

Language Contact:

A Historical Sociolinguistic Reconstruction of

Colloquial Singapore English in Relation to its Chinese

Substrates

Dissertation zur Erlangung des Grades der Doktorin der Philosophie (Dr. phil.)

an der Universität Hamburg
Fakultät für Geisteswissenschaften

vorgelegt von

Lijun Li

aus Guangzhou

Hamburg, 01.03.2021

1. Gutachter:	Prof. Dr. Peter Siemund
2. Gutachter:	Prof. Dr. Michael Friedrich
3. Mitglied der Prüfungskommission:	Prof. Dr. Robert Fuchs
Datum der Disputation:	28.09.2021
Bewertung der Dissertation:	magna cum laude
Bewertung der Disputation:	summa cum laude
Gesamtbewertung:	magna cum laude

1. überarbeitete Version
22.10.2021

路漫漫其修远兮

吾将上下而求索

屈原《离骚》

Acknowledgements

There are many people who contributed to the completion of my dissertation, and I would like to take this opportunity to express my gratitude to these people.

First and foremost, my deepest gratitude goes to my supervisor Prof. Peter Siemund. When I first arrived to embark on the journey of my linguistic studies in Germany in 2011 (to start my MA studies), I never believed that I could accomplish what I have been able to achieve today. He is the one who fully believes in my potentials and abilities, constantly encouraging and offering me professional support, constructive feedback and suggestions.

I am deeply appreciative of my fellow PhD students Richard Bonnie, Kathrin Feindt, Eliane Lorenz, Sharareh Rahbari. We work as an amazing team. I learned so much from our endless discussions and exchanges of information, and it is because of them that I realized the beauty of linguistic studies and the excitement of multilingualism and sociolinguistic research. Special thanks go to my friend Bijia Lou, who gave me extremely helpful support during the final phase of this work. The positive affirmation of all of my colleagues and friends helped bring my dissertation there to fruition. Their enthusiasm kept me going.

I am also grateful to the linguists whom I have met and talked to during conferences and workshops, each one exemplifying the big-hearted selflessness of true scholarship. I would like to express my gratitude here to Prof Ekkehard König, PD Jacob Leimgruber, Prof. Lisa Lim, Prof. Edgar Schneider, Prof. Bao Zhiming, and Prof. Debra Ziegler. Their interest in my work, their profound knowledge and useful comments helped me hone my understanding of key concepts in contact linguistics, World Englishes, and the case in Singapore more specifically.

I am very lucky that I met the love of my life, Stephen, during the course of my academic pursuit. His love, his patience and emotional support, and above all, the sense of home he created for me fortified me for the long road, in work and in life. He kept me well organized in life and getting things done before deadlines. It was his astute advice, encouragement and inspiration that got me over the mountain.

My ambitions for pursuing an academic career abroad, more than 8,000 kilometers away from home could not have been fulfilled without my cherished parents in Guangzhou, China. Although I am far away from them in physical distance and not able to go home often enough, we are still close to each other. Their patient spirit, their understanding and the freedom they gave me tremendously shaped me as who I am today. They listened to my ideas, respected my choices and nurtured me through my uncertainties and frustrations.

Last but not least, I am also grateful to the Faculty of Humanity and the Humanities Doctoral Program at the University of Hamburg for granting me a three-year scholarship which covers significant proportions of living expenses in Germany, and travel expenses for conference participation and workshops.

Table of Contents

<i>Acknowledgements</i>	<i>i</i>
<i>Table of Contents</i>	<i>iii</i>
<i>List of Figures</i>	<i>viii</i>
<i>List of Tables</i>	<i>xi</i>
<i>List of Abbreviations and Symbols</i>	<i>xiii</i>
1 Introduction	1
1.1 Linguistic diversity and language policy in Singapore	2
1.2 CSE and its features	6
1.3 CSE Models.....	12
1.4 Some terminological explanations.....	14
1.4.1 Chinese dialects vs. the Chinese language.....	14
1.4.2 Singapore English vs. Singlish, Colloquial Singapore English	15
1.5 Outline of the research.....	16
2 Research questions	18
2.1 Research background	18
2.2 Objectives and research questions	20
3 A sociohistorical account of multilingualism in Singapore	24
3.1 Historical background.....	25
3.1.1 Precolonial history (before 1819)	26
3.1.2 The colonial port city (1819–1945).....	28
3.1.3 Political change (1945–1965).....	32
3.1.4 The independence era (1965–present)	33
3.2 Population, immigration, and languages.....	35
3.2.1 Overview of demographic trends	36
3.2.2 The Chinese communities and their languages	38
3.2.3 The Peranakan Chinese	43
3.2.4 Multilingualism among the Chinese communities	44
3.2.5 The Malay communities and their languages.....	46
3.2.6 The Indian communities and their languages.....	48
3.2.7 Multilingualism among the Indian communities	48
3.2.8 Eurasians, Arabs, and others	51
3.3 Language policies and language shift in Singapore.....	54
3.3.1 Official languages and the concept of bilingualism in Singapore	54

3.3.2	The origins of language-related policies in Singapore.....	56
3.3.3	The status of English in Singapore.....	61
3.3.4	Speak Mandarin Campaign.....	62
3.3.5	Speak Good English Movement	65
3.3.6	Other language campaigns.....	67
3.4	Language shift in Singapore.....	68
3.5	Summary.....	77
4	<i>Colloquial Singapore English as a contact language.....</i>	78
4.1	A contact perspective on CSE.....	79
4.2	Borrowing, copying, transfer, imposition, shift-induced interference or pattern replication?.....	83
4.3	Scales of borrowability vs. hierarchies of shift-induced interference	91
4.3.1	Scales of borrowability	92
4.3.2	Implicational hierarchies of shift-induced interference	99
4.4	Contact-induced Grammaticalization	109
4.4.1	Ordinary contact-induced grammaticalization vs. replica grammaticalization	111
4.4.2	Replica grammaticalization as recapitulation	114
4.5	Social predictors vs. linguistic predictors of contact-induced change	118
4.6	Dynamic Model of the evolution of postcolonial Englishes	120
4.6.1	The Dynamic Model and Singapore English	123
4.7	Models of CSE	124
4.7.1	The continuum hypothesis	124
4.7.2	Diglossia	126
4.7.3	Cultural orientation and indexicality	127
4.7.4	Systemic transfer and lexifier filter	130
4.8	Linguistic ecology and feature pool model.....	131
4.8.1	The feature pool of CSE.....	134
4.9	Summary and conclusion.....	137
5	<i>The use of already, also, ever and one in CSE.....</i>	139
5.1	Already	140
5.1.1	<i>Already</i> as an aspectual marker in CSE.....	141
5.1.2	Phasal polarity expression <i>already</i>	150
5.1.3	Differences between <i>already</i> in Standard English and CSE <i>already</i>	152
5.1.4	The frequency of <i>already</i> and its preferred sentence position	155
5.1.5	Summary	158
5.2	Also.....	159
5.2.1	Additive marker <i>also</i>	159
5.2.2	The frequency of <i>also</i> and its preferred sentence position	162

