zi1 dak6 jin4 jau5 **zek3** gei1-haai6- jan4* *how know sudden have CL robot* *"What'd you know, a robot suddenly popped up"*
Overgeneralization of *tiu4 條*	*Replace "zi1 枝"*	呢條至係針呀，你自己先至蠢 *Li1 **tiu4** zi3 hai6 zam1 a3, nei3 zi6-gei2 sin1 zi3 ceon2* *this CL just is needle SFP you self also is stupid* *"This is a needle, you are the one who is stupid"*
	Replace "gin6 件"	呢條乜嘢衫？ *Li1 **tiu4** mat1-je3 saam1* *this CL what garment* *"What is this (garment)?"*
Overgeneralization of *gaan1 間*	*Replace "go3 個"*	我媽咪喺間花園到有味道有顏色 *Ngo3 ma1mi4 hai2 **gaan1** fa- jyun2 dou6 jau5 mei6-dou6 jau5 ngan4-sik1* *my mum at CL garden until have smell have colour* *"My mum at the CL garden have smell and colour"*
Types		**Cases in 2010 Cohort (N=143)**
Overgeneralization of *zek3* 隻	*Replace* 'go3 個'	煮咗一隻人 *zyu2 zo2 jat1 **zek3** jan4* *cook ASP one CL person* *"Cooked a person"*
		呢隻係聽筒喎 *nei1 **zek3** hai6 teng1-tung2 wo3* *this CL be stethoscope SPF* *"This is a stethoscope"*

	Replace 'tiu4 條'	落隻魚！炒吓先 *lok6* ***zek3*** *jyu4 caau2-haa5 sin1* *down CL fish fry-ASP first (go3)* *"Put the fish! fry awhile"*
Overgeneralization of *gau6* 嚿	*Replace* 'gaa3 架'	賣嚿汽車畀你！ *maai6* ***gau6*** *hei3-ce1 bei2 nei5* *sell CL car to you* *"Sell you a car!"*
	Replace'gaan1 間'	我想整嚿屋企呀 *ngo5 soeng2 zing2* ***gau2*** *uk1-kei2 aa1* *I want make CL home SFP* *"I wanna make a home"*
	Replace'"gin6 件 /faai3 塊'	撞倒嚿魚柳啦！ *zong6dou3* ***gau6*** *jyu4lau5 laa1* *knock-PRT CL fish-filet SFP* *"Knocked down the fish filet"*
	Replace'個 go3 '	我帶埋嚿西瓜畀你喇 *ngo5 daai3-maai4* ***gau6*** *sai1-gwaa1 bei3 nei5 laa3* *I bring-PRTCL watermelon to you SFP* *"I brought you a watermelon"*
		有嚿收音機啊 *jau5* ***gau6*** *sau1-jam1-gei1 aa3* *have CL radio SFP* *"There is a radio"*
Overgeneralization of *baa2* 把	*Replace 'bou6* 部'	喂!加把電視 *wai3 gaa1* ***baa2*** *din6-si6* *hey add CL TV* *"Hey! Add one TV set"*

Note: LP-Linking Particle; ASP-Aspect Marker; PRT-Particle; SFP-Sentence Final Particle.

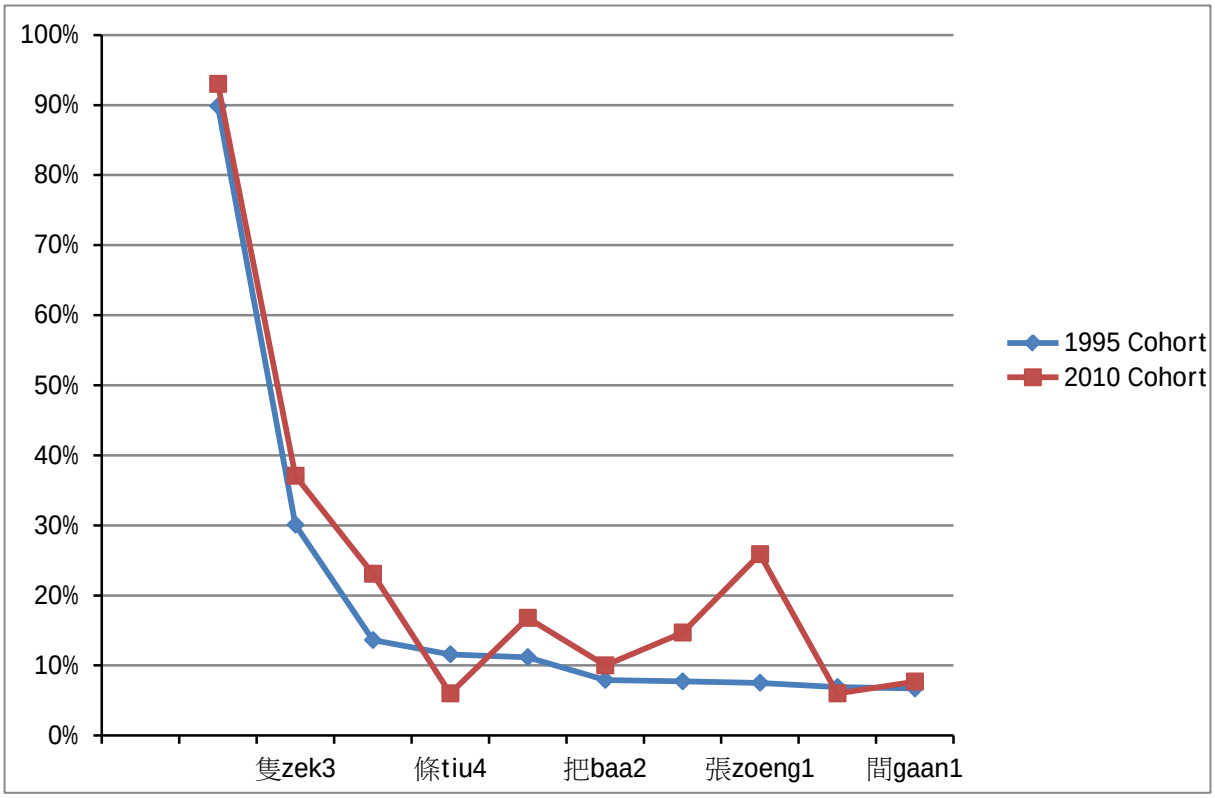

Figure 1 Comparison of Top 10 Classifier Usages in 1995 and 2010 Corpus (%)