

# Multiple Correspondence and Typological Convergence in Contact-Induced Grammaticalization

*Evidence from Cantonese-English Bilingual Development*

*Pui Yiu Szeto*

University of Hong Kong  
*szetopy@hku.hk*

*Stephen Matthews*

University of Hong Kong  
*matthews@hkucc.hku.hk*

*Virginia Yip*

The Chinese University of Hong Kong  
*vcymatthews@cuhk.edu.hk*

## Abstract

This paper examines the emergence of perfective aspect in Cantonese-English bilingual children from the perspective of contact-induced grammaticalization, focusing on the novel use of *already*. Although the adverbial *already* seems to serve a function similar to that of the Cantonese perfective marker *zo2* in the bilingual children, other model constructions suggest that the function of *already* may combine those of several Cantonese particles such as the sentence-final particle *laa3*. The results suggest that in contact-induced grammaticalization, it is possible to develop a new category in the replica language based on multiple different but related categories in the model language. Adopting an evolutionary approach to language transmission (Mufwene, 2001),

---

\* This work was supported by funding from the Chinese University of Hong Kong (CUHK) for the University of Cambridge-CUHK Joint Laboratory for Bilingualism, CUHK-Peking University-University System of Taiwan Joint Research Centre for Language and Human Complexity; General Research Fund from the Hong Kong Research Grants Council [Project no. 14413514 and 14632016] and funding for the Bilingualism and Language Disorders Laboratory at Shenzhen Research Institute of the Chinese University of Hong Kong.

we discuss why grammaticalization in the Cantonese-English bilingual children does not seem to involve coevolution of form and meaning, why the grammaticalization phenomena in the bilingual children are only transient, and how the study of bilingual acquisition can contribute to contact linguistics.

### Keywords

language contact – grammaticalization – bilingual acquisition – perfective aspect – Cantonese – English

## 1 Introduction

Since Weinreich's (1953: 1) statement that the language-using bilingual individuals are the locus of language contact, parallels between bilingual development and language contact have attracted considerable scholarly attention. Transfer in second language acquisition is widely acknowledged to play a role in contact-induced language change (Lefebvre, 1998; Siegel, 2008; Clements, 2009). Some authors have proposed a similar role for transfer in bilingual first language acquisition (Thomason, 2001; Satterfield, 2005; Yip and Matthews, 2007). Adopting Mufwene's (2001, 2008) theoretical framework, a bilingual child's simultaneous exposure to two languages allows the linguistic features of the two languages to compete in the 'feature pool' in the child's mind, which may lead to contact phenomena like those observed in contact languages such as creoles. With an increasing amount of work addressing the connection between bilingual development and language contact in recent years (e.g. Müller, 2006; Soriente, 2007; Yip and Matthews, 2007; Ringblom, 2012; Silva-Corvalán, 2014), we now have a better understanding of the link between the two fields. However, given that contact-induced grammaticalization itself is a relatively new field of study, further research work on the grammaticalization phenomena observed in bilingual development can help illustrate how manifestations of grammaticalization may be affected by various typological and ecological factors.

The emergence of perfective aspect in Cantonese-English bilingual children, a phenomenon revealing the intriguing parallels between bilingual development and language contact, constitutes the focus of this study. The English utterances of Cantonese-English bilingual children show a number of parallels with Singapore Colloquial English (Singlish) (Yip and Matthews, 2007), one of which is the use of bare verb forms together with *already* when inflectional morphemes such as the *-ed* and *-en* verbal suffixes are required in Standard English (1–4).

Cantonese-English bilingual children's English (Yip and Matthews, 2007)

- (1) You wipe your mouth already? (Kasen 3;00)
- (2) You swallow the short teeth already. (Timmy 3;01)

Singlish (Bao, 2005)

- (3) I wash my hand already.  
'I have washed/washed my hand.' (*ibid.*, pp: 239)
- (4) I see the movie already.  
'I have seen/saw the movie.' (*ibid.*, pp: 238)

In their work on contact-induced grammaticalization in Cantonese-English bilingual children, Matthews and Yip (2009) argue that the emergence of *already* as a perfective aspect marker presents a case of ordinary contact-induced grammaticalization, and make an important observation that both the [V *zoz*...*laa3*] and [V *saai3*...*laa3*] constructions may serve as the models for interlingual identification, suggesting that *already* may not correspond to the Cantonese perfective aspect marker *zoz* only. We will take this observation as the point of departure and pursue this line of investigation further. We aim to address the following questions in particular:

- i. Do Cantonese-English bilingual children use *already* as a perfective aspect marker? If so, what constructions in Cantonese may have led to the development of such a new category in their English?
- ii. Is there any sign of typological convergence in grammaticalization tendencies in Cantonese-English bilingual development?
- iii. How can the grammaticalization phenomena observed in bilingual acquisition shed light on contact linguistics?

To address these questions, we first review and discuss issues concerning grammaticalization (especially those occurring in contact situations) in Section 2, thereby outlining a framework for analysis. We then analyze the congruence between *already*, *zoz*, and *laa3* in Section 3, which helps to explain why the bilingual children use *already* as a perfective aspect marker. Section 4 presents the background information of the children in this study and the research methodology involved. Section 5 reports findings on the use of *already* in the bilingual children in terms of function, placement, and frequency. The implications of such findings on the theory of contact-induced grammaticalization and other contact phenomena are then discussed in Section 6. Some conclusions are drawn in Section 7.

## 2 Grammaticalization

### 2.1 *A Global or Type-Specific Phenomenon?*

Grammaticalization theory is concerned with the emergence and development of grammatical forms and constructions. According to Hopper and Traugott (2003: 18), the phenomenon of grammaticalization refers to ‘the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions’. Such a distinction between the two stages of grammaticalization from a semantic perspective may be problematic because it is hard to define what ‘more or less grammatical’ means without additional criteria from other domains of grammar (Bisang, 2015). Thus, typological studies on grammaticalization based on large numbers of languages (e.g. Bybee, 1985; Bybee *et al.*, 1994; Lehmann, 1995; Heine and Kuteva, 2002) generally combine the semantic side of the linguistic sign with its form side. This combination reveals the interesting fact that similar grammaticalization ‘clines’ or ‘pathways’ are found in a wide range of unrelated languages, suggesting that such a phenomenon may be shaped by some universal processes of grammatical change.

Studies like these generally take it for granted that there is a certain degree of interdependence between the meaning-side and the form-side of grammaticalization. However, manifestations of grammaticalization observed in East and Mainland Southeast Asian (EMSEA) languages (Bisang, 2004, 2011, 2015) cast doubt on such assumptions. Languages in this area (including Cantonese) generally belong to the isolating or analytic type, and are characterized by their tone systems and discreteness of syllable boundaries, which contribute to the relative morphophonological stability of grammaticalized items in these languages (Bisang, 2011). For example, no significant phonetic reduction is observed in the grammaticalized form of the Cantonese *gwo3* ‘to pass/cross’ (Ansaldo and Lim, 2004):<sup>1,2</sup>

- 
- 1 Cantonese examples are given in the *JyutPing* Romanization system, developed by the Linguistic Society of Hong Kong (Tang *et al.*, 2002) to meet both linguistic criteria and the constraints imposed by computer applications. IPA correspondences are given in Matthews and Yip (2011: 461). The numbers at the end of each syllable represent the tones: 1 (high level), 2 (high rising), 3 (mid level), 4 (low falling), 5 (high rising) and 6 (low level).
  - 2 The abbreviations used in this paper are as follows: 1SG = 1<sup>st</sup> person singular; 2SG = 2<sup>nd</sup> person singular; 3SG = 3<sup>rd</sup> person singular; CLF = classifier; NEG = negation; PFV = perfective; PROG = progressive; SFP = sentence-final particle; SUR = surpass.

- (5a) *Ng05 gw03 ma5lou6*  
 1SG cross road  
 'I cross the road.'
- (5b) *Ng05 daai6 gw03 nei5*  
 1SG big SUR you  
 'I'm older than you.'

Further, in his rhythm-based typology, Schiering (2006) shows that grammaticalization is accompanied by erosion (phonetic reduction) in languages of stress-based rhythm only. Such findings suggest that typology plays a significant role in determining whether a language undergoes coevolution of form and meaning over time, potentially challenging the assumption that grammaticalization works the same way across language types and areas. One of the major research goals of this study is to find out whether the grammatical development of the Cantonese-English bilingual children's English may be affected by the grammaticalization tendencies of Cantonese.