5.2.3	Summary.....	163
5.3	Ever.....	164
5.3.1	Standard English <i>ever</i>	164
5.3.2	The experiential aspectual marker <i>ever</i> in CSE.....	166
5.3.3	<i>Ever</i> in affirmative responses to polar interrogatives.....	169
5.3.4	Summary.....	170
5.4	One	171
5.4.1	Numeral and pronominal <i>one</i>	171
5.4.2	<i>One</i> as a nominalizer in CSE.....	173
5.4.3	Emphatic <i>one</i>	176
5.4.4	Frequency of <i>one</i> according to its functions.....	179
5.4.5	Summary.....	180
5.5	Chapter summary and conclusion.....	180
6	<i>The Chinese Substrates</i>	183
6.1	Aspectual system of Chinese	183
6.2	了 <i>le</i>	187
6.2.1	Two different <i>le</i> -s.....	187
6.2.2	Aspectual meanings of V- <i>le</i>	189
6.2.3	Aspectual meanings of S- <i>le</i>	191
6.2.4	V- <i>le</i> and S- <i>le</i> in negative sentences	194
6.2.5	Grammaticalization of V- <i>le</i>	195
6.2.6	Grammaticalization of S- <i>le</i>	198
6.2.7	Summary.....	199
6.3	过 <i>guò</i>	200
6.3.1	Aspectual meanings of <i>guò</i>	200
6.3.2	Grammaticalization of <i>guò</i>	202
6.3.3	Summary.....	204
6.4	也/都 <i>yě/dōu</i>	205
6.4.1	The additive marker <i>yě/dōu</i>	205
6.4.2	With universal quantification	206
6.4.3	Concessive <i>lián</i> with <i>dōu/yě</i>	207
6.4.4	Summary.....	208
6.5	的 <i>de</i>	208
6.5.1	The “relativizer” and “nominalizer”.....	208
6.5.2	Sentence-final <i>de</i>	211
6.5.3	Grammaticalization of <i>de</i>	214
6.5.4	Summary.....	215
6.6	Chapter summary and conclusion.....	216
7	<i>Methodology</i>	218

7.1 Database	218
7.1.1 The Oral History Interviews	219
7.1.2 The selection of the informants of OHI	220
7.1.3 ICE-SG	224
7.1.4 CCL Corpus	225
7.2 Procedure	226
7.2.1 Searching the linguistic variables	226
7.2.2 Calculating the relative frequencies of the linguistic variables	227
7.3 Coding the linguistic variables	228
7.3.1 Already	228
7.3.2 Also	232
7.3.3 Ever	234
7.3.4 One	235
7.4 Summary	236
8 Results and analysis	238
8.1 Analysis of <i>already</i>	238
8.1.1 Frequencies of <i>already</i> by sentence position	240
8.1.2 Substrate-influenced <i>already</i> vs. phasal polarity <i>already</i>	244
8.1.3 Aspectual <i>already</i>	248
8.1.4 Extended reading of aspectual <i>already</i>	254
8.1.5 Aspectual <i>already</i> vs. the English perfect	255
8.1.6 <i>Already</i> in negative sentences	260
8.1.7 Double <i>already</i>	262
8.1.8 Other substrate-influenced morphosyntactic features of CSE <i>already</i>	264
8.1.9 <i>Already</i> vs. social variables	270
8.2 Analysis of <i>also</i>	278
8.2.1 Frequencies of sentence-final <i>also</i>	278
8.2.2 Substrate-influenced <i>also</i>	280
8.2.3 <i>Also</i> vs. social variables	287
8.2.4 Positive correlation between frequencies of <i>already</i> and <i>also</i>	291
8.3 Analysis of <i>ever</i>	294
8.3.1 Frequencies of <i>ever</i> according to semantic categories	294
8.3.2 Frequencies of <i>ever</i> vs. social variables	300
8.4 Analysis of <i>one</i>	306
8.4.1 Frequencies of phrase-final <i>one</i> in OHI and ICE-SG	306
8.4.2 Frequencies of phrase-final <i>one</i> and its semantic categories	307
8.4.3 Frequencies of <i>one</i> vs. social variables	311
8.5 Some other findings	315
8.6 Summary and conclusion	317

9	<i>Discussion</i>	319
9.1	The emergence of CSE and its stabilization	319
9.2	CSE in relation to its Chinese substrates	325
9.2.1	Contact-induced grammaticalization.....	328
9.2.2	Grammaticalization of <i>already</i> in relation to Chinese <i>le</i>	330
9.2.3	Grammaticalization of <i>ever</i> in relation to Chinese <i>guò</i>	336
9.2.4	Grammaticalization of <i>one</i> in relation to Chinese <i>de</i>	338
9.2.5	Grammaticalization of <i>also</i> in relation to Chinese <i>yě/dōu</i>	340
9.3	CSE in relation to speakers' social background	342
9.3.1	CSE in different ethnic groups	342
9.3.2	CSE according to educational level	345
9.4	Summary	347
10	<i>Concluding remarks and outlook</i>	349
	<i>References</i>	354
	<i>Appendix</i>	385
I.	Background information of interviewees	385
II.	<i>Already</i> summary	392
III.	<i>Also</i> summary	395
IV.	<i>Ever</i> summary	398
V.	<i>One</i> summary	401
	<i>Abstract</i>	404
	<i>Deutsche Kurzfassung der Ergebnisse</i>	407
	<i>Veröffentlichungen die aus dieser Dissertation hervorgegangen sind</i>	411
	<i>Eidesstattliche Erklärung über das selbstständige Verfassen der vorliegenden Arbeit</i>	412

List of Figures

Figure 3.1: Representation of historical timeline of Singapore from 1299 to present (data from National Library Board 2017).....	26
Figure 3.2: Ethnic composition of Singapore Population from 1824 to 2019 (Saw 1969; Cheng 1985; Singstat 1990, 2000; Wong 2010, 2015, 2019).....	38
Figure 3.3: Map of Chinese dialects in southern and eastern China (adapted from Gong et al. 2011:226).....	40
Figure 3.4: Chinese resident population by dialect group 2015 (Wong 2015)	42
Figure 3.5: Indian resident population by dialect group (data from Wong 2015).....	49
Figure 3.6: Diverse language background to English plus one of the Mother Tongues	69
Figure 3.7: Changes in home language use over time (data from Cavallaro and Ng 2021; Wong 2011, 2016, 2019)	70
Figure 3.8: Differences in home language use according to age (data from Wong 2010, adapted from Siemund and Li 2020)	71
Figure 3.9: Representation of the relative significance of languages in the different ages of linguistic history in Singapore (Lim 2010:46)	76
Figure 4.1: Representation of borrowing scale (Thomason 2001:70–71).....	92
Figure 4.2: Schematic grammaticalization process of CSE objection particle <i>what</i> (adapted from Kuteva et al. 2018:33).....	115
Figure 4.3: Five stages of Schneider’s (2007) Dynamic Model.....	121
Figure 4.4: Relation between socio-economic factors and the range of sub-varieties of SE available to a speaker (adapted from Platt 1975:369)	125
Figure 4.5: Indexical field of SE. Black = cultural orientation, grey = stances (Leimgruber 2013:244).....	128
Figure 4.6: Internal and external ecology of CSE today (adapted from Ansaldo 2009:112; Lim 2009:199)	135
Figure 5.1: Inchoative <i>already</i> and the English simple past (adapted from Bao 2005:240)	154
Figure 5.2: Experiential <i>ever</i> and the English simple past (adapted from Bao 2005:244)	167
Figure 5.3: Temporal schema of inchoative <i>already</i> and experiential <i>ever</i>	168
Figure 6.1: Schematic representation of Mandarin aspectual markers	185
Figure 6.2: Chinese viewpoint aspect according to Xiao and McEnery (2004).....	186
Figure 6.3: Grammaticalization of <i>V-le</i> in historical texts (adapted from Sun 1996:106)	196
Figure 8.1: Frequencies of <i>already</i> in ICE-SG and OHI (ptw=per thousand words, pp=per page)	239
Figure 8.2: The proportion of <i>already</i> in initial, middle, and final position in OHI and ICE-SG	241
Figure 8.3: The proportion of <i>already</i> in initial, middle, and final position in OHI and ICE-SG (according to different text types)	242
Figure 8.4: The proportion of <i>already</i> in initial, middle, and final position in OHI according to educational level.....	243

Figure 8.5: The absolute frequencies and proportion of substrate-influenced <i>already</i> vs. PhP <i>already</i> in OHI and ICE-SG	245
Figure 8.6: Categories of substrate-influenced uses of <i>already</i> in OHI (absolute figures).....	247
Figure 8.7: Categories of substrate-influenced uses of <i>already</i> in ICE-SG (absolute figures)	247
Figure 8.8: The system of semantic relations between <i>not yet</i> , <i>already</i> and <i>no longer</i> (adapted from Löbner 1989:172).....	261
Figure 8.9: Frequency of <i>already</i> (pp) in different ethnic groups	272
Figure 8.10: Frequencies of overall, sentence-final, and aspectual <i>already</i> pp according to educational level	273
Figure 8.11: Frequencies of overall, sentence-final, and aspectual <i>already</i> pp vs. gender.....	274
Figure 8.12: Frequencies of <i>already</i> overall, sentence-final, and aspectual <i>already</i> pp vs. age group	276
Figure 8.13: The proportion of <i>also</i> in non-sentence-final and sentence-final position in OHI and ICE-SG.....	278
Figure 8.14: The proportion of <i>also</i> in non-sentence-final and sentence-final position in OHI and ICE-SG (according to different text types)	279
Figure 8.15: Frequencies of sentence-final <i>also</i> with novel grammatical meanings in OHI	280
Figure 8.16: Frequencies of sentence-final <i>also</i> with novel grammatical meanings in ICE-SG	281
Figure 8.17: Frequencies of <i>also</i> and clause-final <i>also</i> pp in different ethnic groups.....	287
Figure 8.18: Frequencies of overall and clause-final <i>also</i> pp according to educational level	290
Figure 8.19: Frequencies of overall and clause-final <i>also</i> pp according to age group	290
Figure 8.20: Frequencies of overall and clause-final <i>also</i> pp according to gender	290
Figure 8.21: Frequencies of <i>already</i> and <i>also</i> of individual speakers in OHI born between 1899–1983 (pp).....	293
Figure 8.22: Scatter plot of Relative frequency of <i>already</i> vs. <i>also</i> pp	294
Figure 8.23: Scatter plot of Relative frequency of CF- <i>already</i> vs. CF- <i>also</i> pp	294
Figure 8.24: Absolute frequencies of <i>ever</i> according to different functions in OHI.....	295
Figure 8.25: Absolute frequencies of <i>ever</i> according to different functions in ICE-SG ...	295
Figure 8.26: Grammaticalization chain of the French future form (adapted from Hopper and Traugott 2006:9)	297
Figure 8.27: Frequency of <i>ever</i> pp in OHI according to ethnic communities	301
Figure 8.28: Frequency of <i>ever</i> according to educational level	302
Figure 8.29: Frequencies of <i>ever</i> pp according to gender	303
Figure 8.30: Frequencies of <i>ever</i> pp according to age group	303
Figure 8.31: The proportion of substrate-influenced <i>one</i> vs. standard <i>one</i> in OHI and ICE-SG	306
Figure 8.32: Frequencies of <i>one</i> according to different semantic categories in OHI (absolute figures).....	308

Figure 8.33: Frequencies of <i>one</i> according to different semantic categories in ICE-SG (absolute figures)	308
Figure 8.34: Frequency of clause-final <i>one</i> in different ethnic communities in OHI	312
Figure 8.35: Frequencies of CF- <i>one</i> and substrate-influenced CF- <i>one</i> according to educational level	313
Figure 8.36: Frequencies of overall and clause-final <i>one</i> pp according to gender.....	314
Figure 8.37: Frequencies of overall and clause-final <i>one</i> pp according to age group.....	314
Figure 8.38: Frequency of <i>already</i> in 1st half vs. 2nd half of interview (all speakers)	316
Figure 8.39: Frequency of <i>already</i> in 1st half vs. 2nd half of interview (Chinese group)	316
Figure 8.40: Frequency of <i>already</i> in 1st half vs. 2nd half of interview (all other groups)	316
Figure 9.1: Grammaticalization of Chinese <i>le</i> and CSE <i>already</i>	332
Figure 9.2: Grammaticalization of <i>liǎo</i>	334

List of Tables

Table 1.1: Major language-related policies in Singapore since 1963 (Lim 2010:40; Ng 2013:363)	4
Table 1.2: The aspectual categories of Chinese, English, and Singapore English (Bao 2006:242)	10
Table 3.1: Population and ethnic composition of Singapore from 1824 to 2019 (Saw 1969; Cheng 1985; Singstat 1990, 2000; Wong 2010, 2015, 2016, 2019).....	37
Table 3.2: Singapore Chinese speech groups and their alternate names (data based on Cheng 1985:15–23; W.K. Wong 2015)	41
Table 3.3: Literacy among resident population aged 15 years and over (data from Wong 2011; 2016)	67
Table 4.1: Summary of different terminologies for borrowing and interference	84
Table 4.2: Borrowing vs. shift-induced interference	86
Table 4.3: The occurrence of Spanish form-meaning sets in Mexicano (from Field 2002:142)	95
Table 4.4: Analysis of language contact in Brussels and Strasbourg (adapted from Treffer-Daller 1999:3)	100
Table 5.1: Overview of different functions of <i>already</i> as described by various authors ..	141
Table 5.2: The 10 speakers with highest ratios of <i>already</i> in ICE-SG and ICE-GB (adapted from Siemund and Li 2017:22–23).....	156
Table 5.3: Frequency of <i>already</i> per text in ICE-GB and ICE-SG (Bao and Hong 2006:109)	157
Table 5.4: The five speakers with highest <i>also</i> -ratios in ICE-SG and ICE-GB (adapted from Siemund and Li 2017:23)	162
Table 5.5: Frequency of <i>also</i> per text in ICE-GB and ICE-SG (Bao and Hong 2006:110)	163
Table 5.6: Comparison of counts of <i>one</i> in the Private Dialogue sub corpora between ICE-SG and ICE-GB (Bao 2009:344)	179
Table 7.1: Metadata of selected Oral History Interviews (a sample of 15 speakers)	221
Table 7.2: Overview of year of birth and year of interviews conducted	221
Table 7.3: Seven groups of interviewees according to ethnic background	222
Table 7.4: Summary of metadata of selected interviews in OHI.....	223
Table 7.5: Overview of absolute frequencies of <i>already</i> , <i>also</i> , <i>ever</i> , and <i>one</i> in OHI and ICE-SG	226
Table 7.6: Relative frequencies of <i>already</i> , <i>also</i> , <i>ever</i> , and <i>one</i> in OHI and ICE-SG.....	228
Table 7.7: Overview of categories of the semantic interpretations of CSE <i>already</i>	230
Table 8.1: Multiple linear model on relative frequency of aspectual <i>already</i>	277
Table 8.2: Multiple linear model on relative frequency of sentence-final <i>already</i>	277
Table 8.3: Multiple linear model on relative frequency of clause-final <i>also</i>	291
Table 8.4: Multiple linear model on relative frequency of <i>ever</i>	304

Table 8.5: Excluded variables of multiple linear regression model on relative frequency of <i>ever</i>	304
Table 8.6: Transfer and morphosyntactic compliance (adapted from Bao 2005:259).....	305
Table 8.7: Multiple linear model on relative frequency of CF- <i>one</i>	314
Table 8.8: Multiple linear model on relative frequency of CF- <i>one</i> with substrate-influenced features.....	315

List of Abbreviations and Symbols

ADD:	additive marker
AmE:	American English
ASP:	aspectual marker
ATTR:	attributive marker
AUX:	auxiliary
BrE:	British English
CL:	classifier
COP:	copula
CONT:	continuative aspect
CSE:	Colloquial Singapore English
CRS:	current relevant state
ICE-GB:	The British Component of the International Corpus of English
ICE-SG:	The Singaporean Component of the International Corpus of English
ICE:	The International Corpus of English
NEG:	negation
NOM:	nominalizer
NPIs:	negative polarity items
OHI:	The Oral History Interviews
PF:	perfect
PFV:	perfective
PhP:	phasal polarity
PPIs:	positive polarity items
ptw:	per thousand words
pp:	per page
PTC:	particle
RVC:	resultative verb complement
SSE:	Standard Singapore English

1 Introduction

Colloquial Singapore English (CSE), commonly known as “Singlish”, is an English-based creoloid spoken in Singapore which shows traces of all other ethnic languages spoken in the region (i.e. Chinese, Malay, and Tamil). The emergence of Colloquial Singapore English is one of the consequences of language contact and change in a multilingual environment (Leimgruber 2013; Siemund and Li 2017), and its development is influenced by various factors, including nation-wide language-related policies, mechanisms of language acquisition, and speaker’s attitudes. Today, Colloquial Singapore English is spoken by a majority of both young and old generations of Singaporeans, who render it as their native language and a symbol of their national identity (see Chapter 3). It is a language with different variations that anyone travelling in Singapore will definitely encounter. However, it is also a language that the Singaporean government officially labels as “bad English” and “a corrupted form of English” (Goh 1998), and the use of which is considered detrimental for business and for the country’s image, and therefore active steps have been taken to discourage its use (e.g. the annual Speak Good English Movement). Despite such campaigns, the attitudes of Singaporeans towards CSE seem by no means negatively affected (see Siemund, Schulz, and Schweinberger 2014:356; Tan 2014:336) and it continues to enjoy high public visibility and widespread use in Singapore.