## 2.2 *Contact-Induced Grammaticalization: An Overview*

Until recently, grammaticalization has typically been viewed as a 'language-internal' process as there used to be a widespread assumption among linguists that grammatical structure, or syntax, cannot be 'borrowed' (Heine and Kuteva, 2010). However, new findings from a number of fields, particularly areal typology (Ansaldo, 1999; Dahl, 2001; Enfield, 2003; Heine and Kuteva, 2003, 2005) and creole studies (Arends *et al.*, 1994), have provided abundant evidence to demonstrate that grammaticalization could result from 'external' factors such as geographical clustering and substrate influence. In their seminal work on grammaticalization, Hopper and Traugott (2003: 230) also acknowledge the important link between language contact and grammaticalization, concluding that 'Contact has been an important factor for most languages, and a strictly monogenetic view of grammaticalization is ultimately inappropriate'.

Analyzing data from a wide range of languages, an important observation highlighted by Heine and Kuteva (2005) is that the principles of grammaticalization are the same regardless of whether or not language contact is involved. Similarly, Matthews and Yip (2009) see contact as a catalyst driving change along pathways of grammaticalization, and suggest that 'the general principles of grammaticalization are applicable, not only to the substrate language(s) in which grammaticalization originally took place, but also to the contact language affected by it' (Matthews and Yip, 2009: 373).

For example, while the use of equivalents of ‘already’ as a perfective aspect marker (the focus of this study) is observed in contact varieties such as Singlish (Bao, 2005), Chinese-Cuban Pidgin Spanish (Lipski, 1999)/Chinese Coolie Spanish (Clements, 2009), Chinese Immigrant Spanish (Clements, 2003, 2009), Zamboanga Chabacano (realized as *ya-*) (Steinkrüger, 2013) (6), and Papiá Kristang (realized as *ja*) (Baxter, 2013) (7), it is noteworthy that such a grammaticalization process has also developed ‘internally’ in languages like Inuit and Buli (Bybee *et al.*, 1994).

(6) *Éle ya-matá pwérko gat Alyá gránde* (Steinkrüger, 2013)  
 3SG PFV-kill pig really there big  
 ‘S/he killed a really big pig there.’

(7) *Eli ja bai mar onti anoti* (Baxter, 2013)  
 3SG PFV go sea yesterday night  
 ‘S/he went fishing last night.’

Such an observation is consistent with Mufwene’s (2001, 2008) argument against the distinction between ‘internally motivated change’ and ‘externally motivated change’. Contact plays a crucial role in Mufwene’s approach to language evolution as he maintains that all kinds of linguistic change are made possible by the interaction between different idiolects, regardless of whether they are artificially labelled as ‘internally motivated’ or ‘externally motivated’. In short, although grammaticalization may be triggered in contact situations, its underlying principles remain unaffected.

### 2.3 Ordinary Contact-Induced Grammaticalization

Heine and Kuteva (2005: 80) distinguish ‘two main types of contact-induced grammaticalization depending on whether or not there exists already a model source-to-target grammaticalization process to be replicated’. If no such model exists, the process is referred to as ordinary contact-induced grammaticalization (whereas the other type is known as replica grammaticalization). The mechanisms involved in the transfer of grammatical concept or structure from the model language (M) to the replica language (R) in ordinary contact-induced grammaticalization are as follows:

- (8) Ordinary contact-induced grammaticalization (Heine and Kuteva, 2005: 81)
- a. Speakers notice that in language M there is a grammatical category  $M_x$ .
  - b. They develop an equivalent category  $R_x$  in language R on the basis of the use patterns available in R.

- c. To this end, they draw on universal strategies of grammaticalization, using construction Ry in order to develop Rx.
- d. They grammaticalize category Ry to Rx.

To put it another way, speakers of language M would expect the same grammatical distinctions in language R when they acquire the latter as a second L1 or an L2. Consequently they would search for equivalents in language R to categories in language M with which they are already familiar. This process, according to Hickey (2010: 155), is ‘an unconscious one and persists even with speakers who have considerable target language proficiency’.

The use of *already* as a perfective aspect marker in Singlish is an example of ordinary contact-induced grammaticalization, in which a new grammatical category was developed in the replica language (Singlish, a contact variety of English) due to transfer from the model languages (Southern Sinitic and Malay varieties). Compared to Standard English, Singlish seldom employs inflectional morphology. Its aspectual system, as Bao (2005: 237) puts it, is ‘essentially the Chinese system filtered through the morphosyntax of English’. For example, *already* occurs after uninflected verbs to mark perfective aspect without involving any inflectional morpheme. This is consistent with the use of the perfective aspect marker *liau* in Hokkien, one of the most important substrate languages in the formation of Singlish (Ansaldo, 2009):

- (9a) Singlish  
I eat already.  
‘I have eaten.’
- (9b) Hokkien  
*Gua tsiah png-liau*  
1SG eat rice-PFV  
‘I have eaten.’
- (10a) Singlish  
He go to Taiwan already.  
‘He has gone to Taiwan.’
- (10b) Hokkien  
*I khi taiuan-liau*  
3SG go Taiwan-PFV  
‘S/he has gone to Taiwan.’

Lack of a clean typological fit between languages in a contact situation can lead to the re-analysis of certain particles (Clements, 2009). In this case, the Singlish

speakers grammaticalize a lexical item to an aspect marker instead of using inflectional morphology to mark perfective aspect, following the aspect-marking strategy of the substrate languages. It would be interesting to investigate why *already* is re-analyzed as a perfective aspect marker. This study will discuss this question in detail when we talk about the use of *already* in Cantonese-English bilingual children, which is arguably also a case of ordinary contact-induced grammaticalization (Matthews and Yip, 2009).

Although the model proposed by Heine and Kuteva (2005) provides a point of departure, as Matthews and Yip (2009) point out, there are three main problems associated with the model. First, as language contact necessarily involves bilinguals (Mufwene, 2001, 2008), the label ‘Speakers of language R’ seems to show a monolingual bias. Second, it does not sound plausible to assume that people (especially children) have the required metalinguistic awareness to ‘notice that in language M there is a grammatical category Mx’. The third problem, ‘budding historical linguist’ paradox, can be viewed as an extension of the metalinguistic awareness problem – in the model of replica grammaticalization, the notion that people can ‘replicate a grammaticalization process they assume to have taken place in language M’ (Heine and Kuteva 2005: 92) is not quite realistic as the study of grammaticalization process is the work of trained historical linguists; it is next to impossible for the vast majority of the language users to determine the grammaticalization process which has taken place in a language. Nevertheless, as long as we bear in mind that contact-induced grammaticalization is a natural and unconscious phenomenon which occurs in a bilingual speaker’s mind, the problematic wording used in Heine and Kuteva’s model need not affect our understanding and analysis of contact-induced grammaticalization phenomena.

### 3 Congruence between *Already*, *z02*, and *laa3*

To see how contact-induced grammaticalization applies to the case of *already*, we need to consider what the relevant counterparts might be in Cantonese.

#### 3.1 *Analysis*

The development of ‘already’ into a marker of perfective aspect is attested in a number of languages, as seen in (1–4) (also see Section 2.2). Yet it remains unclear why such a development is cross-linguistically common. By analyzing the congruence between *already*, *z02*, and *laa3* in semantic and pragmatic aspects, we argue that *already* is very likely identified with the Cantonese [*z02...laa3*] construction in the bilingual children.



The English adverb *already* is often associated with completed events, which makes it a natural perfective aspect marker. For example, in (11), the event of starting is completed when the event of arriving occurs.

- (11) The performance had already started when we arrived.

However, as Traugott and Waterhouse (1969) and Soh (2009) point out, in addition to a ‘change of state’ interpretation, an important feature associated with the use of *already* is that the assertion made by the sentence is contrary to what one may expect or assume (referred to as the ‘contrary to expectation’ interpretation by Soh, 2009). For example, the sentence in (11) implies that the performance began earlier than expected.

By contrast, neither the ‘change of state’ nor ‘contrary to expectation’ reading is obligatorily associated with the Cantonese perfective aspect marker *z02*. As *z02* is used to encode an event, seen as a whole or as completed (Matthews and Yip, 2011), it may seem unlikely that *z02* alone is involved in the interlingual identification of *already*. However, *z02* commonly co-occurs with the sentence-final particle *laa3*, which functions to express current relevance (Matthews and Yip, 2011) or change of state (Cheung, 2007). For instance, *laa3* has to be used when the event described has current relevance, thus entailing a change of state, as in (12), where there must be a prior state in which the speaker has not returned home.