So far, scholarly work has been dedicated to the phonological and grammatical features of CSE, as well as the social conditions that determine their occurrences (Deterding 2007; Leimgruber, Siemund, and Terassa 2018; Lim, Pakir, and Wee 2010; Wee 2004b). Various models have been postulated to capture the emergence of CSE and its relationship to Standard Singapore English (SSE) (see, amongst others Gupta 1989; Leimgruber 2009, 2013; Platt 1975). Research into CSE has also focused on the tense and aspect system of CSE in relation to its Chinese substrate (see Bao 1995, 2005, 2015; Bao and Hong 2006; Wong 2005). However, there is practically no research that probes into its history, documenting if and to what extent CSE is diachronically stable or how different social variables (i.e. age, ethnicity and socio-economic status) of its speakers interact with the

outcomes of CSE (i.e. salient grammatical markers of CSE). The reason for the lack of breadth in a systematic diachronic approach is primarily the data situation, as there are no readily available diachronic linguistic corpora of CSE that could be consulted.

Therefore, the primary aim of this study is to investigate the extent to which CSE is diachronically stable by examining substrate-influenced uses of CSE *already, also, ever, and one*. These CSE expressions differ significantly from standard Englishes in terms of their semantic functions (e.g. aspectual and emphatic markers) and sentence positions (i.e. sentence-initial, sentence-medial, and sentence-final). The approach is to explore a hitherto completely unused data source for linguistic research, namely the Oral History Interviews (OHI) held by the National Archives of Singapore (OHI-NAS), which allows a significant step back in time, as most of the speakers sampled there were born between the 1890s and 1950s.¹ Furthermore, OHI contain substantial metadata that results in a fairly elaborate characterization of speaker backgrounds (see OHI-NAS 2020). The goal is to establish a diachronic reconstruction of the above-mentioned CSE expressions. For that purpose, the results obtained from OHI will be further compared with the Singaporean component of the International Corpus of English (ICE-SG), a well-explored corpus of Singapore English, which records natural speech among university students starting from 1997 (Nelson 2002:3). In view of the discussions regarding the nativization of what is known as “Postcolonial Englishes” (Schneider 2007:3), in which the stage of Singapore English is subsumed under the stage of endonormative stabilization, it would be quite fascinating to see if such nativization processes could be objectified by using diachronic data.

1.1 Linguistic diversity and language policy in Singapore

To understand the emergence of Colloquial Singapore English, it has to be seen against the linguistic ecology in Singapore arising from a mixture of ethnic, political and educational factors past and present. It has long been recognized that Singapore’s history and

¹ The more recent recordings compiled in the 2000s also include one speaker (003409 LSC) born in 1983.

development are closely related to migrants and migration (Platt and Weber 1980; Saw 1969). The pre-colonial history of Singapore indicates that its importance as a trading port dates back to at least the 3rd century because of its strategic location positioned between the Malacca Straits and the South China sea (National Library Board 2017). Its role as a trading hub was superseded by Malacca in the 15th century, but since 1819 when the British arrived, it has been an attractive destination for Arabs, Chinese, Malays, traders from the Indonesian islands, elsewhere in Southeast Asia, and from European countries (Bolton and Ng 2014; Chew 2012). In the colonial period, Singapore saw an ever-increasing ethnic and linguistic diversity. The colonial census in 1824 listed the following ethnic groups: Europeans, native Christians, Armenians, Arabs, Malays, Chinese, natives of Coromandel and Malacca Coasts, as well as natives of Hindoostan and Bengal, Bugis, Balinese, etc. and by 1836 other groups such as Indo-Britons, Jews, Caffres, Parsees, and Siamese were added to the list. Even after the Second World War, Singapore remained an ethnolinguistically diverse society, with approximately 30 ethnolinguistic groups recorded in the 1950s (Kuo 1983).

The Malays were the majority until 1891 when the Chinese residents achieved a majority, who made up 44 percent of the population while the Malays accounted for 42 percent, followed by Indians (11 percent), and Europeans and Eurasians (approximately 1 percent) (Bolton and Ng 2014:308). Since then, the percentage of Chinese had been on the rise and that of Malays had been on the decline until the early 1900s when the proportions of Singaporean Chinese and Malays started to remain relatively stable. The 1901 census recorded 72.1 percent of Chinese and 15.8 percent of Malays. The Indian laborers and migrants were brought in in large number in the 19th century when the Straits Settlements formed part of British India. The population of the Indians peaked in the mid-1800s at 16 percent, but otherwise ranged from 7 percent to 9 percent from the 1900s onwards (Lim 2010:25). Besides the Chinese, Malays and Indians, there is a distinct ethnic group called Peranakan from Malacca and Penang (Lim 2010:24). They are the descendants of the south Chinese immigrants to Malacca and Penang in the 18th and 19th century, who often married local Malay or Indonesian women. They are often referred to as the Straits-born Chinese, or the Peranakan Chinese (Lim 2010:24).

Since the independence of Singapore from the Federation of Malaya in 1963, the Chinese-Malay-Indian-Others (CMIO) model has been the official ethnic framework in Singapore (see Chapter 3). According to the latest Singaporean census of population,² the Chinese ethnic group comprises 74.4 percent of the entire Singaporean population, the Malays 13.4 percent, and the Indians 9 percent. A particularly heterogeneous group (3.2 percent), including Eurasians, Europeans and Arabs, are subsumed under the category known as Others (Wong 2019).

Perhaps due to such diversity of its population, language-related policies have been regarded as a key pillar in nation building since the independence of Singapore as a nation, which aimed at uniting the multiethnic and multilingual groups, as well as fostering a common sense of Singapore national identity (see Lee 2000). Table 1.1 summarizes the language planning policies implemented in Singapore since 1963, including (i) the policy of four official languages, designating one language as the “mother tongue” of each of the three major ethnic groups (Chinese, Malay, and Indian) and English as the language of inter-ethnic communication; (ii) the bilingualism policy in schools (1966), which requires English to be the medium of instruction, and the “mother tongue” as the second language; (iii) the Speak Mandarin Campaign (SMC) in 1979, which promotes the status of Mandarin to replace Chinese dialects among Chinese Singaporeans; and (iv) the Speak Good English Movement (SGEM), a systematic annual movement with the aim of motivating Singaporeans to speak Standard English and steer away from indigenized CSE.

Year	Language Policies
1963	Four official languages (Chinese, Malay, Tamil, English)
1966	Bilingual policy implemented in schools
1979	Speak Mandarin Campaign (SMC)
2000	Speak Good English Movement (SGEM)

Table 1.1: Major language-related policies in Singapore since 1963 (Lim 2010:40; Ng 2013:363)

² By the end of June 2019.

However, the top-down language planning policies in Singapore do not represent the diversity of the language ecology in Singapore and are in fact a reconceptualization and an oversimplification of the differences among heterogeneous communities, making each definable in terms of one single language. Historical evidence shows that such kind of policies did not appear suddenly but had its roots in the precolonial period (see Chapter 3 for a detailed discussion). The direct result of the implementation of the various language policies by the Singaporean government is a language shift towards English and a decline in the use of Chinese dialects (Siemund and Li 2020). The popularity of English has soared and has been replacing the “mother tongues” as a home language since the 1980s. On the other hand, the promotion of the Speak Mandarin Campaign (SMC) since 1979 has successfully increased the use of Mandarin but has also resulted in a decline in the use of the Chinese dialects (see Chapter 3).

In addition to these changes, CSE as a contact language, which emerged from the complex linguistic ecology in Singapore, has been gaining popularity and has led to policy makers fearing that such a corrupted version of English may jeopardize Singaporeans’ ability to command Standard English (see Goh 1998 Prime Minister’s speech at the Marine Parade National Day Dinner). Such fears, in turn, led to the initiation of the Speak Good English Movement in 2000. However, the popularity of CSE seems unaffected by the campaign and it has even gained more popularity, as recent work on attitudes towards CSE reports generally positive attitudes in the Singaporean population (see Siemund et al. 2014; Tan 2014).

The multi-ethnic composition of Singapore’s population as well as its top-down language-related policies make multilingualism in Singapore an attractive subject of much scholarly research (see Leimgruber et al. 2018; Siemund et al. 2014; S. I. Tan 2016; Tan 2014). Chapter 3 explores in more detail the historical background and multilingualism in Singapore.

1.2 CSE and its features

The main features of Colloquial Singapore English have been described in a proliferation of literature over the past 25 years or so, e.g. among the earliest Gupta (1994), and Foley (1998). As CSE is strongly influenced by southern varieties of Chinese (e.g. Hokkien and Cantonese), and Malay (Ansaldo 2004; Leimgruber 2009, 2013; Siemund and Li 2017), its grammatical system differs from Standard English in many ways but shows similarities with that of the Sinitic and Austronesian languages (Leimgruber 2013; Ziegeler 2021). The most distinctive grammatical features include: (i) the optionality of morphological inflections (e.g. plural, third-person singular present tense, and past tense marking) (Fong 2004), as exemplified in (1); (ii) the optionality of copula *be* and determiners, see (2) (Deterding 2007; Ho and Platt 1993; Li 2014:20); (iii) PRO-drop (Gupta 1994) as shown in (3); (iv) the use of discourse particles derived from Chinese and Malay (Gupta 1992b; Leimgruber 2013; Ler 2006; Lim 2007; Wee 2004a), as exemplified in (4), and (v) topic prominence structure (Bao 2017; Bao and Lye 2005), see (5). All examples below come from the Oral History Interviews.

- (1) My brother just go to work at general post office for \$25 a month.
[OHI-000213-EQ]
- (2) a. Next day, I Ø very clever. (pre-adjective copula deletion)
[OHI-002206-MH]
b. This one Ø, Chairman, Housing Board. (pre-nominal copula deletion)
[OHI-001631-OCN]
c. So then I got Ø job then. (lack of determiner) [OHI-000213-EQ]
- (3) [...] once when you cut the line of your friend's kit and it would "hanyut" as they say in Malay. Ø drift away. Ø fly away. (The pronoun *it* was dropped)
[OHI-002057-AC]
- (4) So now he that was fate, that is life *lah*. [OHI-000213-EQ]
- (5) And then uniforms, everything don't know how many hundred.
[OHI-000213-EQ]

Many of the above examples can be directly translated into Chinese or Malay without having to change their word order or grammatical structure. For instance, the verb *go* is unmarked

in (1), which could be related to the lack of tense marking in substrate languages such as Chinese (see Ziegeler 2021:3). The topic prominence structure in (5) could almost find a one-to-one direct translation in Mandarin Chinese, as shown in (6) (personal knowledge). Topic prominence structures also exist in English, yet they originate in the comment clause, i.e. it has to go through the process of movement (e.g. *I like fruits* → *Fruits, I like*). On the other hand, topic is a basic requirement of the sentence in Chinese, and the topic-comment form is considered as a basic form of sentences in Chinese (Bao 2017:628).

- (6) 然后 制服, 所有东西 不知道 多少 百 件
 ránhòu zhífú suǒyǒu dōngxi bù zhīdào duōshǎo bǎi jiàn
 and then uniform everything don't know how many hundred CL
 'And then there are uniforms. I don't know how many hundred pieces of
 clothes these add up.'

The features in examples (2)–(4) could also be ascribed to the related features in the substrate languages. For instance, copula deletion, especially in syntactic environments such as pre-adjective, pre-V+ing, pre-nominal and pre-locative is a well-attested feature in the Chinese languages. Ho and Platt (1993:31) report 13% of *be*-omission in the same syntactic environments as in Chinese. These features show striking parallels between CSE and Chinese, yet studies have suggested that substrate influence is perhaps not the entire story in a contact situation – universal features such as under- or overgeneralization (Sharma 2009, 2012), typological similarities (Siemund 2013), and grammaticalization patterns (Ziegeler 2014, 2021) – need to be taken into account as well.

At the interface of the lexico-grammatical level, there is the use of pragmatic particles, many of which are borrowings from the Chinese vernaculars, Malay, and Tamil (Gupta 1994; Ler 2006; Li, Lorenz, and Siemund 2021; Lim 2007; Wee 2004a). The assertive marker *lah*, as exemplified in (4), is the most frequently used particle in ICE-SG (Ler 2006), which serves a wide range of pragmatic functions, from conveying assertive meaning (Gupta 1992b:37) to a mood/attitude indicator (Wee 2004a:125). Besides *lah*, other pragmatic particles have been appropriated into CSE, including (i) *lor*, indicating obviousness, (ii) *meh*, close to the pronunciation of [mɜ] as a marker of rhetorical question, (iii) *hor*, a question tag

to verify something while expecting confirmation, (iv) *wat* (also spelled as *what*), presenting obviousness and contradiction to something that has previously been asserted (see Gupta 1992b:37; Lim 2007:464; Leimgruber 2013:88).

The origins of these particles are disputed, as previous scholars have not yet identified a single substrate as the likely source, but suggest, though not convincingly, Malay, Hokkien, and Cantonese equally likely as their sources (Lim 2007:464). There was also a tendency in the past for researchers to identify these particles as coming from the language they happened to be most familiar with. For example, Richards and Tay (1977), specialists of Hokkien, regard these particles as originating in Hokkien, Kwan-Terry (1978, 1989), who studied Cantonese, argues that the majority of particles are of Cantonese provenance. And Baskaran (1987), who has no knowledge in any of the Chinese vernaculars, considers these particles imports from Tamil. A more recent study (Leimgruber 2015) has found traces of the Mandarin particle *bah* in CSE, which conveys uncertainty and tentativeness, via two new online data-sets – the Global Web-Based English corpus (GloWbE) (Davies and Fuchs 2015), and discussion forums posts (from <http://www.sgforums.com>). Using Mandarin particles in CSE seems to be a relatively new development, which is probably due to the resurgence in the prominence of Mandarin Chinese and recent immigration of people from all parts of China.

Another interesting feature for CSE scholars is reduplication. Reduplication is a feature not limited to CSE, but also found in many of the indigenized varieties of English, such as Indian English, Sri Lankan English, Philippine English, African Englishes and Jamaican English (see among others Bamiro 1995; Ho and Platt 1993; Kachru 1992; Platt, Weber, and Ho 1984). What is distinctive about CSE reduplication is that it occurs mainly in three grammatical classes: (i) finite verbs and verb groups, (ii) nouns, and (iii) modifiers (Leimgruber 2009). Their functions range from expressing the meaning ‘a little bit’ or ‘a few’ as shown in (7) and suggesting informality in (8) to conveying emphasis, as exemplified in (9). Bao and Hong (2006), on the other hand, interpret reduplication in (8) as marking the tentative aspect in the sense of trying to do something.