- (12) *Ngo5 faan1-z02 uk1 kei2 laa3*  
 1SG return-PFV home SFP  
 ‘I have returned home.’

In fact, as Soh (2009) demonstrates, the semantics of *already* is similar to the Mandarin sentence-final particle *le*, which corresponds closely to Cantonese *laa3* (Matthews and Yip, 2011).

To refer to a completed event with no current relevance, *z02* is used alone, as in (13), where the fact that ‘it rained yesterday’ has nothing to do with the current situation.

- (13) *Kam4jat6 lok6-z02 coeng4 jyu5*  
 yesterday drop-PFV CLF rain  
 ‘It rained yesterday.’

On the other hand, to emphasize a change of state, *laa3* must be used while *z02* is optional. The context in (14) and (15) is that there had not been any rainfall

for a long time, and it finally rained yesterday. Similarly, *laa3* is also obligatory in (16), where it functions to denote an inceptive reading. From the above examples, we can conclude that *laa3* is closely associated with a ‘change of state’ reading.

(14) *Kam4jat6 zung1jyu1 lok6-zo2 coeng4 jyu5 laa3*  
 yesterday finally drop-PFV CLF rain SFP  
 ‘It finally rained yesterday.’

(15) *Kam4jat6 zung1jyu1 lok6-(zo2) jyu5 laa3*  
 yesterday finally drop-PFV rain SFP  
 ‘It finally rained yesterday.’

(16) *Lok6 jyu5 laa3*  
 drop rain SFP  
 ‘It’s started to rain.’

The ‘contrary to expectation’ reading may still be present in sentences where *zo2* and *laa3* co-occur. But in such cases, the sentence is also associated with a ‘completion’ and ‘change of state’ reading. For example, the utterance of B in (17) is associated with a sense of ‘contrary to expectation’ because the fact that ‘B has eaten’ is in contrast with A’s inviting B to join him/her for a meal. At the same time, B’s utterance also conveys the meaning that the eating event is completed and s/he has changed from a state of ‘not having eaten’ to ‘having eaten’. (18) is the English equivalent of (17), in which *already* carries the ‘change of state’ and ‘contrary to expectation’ readings.

(17a) *Heoi3 m4 heoi3 sik6 faan6 aa3?*  
 go NEG go eat rice SFP  
 ‘Will you go and eat the meal?’

(17b) *Ngo5 sik6-zo2 je5 laa3*  
 1SG eat-PFV thing SFP  
 ‘I’ve already eaten.’

(18) A: Let’s go eat, shall we?  
 B: No, I’ve already eaten.

Although the [*zo2...laa3*] construction shares some properties with *already* in the above example, the differences in acceptability between (19) and (20)

suggest that they differ in the obligatoriness of the ‘contrary to expectation’ reading – when responding to a neutral yes-no question, the use of [z02...laa3] construction is perfectly acceptable, but the use of *already* in (20) sounds unnatural to native speakers. This is because the ‘contrary to expectation’ reading inherently associated with *already* is incompatible with a neutral context; meanwhile, the use of [z02...laa3] construction does not pose a problem because it is not obligatorily associated with a ‘contrary to expectation’ reading. The congruence between *already*, *z02*, and *laa3* is summarized in Table 1.

- (19a) *Maai5-z02 bun2 syu1 mei6 aa3?*  
 buy-PFV CLF book NEG SFP  
 ‘Have you bought the book?’
- (19b) *Maai5-z02 laa3*  
 buy-PFV SFP  
 ‘Yes, I have.’
- (20) A: Have you bought the book?  
 B: ?Yes, I have already bought it.

TABLE 1 *Congruence between already, z02, and laa3.*

	Completion	Change of state <sup>3</sup>	Contrary to expectation
<i>z02</i>	+	-	-
<i>laa3</i>	-	+	-
<i>z02...laa3</i>	+	+	+/-
<i>already</i>	+	+	+

3 Cantonese *laa3/z02...laa3* does not always have the same ‘change of state’ meaning as that associated with the English *already*. Both can convey the meaning that a proposition is true at the time of speaking and there was a time when it wasn’t. For example, by saying ‘Tom has already arrived’ or its Cantonese equivalent ‘aa3 Tom dou3-z02 laa3’, Tom has arrived at the time of speaking and there must be a previous state in which Tom hasn’t arrived yet.

Meanwhile, only Cantonese *laa3/z02...laa3* conveys an inceptive or inchoative meaning which denotes the beginning of an action or state, as shown in (16) and (30b–32b). See Soh (2009) for a detailed comparison between English *already* and Mandarin sentence-final *le* (which corresponds closely to the Cantonese *laa3*).

The use of *laa3* and *already* without any sense of completion or change of state is not considered as they are irrelevant to perfective aspect marking. When used in isolation, *zo2* and *laa3* convey a sense of ‘completion’ and ‘change of state’ respectively. Neither of them shows a high degree of congruence with *already* (as in ‘I’ve already eaten’). However, the co-occurrence of *zo2* and *laa3* results in a high degree of congruence with *already* – conveying both the sense of ‘completion’ and ‘change of state’ – and the only difference lies in the obligatoriness of the ‘contrary to expectation’ reading. Given the high frequency of [*zo2...laa3*] constructions in Cantonese, it is possible for the bilingual children (especially the Cantonese-dominant ones) to identify *already* with [*zo2...laa3*].

### 3.2 Predictions

Based on the above analysis, we hypothesize that the Cantonese [*zo2...laa3*] construction serves as the model for interlingual identification, leading to the development of *already* into a perfective aspect marker in the bilingual children. If this is the case, two predictions follow. First, *already* will undergo semantic bleaching and no longer be obligatorily associated with a ‘contrary to expectation’ interpretation. In their acquisition of English, the bilingual children may map some features of [*zo2...laa3*] onto *already*. Because of this mapping, *already* in their speech will have different features than it has in monolingual English. Therefore, we predict that the bilingual children will use *already* in neutral contexts where no sense of ‘contrary to expectation’ is involved, such as the one described in (21), which is considered unnatural by native speakers.

- (21a) Have you watched the movie?  
 (21b) ?Yes, I have already watched it.

Second, given that Cantonese [*zo2...laa3*] always occurs post-verbally, we predict that *already* will have a strong tendency to occur post-verbally in the bilingual children’s utterances, unlike the monolingual use of *already*, which can also occur pre-verbally. It is also noteworthy that *zo2* always occurs immediately post-verbally in a verb phrase, while *laa3* always occurs in a sentence-final position, as shown in (22–23).

- (22) *Ng05 sik6-zo2 maan5faan6 laa3*  
 1SG eat-PFV dinner SFP  
 ‘I’ve had dinner (already).’

- (23) *Ng05 dou3-z02 hok6haau6 laa3*  
 1SG arrive-PFV school SFP  
 'I've arrived at school (already).'

If *already* is identified with [z02...laa3], it should occur either in immediately post-verbal or in sentence-final position.

4 Data for this Study

The data for this study comes from the Hong Kong Bilingual Child Language Corpus in the Child Language Data Exchange System (CHILDES) (MacWhinney, 2000). The corpus documents the longitudinal development of nine simultaneous bilingual children who were exposed to Cantonese and English from birth. The data was collected when the children were between ages 1;03 and 4;06. The background information of the nine bilingual children involved in this study is shown in Table 2. Further details of the children and data collection are given in Yip and Matthews (2007). Diary data of Timmy (1;03–6;00), Sophie (1;06–5;06), and Alicia (1;00–5;04) will also be used to complement the corpus data when necessary, but they will not be used in the quantitative analyses because of the inevitable selection bias of the diary data (see Yip and Matthews, 2007: 72 for relevant discussion).

TABLE 2 Background information of the Cantonese-English bilingual children.