- (7) His children, they still have those *little little* books from childhood. [OHI-001163-RZ]
- (8) Satellites just *look look look*. You don't go down and act. [OHI-003223-CSS]
- (9) *See, see, see* there's a fire, we could see you know. We were in the sea. Burning, smells like rubber. [OHI-002068-LKC]

Besides grammatical features, lexical borrowings from Chinese, Malay and Tamil were included into the CSE lexicons, designating items of daily life. For example, the word *ang moh* from Hokkien literally meaning 'red hair' is used in CSE to describe Caucasians. Malay contributes many lexical borrowings related to Malay cuisine, such as *teh-tarik*, literally 'pulled tea', which is a typical Malay drink prepared with tea and condensed milk; *ice kacang* (lit. 'ice beans'), a colorful dessert made of ice, condensed milk, red beans and jelly cubes and *otah*, a typical Malay dish made with fish paste wrapped in banana leaves and cooked over charcoal. In fact, the Malay word *makan* is commonly used to mean food, as shown in (11), or the act of eating (Faraci 2014). In Tamil, the word *goondu* means 'fat', but in CSE it refers to a person who is not very smart. Many cases of these lexical borrowings are presented in the OHI, see (10) and (11):

- (10) There was less *ang moh already*, less soldiers come in so the business also slowing down. [OHI-002951-JY]
- (11) Because, mostly shift duty we don't have the proper time for *makan* time. Sometimes, we have a lot of things to do, very busy, we forget about the *makan already*. [OHI-001953-LAS]

The examples in (10) and (11) also introduce the use of *already* as a completive/perfective aspect marker in CSE, which cuts across the lexically-based limitations of tense and aspect marking (Bao 2015; Teo 2019; Ziegeler 2021). Like Chinese, CSE verbs are often not inflected and temporal adverbs such as *last time*, *next time*, *before* and *after* are used to give temporal clues (Leimgruber 2009). As exemplified in (11) "we *forget* about the *makan already*", CSE *already* adopts the aspectual meaning of 了 *le* in Chinese which is concerned with the lexical aspect rather than grammatical aspect. Additionally, CSE *already* prefers phrase/sentence-final position (Bao and Hong 2006; Li 2014). The parallels between CSE

and Chinese are so striking that Bao argues that the entire Chinese system of aspectual marking has been replicated into CSE by using English morphosyntactic material (Bao 1995, 2005, 2015).

		Chinese	CSE	English
(a)	Perfective			
	i. Completive	V <i>le</i>	S <i>already</i>	V-ed, V-en
	ii. Experiential	V <i>guò</i>	<i>ever</i> V	≈ <i>ever</i> V-en
	iii. Emphatic	<i>yǒu</i> V	<i>got</i> V	—
		V <i>wán</i>	<i>finish</i> V	—
(b)	Inchoative	S <i>le</i>	S <i>already</i>	—
(c)	Inceptive	S <i>le</i>	S <i>already</i>	—
(d)	Imperfective			
	i. Dynamic	<i>zài</i> V	V-ing	V-ing
	ii. Stative	V <i>zhe</i> ... (<i>ne</i>)	≈ V-ing	≈ V-ing
	iii. Stative	V <i>zhe</i>	≈ V-ing	≈ V-ing
(e)	Tentative	V-V	—	—

Table 1.2: The aspectual categories of Chinese, English, and Singapore English (Bao 2006:242)

As illustrated in Table 1.2, not only the perfective/completive aspectual meaning from Chinese *le* was incorporated into CSE *already*, but also the inchoative “the beginning of a new situation” (Bao 2005:200) and inceptive “the recent start of a new event” were imported (Bao 2015:20, see (12)), which is not overtly marked by functional morphemes in English (Bao 2005, 2015). Other substrate-influenced aspectual markers in CSE include: (i) the experiential *ever* modeled on the use of an experiential marker *guò* in Chinese, which will also be examined in this study (see Chapter 5); (ii) the emphatic *got V* and *finish V* are transferred from *yǒu V* and *V wán* respectively, and (iii) the tentative *V-V* which had not yet been detected in Bao (2005). According to Bao (2005), the tentative *V-V* should not be transferred into CSE, because a *V-V* structure violates the morphosyntactic rules of English, but it was later found in Leimgruber (2009, 2013), Li (2014) and Bao (2015). The

imperfective *zài V* (dynamic) and *V-zhe* correspond to *V-ing* in CSE, according to Leimgruber (2009, 2013), though the imperfective in English has the same form *V-ing*.

- (12) a. Completive *already*: [...] my uncle will drive, then we go and eat. Eat *already* come back they will play until morning. [OHI-002951-DY]
 b. Inchoative *already*: From the time they are ready for mainstream school, you can *already* assess them. [OHI-002749-DT]
 c. Inceptive *already*: It rain *already*. (Bao 2005:241)

Such uses of *already*, along with Mandarin Chinese *le*, Indonesian/Malay *sudah*, Tai *léew*, and Vietnamese *đã* and *rồi* are considered as “iamitive” (from Latin *iam* ‘already’), which is a cover term for inchoative, perfective/completive aspects (see Dahl and Wälchli 2016 for more detailed descriptions of iamitive). There is reason to believe that CSE *already* may origin from Malay *sudah*, yet a recent study of Teo (2019) reveals that Singaporean Chinese tend to use *already* in sentence-final position, in contrast to Malay speakers, who have a much weaker preference for using *already* in sentence-final position instead of sentence-initial or sentence-medial position. Moreover, CSE *already* occurs more frequently in negative contexts among the Chinese than Malay speakers. It is also worth mentioning that the aspectual use of *already* appears to be a common linguistic phenomenon in other Asian and African varieties of English (e.g. Hong Kong English, Cameroon English, Nigerian English, Ghanaian Pidgin), which is believed to have been introduced by the substrate languages spoken in the relevant areas (Li and Siemund 2021).

There are many phonological features of CSE that are strikingly different from mainstream varieties of English. They also demonstrate influences from the local substrate languages. Features include shorter vowels, so that the following minimal pairs are pronounced as homophones: *beat* and *bit* as [bit], *bed* and *bad* as [bed], *pet* and *pat* as [pet], *cart* and *cut* as [kat], although this phenomenon depends on the condition of its phonological environment. Other features are reduced consonant clusters and unreleased or glottalized final consonants (Leimgruber 2013:66). The dental fricative [θ] in *health* is realized as the labiodental fricative [f] ([hɛlf]) (Bao 1998:154). For more details on the phonological

features and other grammatical features in CSE, interested readers are referred to Bao (1998), Deterding (2005), Lim (2004), and Leimgruber (2013).

1.3 CSE Models

The distinctiveness of CSE within its sociolinguistic context and its complex relationship with Standard Singapore English have been the subject of many scholarly debates. During the early 1970s, a purist interpretation of the linguistic features tended to dominate the field (Crewe and Vargish 1977; Tongue 1974), where CSE with its salient linguistic features was regarded as a divergence from British English, considered “erroneous” and “non-standard” (S. I. Tan 2016:70). By the late 1970s to 1980s, there was a growing recognition that these features were natural results of the evolution of English in a postcolonial and multilingual setting. And according to one of the earliest proposals in this period, the linguistic situation can be captured in terms of a “post-creole continuum” (Platt 1975). Within this proposal, CSE was considered to be the outcomes of the development of English-medium education during the British colonial period, as using English as the primary medium of instruction resulted in the penetration of English into the local sociolinguistic landscape. Another prominent model to represent the diversity regarding CSE more adequately is Gupta’s (1989, 2001) explanation in terms of her diglossia hypothesis. Her observation is that CSE and SSE serve different functions: CSE is used as a home language and amongst friends whereas SSE is used in more official contexts like administration, business, and education.

More recent models include the “culture orientation model” developed by Alsagoff (2007, 2010) and the “model of indexicality” as expounded in Leimgruber (2009, 2013). Alsagoff (2007) argues that two roles of English – as a global language and as an inter-ethnic *lingua franca* – are representative of, and closely related to Singaporean macro-cultural perspectives and identity; and that the variation in the use of English in Singapore is a reflection of the conflict between “being global” and “being local” (Alsagoff 2007:34). In Leimgruber’s (2009, 2013) model of indexicality in Singapore English, he argues that Singaporean speakers create their utterances drawing from a pool of features since there is

no identifiable matrix language (Leimgruber 2009:161). We can see that the latter two models allow for more flexibility and intentionality in the linguistic behavior of the Singaporean speakers.

Besides the previous models of CSE which concern the relation of CSE to SSE, the “systemic transfer and lexifier filter model” proposed by Bao (2005, 2015) offers interesting insights into the similarities between Mandarin Chinese and CSE, arguing that Chinese provides the whole grammatical system for CSE while English acts as a filter to prohibit the linguistic features that do not meet the morphosyntactic rules of English (Bao 2015:58).

We can see that the “post-creole continuum”, the “diglossia hypothesis”, the “cultural orientation”, and the “indexicality” models are not strictly limited to the Singaporean contexts but have been established on models proposed for other varieties of English. However, they often neglect the role that individual speakers and their different social variables, i.e. age, gender, ethnicity, and educational level, play in contact-induced change. Moreover, they are limited to the synchronic contexts.

The “dynamic model” of Schneider (2003, 2007) depicts a unified theoretical and diachronic account of the development and evolution of postcolonial Englishes (PCEs), pointing out that CSE manifests many characteristics of the model’s stage 4: endonormative stabilization, in which the innovative linguistic features of the local norms become increasingly accepted (see Chapter 4.6). However, the data that he uses to support the model has been drawn from mainly synchronic sources and secondary resources. Therefore, besides discussing these previous models of CSE in Chapter 4, this study also tests to which extent these models are applicable with primary diachronic data, and the empirical part of the study (see Chapter 7 and Chapter 8) puts CSE in a wider sociolinguistic context, with a focus on various social factors of individual speakers that contribute to the emergence and formation of CSE.

1.4 Some terminological explanations

This section discusses two terminological issues that are important to the present study, namely language vs. dialect, Singapore English vs. Singlish, Colloquial Singapore English.

1.4.1 Chinese dialects vs. the Chinese language

The boundary between “dialect” and “language” is not self-evident. There is no sharp dividing line between “possible to understand” and “impossible to understand” and dialects will turn into languages given enough time and the right social circumstances (Thomason 2001:2). Haugen (1966:922) also emphasizes the “ambiguities” and “obscurities” attached to the distinction between language and dialect, making the identification of languages an extremely difficult task.

These ambiguities and obscurities manifest themselves most obviously in the Chinese language. For example, Wardhaugh and Fuller (2015) compare the difference between Mandarin Chinese and Cantonese Chinese to the difference between German and Dutch as well as Portuguese and Italian. Though both Mandarin Chinese and Cantonese Chinese belong to the Sinitic language family, these two varieties are not mutually intelligible (e.g. Pronouns such as *I*, *you*, and *he* in Mandarin are *wǒ*, *nǐ*, and *tā*, but in Cantonese they are, respectively, *ngóh*, *léih*, and *kéuih*), and referring to Cantonese Chinese and Mandarin Chinese as dialects may not make much sense from a linguistic perspective, as traditionally, the linguistic distinction between languages and dialects may hinge on the degree of mutual intelligibility (see Lyons 1981).³ However, all Chinese “dialects” share the same standard writing system, where each morphem or syllable has its own symbol that, more or less, stays the same from one Chinese “dialect” to another (McWhorter 2012).

although in terms of the chronology of language contact, Southern Min, Cantonese, and Hakka rather than Mandarin are the substrates that are more likely to exert a strong influence on Colloquial Singapore English (Lim 2007), the features that are investigated in

³ Other criteria to distinguish languages from dialects include sociopolitical reasons (e.g. national boundaries and prestige attached to a certain variety), cultural differences, distinctive writing systems (McWhorter 2012).

this study are pan-Chinese. Here, the word “language” is used as a collective noun which is often used to refer to a group of closely related languages. The Chinese language, as it is often called, is a group of related dialects or varieties, which share close syntactic and morphosyntactic similarities (Bao 2015:51). While Mandarin Chinese is the official representative, the other constituting members are referred to as dialects or vernaculars. For example, Cantonese and Hokkien are both Chinese dialects/vernaculars.

1.4.2 Singapore English vs. Singlish, Colloquial Singapore English

Since terms like Singapore English, Standard Singapore English and Colloquial Singapore English are very often referred to in this study, it is necessary to define and draw some distinctions between them. Unlike some authors who use the term Singapore English as an equivalent to Colloquial Singapore English (or Singlish as the local code) (see Salazar 2014), this study uses the term “Singapore English” in a broader sense and defines it as “English in Singapore”, and in Gupta’s (1998:2) words “a range of Englishes”, covering different varieties of English spoken in Singapore, from the standard norm promoted by the language planners to the non-standard, local forms of English spoken and adopted by the Singaporean population. It generally refers to any English/Englishes spoken by Singaporeans in the city-state. On the other hand, Standard Singapore English (SSE) and Colloquial Singapore English (CSE) refer to the High and Low varieties in diglossic opposition, respectively. As mentioned earlier, CSE is also known as Singlish, but I will avoid using the term Singlish, as it is usually associated with strong social meanings (i.e. stigmatized local code).

It is difficult to pinpoint whether the interviewees selected in this study are CSE speakers, as speakers often mix standard and non-standard codes in one single utterance, which is a common phenomenon observed in recent CSE studies (Leimgruber 2013; Siemund and Li 2017). Moreover, the term “Colloquial Singapore English”, or “Singapore Colloquial English” was coined in the 1970s by contact linguists, e.g. among the earliest Platt (1975) and Gupta (1989). However, as evident in the interviews drawn from the National Archives of Singapore, the emergence of Colloquial Singapore English dates back

even earlier than the creation of the term itself. The term CSE can be safely assumed to refer to the utterances of speakers, instead of individual speakers themselves in the Oral History Interviews. For example, the substrate-influenced words (*already*, *also*, *ever*, and *one*) under investigation are salient markers of CSE.

1.5 Outline of the research

This study addresses the sociolinguistic reconstruction of Colloquial Singapore English in 10 chapters. Following this introduction, **Chapter 2** describes the research questions, the objectives and rationales of this study. **Chapter 3** sketches a sociohistorical overview of Singapore, including a brief history of Singapore, the development of its population composition, immigration patterns with a special focus on its top-down multilingual policies, which have resulted in a dramatic change in Singapore's linguistic situation and an unprecedented transformation of linguistic repertoires of individual Singaporeans.

Chapter 4 then presents a discussion of general language contact theories, including “contact-induced change” (Thomason 2001), “scale[s] of borrowing and pattern replication” (Field 2002; Matras 2009), “contact-induced grammaticalization” (Heine and Kuteva 2005), and various CSE models from “post-creole continuum” (Ho and Platt 1993; Platt 1975) and “diglossia” (Gupta 1989), to the more recent “cultural orientation model” (Alsagoff 2007, 2010) and “indexicality model” (Leimgruber 2009, 2013), to “system transfer and lexifier filter” (Bao 2005, 2015), as well as the “Dynamic Model of Postcolonial English” (Schneider 2007), which provides a diachronic outline for the emergence and development of Singapore English.

Chapter 5 focuses on the extended uses of CSE adverbials *already* and *ever* as aspectual markers, the adverb *also* with additive and concessive readings, and the pronominal *one* as a relative clause pronoun, and an emphatic marker. Interestingly, there are restrictions on the contexts of use of these CSE markers, which differ from those in the Chinese substrates. These are explored in **Chapter 6**.

Chapter 7 presents the main database of this study – the Oral History Interviews drawn from the National Archives of Singapore (OHI-NAS) that span nearly the entire 20th century, which help to fill a gap in the literature on the diachronic development of CSE markers (i.e. *already*, *also*, *ever*, and *one*). Corpora such as the Singaporean component of the International Corpus of English (ICE-SG) and the Chinese Corpus held by the Center of Chinese Linguistics (CCL) serve as complementary databases for comparison purposes. This is followed by a description of the methods applied in the empirical part.

Chapter 8 is a much-expanded treatment of the topic, which presents the data analysis, discusses the results relating to the previous theories, and gives answers to the research questions. Questions will be addressed concerning the stability of Colloquial Singapore English, the correlations between the production of CSE markers and social backgrounds of the interviewees as well as the effects of the Chinese substrates on these four CSE markers. **Chapter 9** discusses these results in relation to the previous theories of the emergence of Colloquial Singapore English. **Chapter 10** concludes the study and presents the outlook of further CSE studies.

2 Research questions

This chapter comprises two sections. First, I introduce the research background of varieties of English with a focus on Singapore English and discuss the rationale of this study. Secondly, I present my research questions, which guide this study of Colloquial Singapore English in relation to their Chinese substrates from a diachronic perspective.