Name	Age	No. of English files	No. of utterances in English files
Timmy	2;01.22–3;06.25	38	6,241
Sophie	1;06.00–3;00.09	40	6,717
Alicia	1;03.10–3;00.24	40	5,109
Llywelyn	2;00.12–3;04.17	17	4,121
Kathryn	3;01.05–4;06.07	17	4,202
Charlotte	1;08.28–3;00.03	19	4,621
Janet	2;10.16–3;11.11	22	5,043
Kasen	2;04.07–4;00.09	20	5,318
Darren	1;07.23–3;11.24	28	5,304
Total		241	46,676

TABLE 3 *Background information of the English monolingual children.*

Name	Corpus	Age	No. of English files	No. of utterances in English files
Adam	Brown (1973)	2;03.04–5;02.12	55	45,520
Eve	Brown (1973)	1;06–2;03	20	10,856
Sarah	Brown (1973)	2;03.05–5;01.06	139	30,945
Lara	Rowland and Fletcher (2006)	1;09.13–3;03.25	118	47,687
Thomas	Lieven <i>et al.</i> (2009)	2;00.12–4;11.20	379	201,628
Total			711	336,636

We also make use of corpus data for five English monolingual children taken from three corpora in the CHILDES database (Brown, 1973; Rowland and Fletcher, 2006; Lieven *et al.*, 2009) to conduct systematic comparison between the longitudinal development of the two groups of children. The background information of the five English monolingual children in this study is shown in Table 3.

To study the use of *already* in the children, the command 'kwal' in the CLAN program is used to search for utterances containing the word *already* in both the monolingual and bilingual data. Instances of imitation, repetition, and unclear utterances in the search results were discarded.

## 5 The Use of *Already* in the Bilingual Children

### 5.1 *The Function of Already*

If Cantonese-English bilingual children use *already* as a perfective aspect marker, we would expect to find examples in which *already* occurs alongside a verb to mark completed actions, functioning like the Cantonese perfective aspect marker *zo2*. A number of such examples are found in the corpus data (24a–26a), for which the Cantonese counterparts are given in (24b–26b).

- (24a) She wake already. (Sophie 2;06)
- (24b) *Keoi5 seng2-zo2 (laa3)*  
 3SG wake-PFV SFP  
 'She has woken up (already).'

(25a) I say already the neck. (Kasen 3;05)

(25b) *Ngo5 waa6-zo2 hai6 tiu4 geng2 (laa3)*  
 1SG say-PFV be CLF neck SFP  
 ‘I’ve (already) said it’s the neck.’

(26a) I turn on already. (Darren 2;10)

(26b) *Ngo5 hoi1-zoek6-zo2 (laa3)*  
 1SG open-on-PFV SFP  
 ‘I’ve turned it on (already).’

Whether the above utterances are associated with a sense of ‘change of state’ or ‘contrary to expectation’ remains uncertain, as it would be perfectly acceptable to add *laa3* to the sentence-final position of the Cantonese equivalents of the utterances. But it is certain that the use of *already* in the above utterances convey a sense of ‘completion’. In addition, examples of the combination ‘already now’ are found in the Cantonese-English bilingual children’s speech data but not in the monolingual data. As discussed in Matthews and Yip (2009), a possible model for such a combination is the [V *zo2...laa3*] construction. In these examples, *already* functions to indicate that the events concerned are completed, and *now* functions to mark current relevance.

(27a) The police car crash, broken already now. (Timmy 2;10)

(27b) *Gaa3 ging2ce1 zong6-laan6-zo2 laa3* (Cantonese counterpart)  
 CLF police-car crash-broken-PFV SFP  
 ‘The police car has crashed (already).’

(28a) Gaagaa eat the congee already now. (Charlotte 2;04)

(28b) *Gaagaa sik6-zo2 di1 zuk1 laa3* (Cantonese counterpart)  
 Gaagaa eat-PFV CLF congee SFP  
 ‘Gaagaa has eaten the congee (already).’

There is also evidence suggesting that *already* functionally corresponds to *laa3* instead of *zo2*. First, as discussed in Matthews and Yip (2009), the ‘all already’ combination found in some bilingual children’s speech data as a result of interlingual identification with the Cantonese [V *sai3...laa3*] construction, where *all* corresponds to the quantifying particle *sai3* and *already* corresponds to the sentence-final particle *laa3*, denoting a sense of ‘change of state’. In the following example, the child uses such a combination to indicate that all the food has been eaten up, implying a change of state where the food concerned no longer exists.

- (29a) He eat all already. (Sophie 2;09)  
 (29b) *Keoi5 sik6 saai3 laa3* (Cantonese counterpart)  
 3SG eat all SFP  
 'He has eaten up all the food.'

Furthermore, the use of *already* in progressive sentences suggests that it sometimes corresponds to *laa3* to convey an inchoative reading (30a–32a), where the events concerned began before the time of utterance and are ongoing at the time of utterance. *Already* does not mark perfective aspect in these examples, otherwise it would be incompatible with the *-ing* progressive aspect marker. By the same token, the co-occurrence of the Cantonese perfective aspect marker *zo2* with the progressive marker *gan2* would result in an ill-formed sentence (33).

- (30a) Sleeping already. (Sophie 2;05)  
 (30b) *Fan3-gan2 laa3* (Cantonese counterpart)  
 sleep-PROG SFP  
 '(Someone) is sleeping already.'

- (31a) She going already. (Charlotte 2;09)  
 (31b) *Heoi3-gan2 laa3* (Cantonese counterpart)  
 go-PROG SFP  
 'She is going already.'

- (32a) xxx dancing already. (Janet 3;04)  
 (32b) xxx *tiu3-gan2 mou5 laa3* (Cantonese counterpart)  
 xxx jump-PROG dance SFP  
 'xxx is dancing already.'

- (33) \**Ngo5 sik6-zo2-gan2 faan6*  
 1SG eat-PFV-PROG rice  
 Intended meaning: 'I'm eating already.'

If *already* is identified with [*zo2...laa3*] by the bilingual children as expected, it is not surprising that both *zo2* and *laa3* contribute to the meaning of *already*, leading to its polyfunctionality. An important prediction made in Section 3.2 is that *already* will undergo semantic bleaching and be no longer obligatorily associated with a 'contrary to expectation' reading. Consistent with our prediction, there are cases in which the bilingual children's use of *already* is unlikely to be associated with a 'contrary to expectation' interpretation. In (34a), the child appears to be replying to a neutral question by indicating that



the sending action is completed. There appears to be no specific presupposition suggesting a ‘contrary to expectation’ interpretation.

- (34a) Investigator: Did you send him to the hospital?  
 Child: I send already. (Timmy 3;01)
- (34b) A: *Sung3-zo2 keoi5 jap6 ji1jyun2 mei6?* (Cantonese counterpart)  
 send-PFV 3SG enter hospital NEG  
 ‘Have you sent him to the hospital?’  
 B: *Sung3-zo2 laa3*  
 send-PFV SFP  
 ‘Yes, I have.’

Moreover, one child is observed to use *already* when asking yes/no questions with no presupposition about whether the enquired actions should have been done, as illustrated in (35a). In its Cantonese counterpart (35b), the perfective aspect marker *zo2* is used to indicate that the questions are about completed actions. As such usage is not found in the monolingual data, the results support our hypothesis that *already* is identified with the [*zo2...laa3*] construction and undergo semantic bleaching in the bilingual children, losing its obligatory association with the ‘contrary to expectation’ reading.

- (35a) Child: You already eat?  
 Investigator: Mm?  
 Child: Okay, you wipe your mouth already? (Kasen 3;00)
- (35b) *Nei5 sik6-zo2 je5 mei6? Nei5 maat3-zo2 zeoi2 mei6?*  
 2SG eat-PFV thing NEG 2SG wipe-PFV mouth NEG  
 ‘Have you eaten? Have you wiped your mouth?’

Although it is difficult to determine whether the monolingual children have the pragmatic capacity to use *already* to convey a counter-expectation sense at such an early stage of development, the lack of data clearly showing that the monolingual children use *already* to mark perfectivity in neutral contexts strongly suggests a fundamental difference between the bilingual and monolingual children in the acquisition of *already*.

### 5.2 The Placement of *Already*

*Already* is expected to have a strong tendency to occur post-verbally in the bilingual children due to influence from the Cantonese postverbal [*zo2...laa3*] construction. The corpus data confirm this prediction. As shown in Table 4, *already* predominantly occurs post-verbally (90%) in the bilingual children. Five out of nine children produced *already* only in postverbal position.

TABLE 4 *The placement of already in the Cantonese-English bilingual children.*

Name	Pre-verbal <i>already</i>	Post-verbal <i>already</i>
Timmy	1	12
Sophie	0	20
Alicia	0	4
Llywelyn	0	3
Kathryn <sup>4</sup>	4	3
Charlotte	1	11
Janet	0	24
Kasen	5	18
Darren	0	4
Total	11 (10%)	99 (90%)

TABLE 5 *The placement of already in the English monolingual children.*

Name	Pre-verbal <i>already</i>	Post-verbal <i>already</i>
Adam	14	4
Eve	2	2
Sarah	14	3
Lara	5	5
Thomas	16	8
Total	51 (69.9%)	22 (30.1%)

Bilingual children's strong tendency to place *already* in postverbal position is contrasted with monolingual children's clear preference for preverbal placement of *already*. As shown in Table 5, pre-verbal *already* is used more often than post-verbal *already* (69.9% vs. 30.1%) in the monolingual children. Fisher's

4 Kathryn appears to be an outlier in this regard – unlike all the other bilingual children, she has slightly more preverbal use, like the monolinguals do. We suspect such a difference may be due to the fact that Kathryn's recording started at a much later stage than the other children (see Table 2). At such an age period, the target-like tense/aspect forms may gradually emerge in the bilingual child at the expense of the non-target perfective *already* (see Section 6.2).

TABLE 6 *The use of immediately post-verbal and sentence-final already in the Cantonese-English bilingual children..*

Name	Immediately post-verbal	Sentence-final
Timmy	0	2
Sophie	1	4
Alicia	1	0
Llywelyn	0	2
Kathryn	0	2
Charlotte	0	6
Janet	2	7
Kasen	1	7
Darren	0	1
Total	5 (13.9%)	31 (86.1%)

(1954) exact test shows that the differences between the two groups are highly significant ( $p < .0001$ ). The results support our hypothesis that the bilingual children have identified *already* with the Cantonese [zɔ2...laa3] construction.

Further, as discussed in Section 3.2, we predict the bilingual children to use *already* either immediately post-verbally (like zɔ2) (36) or in a sentence-final position (like laa3) (37). Examples like (28) are considered ambiguous in this respect and not included in the comparison.

(36) I make already this. (Janet 3;10)

(37) I put her clothes already. (Charlotte 2;10)

(38) No, I sleep already. (Llywelyn 2;08)

To investigate this issue, the post-verbal *already* in the bilingual children is further subdivided into two categories, namely ‘immediately post-verbal’ and ‘sentence-final’. Table 6 reveals that immediately post-verbal *already* is quite rare in the corpus data, suggesting that the Cantonese sentence-final particle *laa3* may have played a significant role in the interlingual identification of *already*.

The five examples of immediately post-verbal *already* found in the corpus data are shown in (39a–43a), and their Cantonese equivalents are given in (39b–43b).

- (39a) I say already the neck. (Kasen 3;05)
- (39b) *Ng05 waa6-z02 hai6 tiu4 geng2 laa1*  
 1SG say-PFV be CLF neck SFP  
 'I've (already) said it's the neck.'
- (40a) I make already this. (Janet 3;10)
- (40b) *Ng05 zing2-z02 ni1-g03 laa3*  
 1SG make-PFV this-CLF SFP  
 'I've (already) made it.'
- (41a) Father: Who comes on Saturday usually?  
 Child: Miss Michelle, come already Miss Michelle. (Janet 3;08)
- (41b) Miss Michelle *lei4-z02 laa3* Miss Michelle  
 Miss Michelle come-PFV SFP Miss Michelle  
 'Miss Michelle has (already) come.'
- (42a) [answering the question 'Where is the monster?']  
 He go already the monster. (Sophie 2;10)
- (42b) *Keoi5 zau2-z02 laa3 zek3 gwaai3sau3*  
 3SG go-PFV SFP CLF monster  
 'The monster has (already) left.'
- (43a) Have already here. (Alicia 2;09)
- (43b) *Jau5 laa3 ni1dou6*  
 have SFP here  
 'It's already there.'

It is noteworthy that the immediately post-verbal *already* corresponds to *z02* only in two of the above examples (39a and 40a). The other examples (41a-43a) involve the dislocation focus construction (also known as 'right-dislocation'), in which the focus phrase ( $\beta$ ) is placed at the beginning of the sentence, followed by a sentence particle (SP) and then the remnant ( $\alpha$ ) (Cheung, 2009). For example, (46a) is a Cantonese sentence following the canonical word order; (46b) is its counterpart with dislocated focus construction, where the focus phrase *zau2-z02* is dislocated to the front, followed by the sentence particle *laa3* and finally the remnant *di1 hok6sang1* 'the students'.

- (44a) *Di1 hok6sang1 zau2-z02 laa3*  
 CL student leave-PFV SFP  
 'The students have left.'

- (44b) *Zau2-zo2 laa3 di1 hok6sang1*  
 leave-PFV SFP CL student  
 ‘The students have left.’

Thus, the utterances in (41a–43a) can be analyzed as follows (41’–43’). (41’) is a special case of dislocation focus construction known as discourse copying construction, where  $\alpha$  is a ‘copy’ of an element in  $\beta$  (Cheung, 2009). Such a use of *already* is not attested in monolingual usage but is reminiscent of the use of the Cantonese construction in (41b–43b) where a sentence particle is required between the  $\beta$  and  $\alpha$  strings.

- (41’) Miss Michelle come already Miss Michelle. (Janet 3;08)  
 $\beta$  SP  $\alpha$
- (42’) He go already the monster. (Sophie 2;10)  
 $\beta$  SP  $\alpha$
- (43’) Have already here. (Alicia 2;09)  
 $\beta$  SP  $\alpha$

5.3 *The Frequency of Already*

Another interesting finding of this study is the notable difference in the usage frequency of *already* between the bilingual and monolingual children. Comparing Table 7 with Table 8, all the bilingual children use *already* more frequently than the monolingual children (0.73–4.76 vs. 0.12–0.55 tokens per 1000 utterances); and overall, *already* is used more than ten times as often in the bilingual children as in their monolingual counterparts (2.36 vs. 0.22 tokens per 1000 utterances). An unpaired *t* test shows that the differences between the two groups are statistically significant ( $p < .05$ ).

Such a remarkable difference in frequency is arguably a result of the development of *already* from minor to major use pattern in the bilingual children’s English, a phenomenon commonly observed in situations of language contact (Heine and Kuteva, 2005: 44–62), where an existing use pattern in the replica language is used more frequently, extended to new contexts, and may become associated with a new grammatical function. In monolingual usage, the use of *already* can be regarded as a minor use pattern restricted to specific contexts and associated with a sense of ‘contrary to expectation’. However, as discussed in Section 3.2, because of the considerable overlap between the functions of *already* and the Cantonese [zo2...laa3] construction, the former is likely identified with the latter in the Cantonese-English bilingual children. Given that the

TABLE 7 *The frequency of already in the Cantonese-English bilingual children.*

Name	No. of <i>already</i>	No. of utterances	Frequency (per 1000 utterances)
Timmy	13	6,241	2.08
Sophie	20	6,717	2.98
Alicia	4	5,109	0.78
Llywelyn	3	4,121	0.73
Kathryn	7	4,202	1.67
Charlotte	12	4,621	2.60
Janet	24	5,043	4.76
Kasen	23	5,318	4.32
Darren	4	5,304	0.75
Total	110	46,676	2.36

TABLE 8 *The frequency of already in the English monolingual children.*

Name	No. of <i>already</i>	No. of utterances	Frequency (per 1000 utterances)
Adam	18	45,520	0.40
Eve	4	10,856	0.37
Sarah	17	30,945	0.55
Lara	10	47,687	0.21
Thomas	24	201,628	0.12
Total	73	336,636	0.22

Cantonese [zɔ2...laa3] construction is not necessarily associated with a sense of 'contrary to expectation', *already* undergoes semantic bleaching and is no longer obligatorily associated with such a pragmatic function, eventually turning into a more widely used major use pattern for marking perfective aspect.

The results suggest a close connection between semantic bleaching and frequency increase in contact-induced grammaticalization. The semantic or

pragmatic meaning obligatorily associated with a lexical item in the replica language (English) limits its use in particular contexts. Once the lexical item is identified with a grammar word in the model language (Cantonese), it will gradually lose its original meaning and can be used in new contexts, resulting in an increase in usage frequency. A similar trend is observed in Colloquial Singapore English, where the usage frequency of *already* is remarkably higher than that in its acrolectal counterpart (Bao and Hong, 2006). The prevalence of the use of equivalents of 'already' as a perfective aspect marker in contact languages is likely a result of the association of 'already' with completed events (see example 11 in Section 3.1). Our results suggest that the grammaticalization of 'already' in contact languages may need to incorporate the loss of the 'contrary to expectation' feature.

## 6 Discussion

### 6.1 *Implications for Contact-Induced Grammaticalization*

Earlier studies often assume that contact-induced grammaticalization is triggered by a single source, e.g. for morphemes in Singlish (Bao, 2005). Through investigating the congruence of *already*, *zo2*, and *laa3*, and analyzing the use of *already* in the bilingual children, an important finding of this study is that the properties of *already* in the bilingual children's English combine those of several Cantonese particles, suggesting that one-to-one correspondence may not be the norm in ordinary contact-induced grammaticalization. As reviewed in Section 2.3, there is no model source-to-target grammaticalization process to be replicated in ordinary contact-induced grammaticalization. Given that the development of a new grammatical category in the replica language is involved, multiple lexical items or constructions in the model language are likely to be involved in the process. As a result, the new category developed in the replica language may not have an exact equivalent in the model language. In our case, in addition to the individual contribution of *zo2* and *laa3*, [V *zo2*... *laa3*] may also be processed holistically and identified with [V *already*]; and the grammaticalized *already* appears to function like a combination of *zo2* and *laa3*. This study highlights the importance of analyzing the congruence of multiple related lexical items or constructions in the model language in cases of ordinary contact-induced grammaticalization.

In analyses based on the model of contact-induced grammaticalization, one needs to first establish the model (M) and the recipient (R) languages. Given that the bilingual children in this study were exposed to both

Cantonese and English simultaneously since birth, it may seem difficult to tell which language serves as the M language. Nonetheless, we maintain that Cantonese is the M language in Heine and Kuteva's (2003, 2005) sense, at least in perfective aspect marking. Why? Because these children know how to express perfective aspect in their Cantonese by age 2, at which point they are far from acquiring the English *have* perfect. Note that this could be true even for children for whom Cantonese is not dominant. Since Cantonese monolinguals use *z02* from before age 2, whereas English monolingual children acquire the *have* perfect around 2;07 (Szeto, 2015), there is a developmental asynchrony involved. Consequently, the perfective *already* emerges in the bilingual children based on its Cantonese counterpart to fill the gap. Thus, both language dominance and developmental sequence have to be taken into account when studying contact-induced grammaticalization in bilingual development.

The results also suggest that typological convergence in grammaticalization is possible in bilingual development. While it is widely acknowledged that language contact may trigger the development of grammatical markers in the replica language along the pathways of grammaticalization present in the model language, the effect of language contact on the manifestations of grammaticalization remains under-investigated. As a typical EMSEA language, Cantonese undergoes limited coevolution of form and meaning over time, producing grammaticalized items with a high degree of morphophonological stability (see Section 2.1). In addition, the areal patterns of grammaticalization in this sprachbund include (i) polyfunctionality of grammaticalized items and (ii) lack of obligatory grammatical categories (Bisang, 2004, 2011, 2015). Interestingly, the English of the bilingual children in this study seems to follow such tendencies, which is a clear sign of typological convergence in grammaticalization.

First, among the three grammaticalized items in the bilingual children's English which emerged as a result of Cantonese influence (namely *already*, *give*, *one*) (Yip and Matthews, 2007), none of them is found to have undergone significant phonetic reduction. If the lexical meaning of the grammaticalized item is retained, a lexical item with its various derived grammatical items with limited or no phonological change will coexist in the bilingual children's speech (cf. Hopper's 1991 concept of layering). In other words, one and the same surface structure can have more than one function. For example, while some bilingual children are found to use *give* as a permissive marker (46) and passive marker (47), its original lexical meaning is also retained (45):



Lexical give

(45) Give me the pencil. (Sophie 2;10)

Permissive give

(46) [appearing in swimsuit] Daddy I give you see. (Sophie 3;04)  
 'Daddy I let you see.'

Passive give

(47) [points to scratched leg] Here is give Timmy scratch. (Sophie 3;06)  
 'Here is scratched by Timmy.'

Similarly, although our findings show that *already* is not necessarily associated with a 'contrary to expectation' reading in the bilingual children, there are still examples showing that such a reading may be present. In (48–49), the children seem to use *already* to correct a wrong assumption, which is a function of the lexical *already*. These examples suggest that the lexical *already* coexists with the grammaticalized *already* in the bilingual children.

(48) Investigator: Hey Sophie, she said she wants to listen to a story.  
 Child: She sleep already. (Sophie 2;06)

(49) Investigator: Are you sleeping now?  
 Child: No, I sleep already. (Llywelyn 2;08)

Another areal feature of EMSEA languages observed in the bilingual children is the lack of obligatory grammatical marking, which includes the frequent omission of tense-aspect markers. In Standard English, tense/aspect markers like the *-ed* and *-en* verbal suffixes are obligatory when we talk about past events or completed actions. If the bilingual children's English follows the grammaticalization tendencies of Standard English, we would expect *already* to be an obligatory category for marking perfective aspect. However, as shown in the corpus data, *already* does not seem to be an obligatory marker of perfectivity (50–52):

(50) You know yesterday I watch Incredible and... (Kasen 2;11)

(51) [Responding to the question 'Where did you go?']  
 We go to the church and dance. (Charlotte 3;00)

(52) I make that yesterday. (Janet 3;10)

In contrast with the obligatory grammatical categories commonly found in Standard Average European languages, grammatical markers of tense/aspect are optional in the bilingual children, suggestive of the grammaticalization tendencies observed in EMSEA languages.

Such a finding has significant implications – when languages of different typological profiles come into contact, typological convergence in grammaticalization tendencies may occur. In our case, given the isolating typology of Cantonese, it is natural that the bilingual children take longer to fully acquire the relatively morphosyntactically complex English tense/aspect forms. As a result, after having acquired the basic aspectual distinction between perfective and imperfective in Cantonese, the bilingual children resort to using *already* after uninflected verbs to mark perfective aspect in order to fill the gap in their English. Following the grammaticalization tendencies of Cantonese, *already* serves as a polyfunctional marker which shows no sign of obligatoriness or phonetic reduction.

## 6.2 *Transient Grammaticalization*

The Cantonese-English bilingual children in Yip and Matthew's (2007) study show a number of interesting grammatical peculiarities in their English not attested in monolingual development, such as the use of prenominal relative clauses, passive-marking *give*, and the perfective-marking *already* discussed in this study. While close parallels of such 'non-target' development are also observed in English-lexified contact languages such as Singlish and Chinese Pidgin English, a crucial difference between the two cases is that such phenomena persist in the contact languages even among adult speakers, but gradually disappear in the bilingual children after they attend school. In our case, the grammaticalized use of *already* is later replaced by the standard English *have* perfect when the bilingual children have acquired the target-like tense/aspect forms (Szeto 2015). We coin the term 'transient grammaticalization' to describe the transient use of grammaticalized items in the replica language based on model constructions provided by the model language in the early developmental stage of the bilingual children.

The reasons why the grammaticalization phenomena in the bilingual children are only transient merit further discussion. We focus on the use of *already* as a perfective aspect marker in our discussion as this is the main theme of this study. In both the Cantonese-English bilingual children and Singlish speakers, the motivation for developing *already* into a perfective aspect marker is arguably the same – the isolating typology of the substrate/replica language(s) poses an obstacle to the acquisition of inflectional morphology involved in the target-like tense/aspect marking, triggering the development of *already* into

a perfective aspect marker to fill the gap.<sup>5</sup> Given the similarities between the internal ecology of the formation of Singlish and the development of English in the bilingual children, the transience of the contact phenomena found in the bilingual children is likely due to the external ecology under which their English develops.

Instead of acquiring their parents' language *in toto*, recent studies in language evolution generally agree that language is characterized by variation and is acquired piecemeal and selectively (Croft, 2000; Mufwene, 2001, 2008; Ansaldo, 2009); one's communication network plays a pivotal role in one's language development because one will select and recombine linguistic features encountered in different instances of 'linguistic interbreeding' (idiolectal interaction) to arrive at an individual grammar. In a multilingual and informal ecology of transmission, one will typically receive diverse and rich input, which favors creative innovations. By contrast, in a monolingual and normative ecology of transmission, where educational institutions may enforce grammatical norms to eliminate deviant features, idiolectal variation tends to be limited and creative innovations may be kept at bay. Given that language acquisition is an incremental process involving several stages of 'self-correction' due to continuous input and feedback, it is important to take into account the linguistic environment to which the bilingual children are exposed when they start to have more opportunities to interact with people other than their family members.

While Hong Kong English is increasingly recognized as a post-colonial English variety with Cantonese influence (Bolton, 2002; Schneider, 2007), it is arguably of little relevance to the present study. Despite its official status, English is only used by 3.5% of the Hong Kong population as their usual language (Census and Statistics Department, Hong Kong Special Administrative Region, 2012) and is rarely used for everyday communication among the vast majority of Hong Kong people. Hong Kong English generally refers to the L2 English variety spoken by Hong Kong Cantonese speakers in limited contexts, but not the 'standard' English spoken by the small number of native speakers. Therefore, although the bilingual children in this study receive input of two languages from birth, after attending an international school at around age 3, their English essentially developed in a monolingual and normative ecology: the children mostly received input from more or less monolingual

---

5 A thorough, systematic study of the acquisition of inflectional morphology must take both the verbal and nominal paradigms into account. In this study we only focus on the verbal paradigm. We have no data on the bilingual children's emergence of nominal inflection yet and we'll leave this issue for future studies.

English-speaking teachers and peers. The deviant features of the bilingual children's English may gradually disappear under such an environment, where 'standard' English grammatical features predominate in the feature pool, while the deviant features are marginalized or even stigmatized. In this regard, the development of the bilingual children's English after attending school may have close parallels with the emergence of an acrolectal variety in a contact language, which shows convergence towards the lexifier language from which it is descended. Comparative studies of the English development of the Cantonese-English bilingual children in late childhood and the acrolectal variety of Singapore English may help to validate this prediction. In short, the transient grammaticalization phenomena observed in the bilingual children are consistent with Mufwene's (2006) notion that grammaticalization stems from idiogrammatization, and whether the latter can spread to the community depends on the contact ecology.

Meanwhile, as noted by Lim (2015), the use of English in Hong Kong is on the rise thanks to the great prominence of computer-mediated communication (CMC) in recent years. Because of input convenience, many Hong Kong people who are Cantonese-dominant in non-CMC domains prefer using English for CMC. The increasing prominence of CMC may therefore drive the Hong Kong community to become more English-dominant, first in CMC domains and subsequently in other domains as well. A possible outcome of the increasing use of English for everyday communication is the emergence of a more nativized and stabilized New English variety in Hong Kong. Given the parallels between bilingual development and language contact, if such a contact variety emerges, we predict that many phenomena observed in Cantonese-English bilingual development may also be present in this emergent variety. The possible emergence of such a New English variety is an area which merits close attention.

### 6.3 *Implications for Contact Language Formation*

As reviewed in Section 2.2, perfective aspect develops through grammaticalization of an adverb with the lexical meaning 'already' in a number of contact languages. Such developments are not only found in Singlish but also in some Ibero-Romance creoles, such as the Spanish-lexified Zamboanga Chabacano (realized as *ya-*), and the Portuguese-lexified Papiá Kristang (realized as *ja*). In addition to reaffirming the strong link between bilingual development and language contact, we argue that the close parallels between the two areas provide support for the evolutionary account of creole formation (Mufwene, 2001, 2008), where the feature pool is the sum of every individual linguistic system in a given linguistic setting, and the linguistic features which are prominent, frequent, and typologically congruent in the given setting are likely to get replicated and propagated (Ansaldò, 2009). The specific linguistic composition

and social factors vary across time, space, and situations, resulting in a unique feature pool and thus a unique creole in each contact situation. A great appeal of this theoretical framework lies in its universality – it can account for change in all kinds of linguistic varieties, regardless of whether they are (artificially) labelled as languages, dialects, vernaculars, creoles, or pidgins, and whether language transmission at the individual or communal level is involved. The remarkable differences between a creole and its input languages can be attributed to the fact that creoles emerge in highly multilingual environments with low normative tendencies, leading to elevated mutation rates in the replication process (Aboh and Ansaldo, 2007). Similarly, in bilingual development, the emergence of innovative and non-target-like features can be perceived as the result of linguistic features from different input languages competing in the child's mind. It is therefore not surprising to find parallels between bilingual development and language contact, especially if the input languages involved are typologically similar. Moreover, under this framework, language is a dynamic system; the later reconvergence of the bilingual children's English with the standard variety is a result of the high degree of monolingualism and normative tendencies in the ecological setting of an international school.

Meanwhile, other well-known theories on creole formation, such as the Language Bioprogram Hypothesis (Bickerton, 1984) and Creole Prototype theory (McWhorter, 2002, 2005), both of which maintain that creoles are a structurally distinct class of languages created by children with impoverished linguistic input, cannot adequately explain the notable differences between different creoles and do not take multilingualism into account, making it impossible to link bilingual development with language contact (for a detailed critique, see Ansaldo *et al.*, 2007). The Relexification Hypothesis (Lefebvre, 1998, 2004) argues that immigrant adults retain the grammar of the substrate language while adopting the lexicon of the lexifier (superstrate) language, forming a new language (i.e. a creole). This hypothesis downplays the role of child language acquisition in creole formation, making it difficult to link up with bilingual development.

## 7 Conclusions

Both quantitative and qualitative differences are observed between the Cantonese-English bilingual children and English monolingual children in the use of *already* – the bilingual children use *already* much more frequently and show a much stronger tendency to place it post-verbally. In addition, they are found to use *already* without any sense of 'contrary to expectation', which is different from the monolingual usage. The loss of the 'contrary to expectation'

reading of 'already' may be a common consequence of its grammaticalization in bilingual development as well as contact languages.

While the consistent association of the bilingual children's *already* with a perfective reading suggests that *already* functions as a perfective aspect marker, it is not in one-to-one correspondence with the Cantonese perfective aspect marker *zo2*. The meaning and placement of the bilingual children's *already* show signs of influence from both *zo2* and the sentence-final particle *laa3*, suggesting that the English [V *already*] construction may have been identified with the Cantonese [V *zo2...laa3*] construction, leading to the grammaticalization of *already*. The results suggest that multiple correspondence may be commonplace in contact-induced grammaticalization. We see the results as validating aspects of the theory of contact-induced grammaticalization as proposed by Heine and Kuteva (2003, 2005): there is a model, but there may be multiple models, so a one-to-one relationship should not be assumed.

Another important finding of this study is that typological convergence in grammaticalization tendencies is a possible scenario in language contact – manifestations of grammaticalization commonly observed in Cantonese are evident in the bilingual children's English grammar, producing morphophonologically stable grammatical markers which are polyfunctional and optional. Our results support the notion that grammaticalization is a type-specific phenomenon which can spread to other languages through bilingual development and/or areal diffusion.

The close parallels between the bilingual children and Singlish in terms of the grammaticalization of *already*, as well as the similar developments found in other contact languages like Chinese-Cuban Pidgin Spanish, Zamboanga Chabacano, and Papiá Kristang, further confirm the strong link between bilingual acquisition and contact language formation. Moreover, the transience of the contact phenomena found in the bilingual children highlights the importance of the role of external ecology in language change.

## References

- Aboh, Enoch O. and Umberto Ansaldo. 2007. The role of typology in language creation: A descriptive take. In Umberto Ansaldo, Stephen Matthews and Lisa Lim (eds.), *Deconstructing Creole* (pp. 39–66). Amsterdam: John Benjamins.
- Ansaldo, Umberto. 1999. *Comparative constructions in Sinitic: areal typology and patterns of grammaticalization*. Stockholm: Allduplo Tryckeri.
- Ansaldo, Umberto. 2009. *Contact Languages: Ecology and Evolution in Asia*. Cambridge: Cambridge University Press.

- Ansaldò, Umberto and Lisa Lim. 2004. Phonetic absence as syntactic prominence: Grammaticalization in isolating tonal languages. In Olga Fischer, Muriel Norde and Harry Peridon (eds.), *Up and down the cline – The nature of grammaticalization*, 345–362. Amsterdam: John Benjamins.
- Ansaldò, Umberto, Stephen Matthews and Lisa Lim (eds.). 2007. *Deconstructing Creole*. Amsterdam: John Benjamins.
- Arends, Jacques, Pieter Muysken and Norval Smith (eds.). 1994. *Pidgins and Creoles: An Introduction*. Amsterdam: John Benjamins.
- Bao, Zhiming. 2005. The aspectual system of Singapore English and the systemic substratist explanation. *Journal of Linguistics*, 41: 237–267.
- Bao, Zhiming and Hong Huaqing. 2006. Diglossia and register variation in Singapore English. *World Englishes* 25(1): 105–114.
- Baxter, Alan N. 2013. Papiá Kristang structure dataset. In Susanne M. Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at <http://apics-online.info/contributions/42>, Accessed on 2016-02-22).
- Bickerton, Derek. 1984. The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7(2): 173–188.
- Bisang, Walter. 2004. Grammaticalization without coevolution of form and meaning: the case of tense-aspect-modality in East and mainland Southeast Asia. In Walter Bisang, Nikolaus P. Himmelmann and Björn Wiemer (eds.), *What Makes Grammaticalization? – a Look from its Fringes and its Components*, 109–138. Berlin: Mouton de Gruyter.
- Bisang, Walter. 2011. Grammaticalization and typology. In Heiko Narrog and Bernd Heine (eds.), *The Oxford Handbook of Grammaticalization*, 105–117. Oxford: Oxford University Press.
- Bisang, Walter. 2015. Problems with primary vs. secondary grammaticalization: the case of East and mainland Southeast Asian languages. *Language Sciences*, 47: 132–147.
- Bolton, Kingsley (ed.). 2002. *Hong Kong English: Autonomy and Creativity*. Hong Kong: Hong Kong University Press.
- Brown, Roger. 1973. *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.
- Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
- Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Census and Statistics Department, Hong Kong Special Administrative Region. 2012. *2011 Hong Kong Population Census*. (Available online at <http://www.census2011.gov.hk/en/index.html>, Accessed on 2016-02-22).

- Cheung, Samuel Hung-nin. 2007. 香港粵語語法的研究 [*A Grammar of Cantonese as Spoken in Hong Kong*] (Revised ed.). Hong Kong: The Chinese University Press.
- Cheung, Y.-L. Lawrence. 2009. Dislocation focus construction in Chinese. *Journal of East Asian Linguistics*, 18(3): 197–232.
- Clements, J. Clancy. 2003. The tense-aspect system in pidgins and naturalistically learned L2. *Studies in Second Language Acquisition*, 25: 245–281.
- Clements, J. Clancy. 2009. *The Linguistic Legacy of Spanish and Portuguese: Colonial Expansion and Language Change*. Cambridge: Cambridge University Press.
- Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. Edinburgh: Longman.
- Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.
- Enfield, Nicholas J. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge Curzon.
- Fisher, Ronald A. 1954. *Statistical Methods for Research Workers*. Oliver and Boyd (ISBN 0-05-002170-2).
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva. 2003. Contact-induced grammaticalization. *Studies in Language*, 27: 29–572.
- Heine, Bernd and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva. 2010. Contact and Grammaticalization. In R. Hickey (ed.), *The Handbook of Language Contact*, 86–105. West Sussex: Blackwell Publishing Ltd.
- Hickey, Raymond. 2010. Contact and Language Shift. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 151–169. West Sussex: Wiley-Blackwell.
- Hopper, Paul. 1991. On some principles of grammaticalization. In Elizabeth Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization* (Vol. 1, 17–35. Amsterdam: John Benjamins.
- Hopper, Paul and Elizabeth Traugott. 2003. *Grammaticalization* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Lefebvre, Claire. 1998. *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole*. Cambridge: Cambridge University Press.
- Lefebvre, Claire. 2004. *Issues in the Study of Pidgin and Creole Languages*. Amsterdam: John Benjamins.
- Lehmann, Christian. 1995. Grammaticalization: synchronic variation and diachronic change. *Lingua e stile*, 20(3): 303–318.
- Lieven, Elena, Dorothe Salomo and Michael Tomasello. 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20(3): 481–508.



- Lim, Lisa. 2015. *Catalysts for change: On the Evolution of New Contact Varieties in the Multilingual Knowledge Economy*. Unpublished manuscript. The University of Hong Kong.
- Lipski, John M. 1999. Chinese-Cuban Pidgin Spanish: implications for the Afro-creole Debate. In John R. Rickford and Suzanne Romaine (eds.), *Creole Genesis, Attitudes and Discourse: Studies Celebrating Charlene J. Sato*, 215–234. Amsterdam: John Benjamins Publishing.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk* (3<sup>rd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum.
- Matthews, Stephen and Virginia Yip. 2009. Contact-induced grammaticalization: Evidence from bilingual acquisition. *Studies in Language*, 33(2): 366–395.
- Matthews, Stephen and Virginia Yip. 2011. *Cantonese: A Comprehensive Grammar* (2<sup>nd</sup> ed.). New York: Routledge.
- McWhorter, John. 2002. The rest of the story: Restoring pidginization to creole genesis theory. *Journal of Pidgin and Creole Languages*, 17(1): 1–48.
- McWhorter, John. 2005. *Defining Creole*. Oxford: Oxford University Press.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Mufwene, Salikoko S. 2006. Grammaticization is part of the development of creoles. *Papia*, 16: 5–31.
- Mufwene, Salikoko S. 2008. *Language Evolution: Contact, Competition and Change*. London: Continuum International Publishing Group.
- Müller, Natascha. 2006. Emerging complementizers: German in contact with French/Italian. In Claire Lefebvre, Lydia White and Christine Jourdan (eds.), *L2 Acquisition and Creole Genesis: Dialogues*, 145–165. Amsterdam: John Benjamins.
- Ringblom, Natasha. 2012. The emergence of a new variety of Russian in a language contact situation: The case of a Russian-Swedish bilingual child. In Kurt Braunmüller and Christoph Gabriel (eds.), *Multilingual Individuals and Multilingual Societies*, 63–80. Amsterdam: John Benjamins.
- Rowland, Caroline F. and Sarah L. Fletcher. 2006. The effect of sampling on estimates of lexical specificity and error rates. *Journal of Child Language*, 33, pp. 859–877.
- Satterfield, Teresa. 2005. The bilingual bioprogram: evidence for child bilingualism in the formation of creoles. *Proceedings of the 4th International Symposium on Bilingualism*, 2075–2094.
- Schiering, René. 2006. *Cliticization and the Evolution of Morphology: a Cross-linguistic Study on Phonology and Grammaticalization*. PhD dissertation, University of Konstanz.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Siegel, Jeff. 2008. *The Emergence of Pidgin and Creole Languages*. Oxford: Oxford University Press.

- Silva-Corvalán, Carmen. 2014. *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge: Cambridge University Press.
- Soh, Hooi Ling. 2009. Speaker presupposition and Mandarin Chinese sentence-final-*le*: a unified analysis of the “change of state” and the “contrary to expectation” reading. *Natural Language and Linguist Theory*, 27: 623–657.
- Soriente, Antonia. 2007. Cross-linguistic and cognitive aspects in the acquisition of WH-questions in an Italian-Indonesian bilingual child. In Istvan Kecskes and Liliana Albertazzi (eds.), *Cognitive Aspects of Bilingualism*, 325–362. Dordrecht: Springer.
- Steinkrüger, Patrick O. 2013. Zamboanga Chabacano structure dataset. In Susanne M. Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at <http://apics-online.info/contributions/46>, Accessed on 2016-02-22).
- Szeto, Pui Yiu. 2015. *The Emergence of Perfective Aspect in Cantonese-English Bilingual Children: Bilingual Development and Language Contact*. MPhil Thesis, The Chinese University of Hong Kong.
- Tang, Sze-wing, Fan Kwok, Thomas Hun-Tak Lee, Caesar Lun, Kang Kwong Luke, Peter Tung and Kwan Hin Cheung. 2002. *Guide to LSHK Cantonese romanization of Chinese characters* (2nd ed.). Hong Kong: Linguistic Society of Hong Kong.
- Thomason, Sarah G. 2001. *Language Contact: An introduction*. Edinburgh: Edinburgh University Press.
- Traugott, Elizabeth C. and John Waterhouse. 1969. ‘Already’ and ‘yet’: a suppletive set of aspect-markers? *Journal of Linguistics*: 193–320.
- Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York.
- Yip, Virginia and Stephen Matthews. 2007. *The Bilingual Child: Early Development and Language Contact*. Cambridge: Cambridge University Press.