2.1 Research background

The enormous expansion of English worldwide since the 19th century has resulted in variations of English in form and function in different multilingual contexts, and the emergence of new varieties of English. One of the most influential models capturing the stratification of varieties of English remains Kachru's (1985) three concentric circles. This model visualizes the complex reality of World Englishes by categorizing countries (or regions) into the Inner Circle (English as a mother tongue), the Outer Circle (English as an additional institutionalized language), and the Expanding Circle (English as a foreign language) (Kachru 1985, 1986, 1992). Based on this model, groundbreaking research has sought to describe and legitimize this categorization of varieties English, particularly in Asia, including mainland China, Hong Kong, Malaysia and Singapore (see e.g. Bolton and Graddol 2012 on mainland China; Chan 2013 on Hong Kong; Hashim and Leitner 2011 on Malaysia). Kachru also underlines the dynamics of Asian Englishes by stating that "the karmic cycle of the English language has gained an unprecedented momentum," and that "the functional dynamics of Asian Englishes [...] are in constant change" (Kachru 2005:xvii-1).

However, Kachru's model is not without its limitations. Especially in the modern era, the sociolinguistic reality is getting more and more complex, fostered by the unprecedented rate and scale of globalization, the advancement of technology, and hence the high speed of information flow through the Internet and social media (Omoniyi and Saxena 2010). A growing number of recent studies have challenged the applicability of Kachru's model (e.g.

Bruthiaux 2003; Rajadurai 2005; Ferguson 2006; Jenkins 2009; Mair 2013, 2017; Siemund 2018). One of the limitations, for example, includes the problematic category of “national standards” (Mair 2017:5), as migration and media have weakened the boundary of the “national links” (Mair 2013:253–54) between vernaculars and their territories or primary communities of speakers. Siemund (2018:133–34) reminded us that more than one variety may be found in one territory that are principally distinguishable (e.g. more than one varieties of Indian or Chinese Englishes); and in turn, several territories may share one variety of English (e.g. Nigerian Pidgin across West African countries, or Colloquial Singapore English in Singapore, Malaysia, and Indonesia). A further point that was neglected in Kachru’s model is the sociolinguistic complexity within the nation state and sociolinguistic change over time (Ferguson 2006), and therefore Mair (2013) called for adding to research in varieties of English “a focus on social styles in interaction to our established concern with varieties, sounds, words, and constructions” (2013:256). Moreover, English has increased its “fluidity” (Jenkins, Cogo, and Dewey 2011) as it becomes more intertwined in terms of economic, cultural, political, and social spaces. It continues to be used as a *lingua franca* in a multitude of international contexts, both in professional arena, e.g. international organizations, such as the European Union (EU), and the Association of Southeast Asian Nations (ASEAN), and in private domains, e.g. cross-cultural marriage as well as communication via social media such as YouTube, Facebook, Twitter, and Instagram. In sum, there is an increasing hybridization of language and culture in all three circles of Kachru’s (1985) model.

With regards to the evolution of English varieties in the postcolonial context, Schneider’s (2007) innovative Dynamic Model is perhaps one of the most inspiring and well-described theoretical frameworks. Schneider (2007) captures five developmental stages for all postcolonial Englishes, which are: (1) foundation; (2) exonormative stabilization; (3) nativization; (4) endonormative stabilization; and (5) differentiation. Four sets of parameters characterize these different stages, which are: (a) extralinguistic background; (b) identity construction; (c) sociolinguistic conditions; and (d) linguistic consequences (Schneider 2007:56).

Schneider (2007) draws special attention to Englishes in the Outer Circle in Kachru's model, and in his view, they represent the most interesting and critical stage (stage 3 to 4). Belonging to the Outer Circle in Kachru's model, Singapore English has developed into a norm-developing variety within just a few decades. According to Schneider, Singapore English "has gone through a vibrant process of structural nativi[z]ation" and "clearly reached stage 4 of the cycle" (Schneider 2007:158–160). However, research into Singapore English so far has predominantly focused on synchronic studies which investigate the structural properties on the different levels of language organization. Even Schneider (2007) himself provides little diachronic empirical evidence to strengthen this argument about Singapore English entering stage 4 (endonormative stabilization) and possibly stage 5 (differentiation). Recently, investigations in Singapore English started to focus on diachronic perspectives (e.g. Ziegeler 2015, 2016 on historical replication in Singapore English), but their analyses have been based on synchronic material as well.

This research gap is surprising, in particular because English historical linguistics has a rich and long-standing tradition of corpus-based work (see the surveys in Kyto and Pahta 2016). By using diachronic corpus resources, we can learn about diachronic variation in genre, registers, and varieties (Biber and Gray 2011; Hundt and Mair 1999; Tagliamonte 2012). Yet, so far, most corpus-based research into English historical linguistics has focused on Englishes in the Inner Circle (see Bergs and Brinton 2017). In view of both Kachru's (1985) and Schneider's (2007) discussions, Singapore English is currently in the stage of endonormative stabilization. It would be quite fascinating to see if such nativization processes could be objectified by using diachronic data.

2.2 Objectives and research questions

Therefore, this study seeks to investigate the diachronic development of CSE that includes data produced by speakers who were born in the late nineteenth and early 20th century. The objective is to offer some first steps towards a diachronic reconstruction of CSE. Since it is impossible to provide a comprehensive picture of the diachronic developments, this study

exclusively focuses on four highly salient markers of CSE, namely the aspectual marker *already*, the additive particle *also*, the experiential marker *ever*, and the emphatic marker *one*. A comparison of Oral History Interviews (OHI) drawn from the National Archives of Singapore (NAS) with the more well-known International Corpus of English (ICE-SG) will be provided. As the Oral History Interviews represent a type of CSE four to five decades earlier than that in ICE corpus, a diachronic trend can be outlined.

Furthermore, studying both Colloquial Singapore English – an intriguing contact language which is derived from varieties of languages coming from different language families – and its Chinese substrates from a diachronic perspective could offer us more insights into the patterns of language contact (e.g. contact-induced grammaticalization, see for examples Pietsch 2009; Ziegeler 2015), which in turn could help tremendously in predicting future language change in contact situations. Additionally, with the help of the Oral History Interviews, we can compare chains of grammaticalization of the above CSE grammatical markers with that of the Chinese substrates. Such a comparison allows us to validate or invalidate previously proposed contact-induced grammaticalization models, e.g. “replica grammaticalization” (see Heine and Kuteva 2005).

A further goal of this study is to examine the social factors, such as gender, educational level and ethnicity of individual CSE speakers in conditioning the production of the linguistic variables in CSE. As Sapir (2008:147) puts it, “Everyone knows that language is variable”. And at the heart of language variation and change, it is the individuals’ conscious, or more usually, subconscious adjustments in their linguistic behavior that constitute variation, and potentially change (see Labov 1963). Some sociolinguists have argued that it is the position of the standard language variety in the linguistic market – that is, the extent to which the standard variety is valued in people’s daily life – that plays a more significant role in shaping patterns of language variation and change than their social background (Sankoff and Laberge 1978:239–50). Yet, over the past four decades, the primary interest of sociolinguists lies in language variation according to social factors, that is, to correlate the patterned linguistic heterogeneity as the dependent variable with extra linguistic phenomena (the social and cultural) as the independent variables (Chambers 2003;

Sankoff 1988; Tagliamonte 2012). And in order to shed light on social variation in language they gather their data from sociolinguistic interviews with speakers whose social characteristics constitute a representative sample of the society or community. As Labov (2001:33) notes, “[t]o understand the forces operating in linguistic change, we will necessarily be focusing upon a handful of individuals”.

New studies situated at the interface of World Englishes and variationist sociolinguistics have investigated the impact of social factors on linguistic variation by exploring the ICE corpora (e.g. Mair 2009, 2009; Fuchs and Gut 2012, 2015; Schweinberger 2012). While there have been studies on self-reported attitudes of Singaporeans towards Colloquial Singapore English (e.g. Siemund et al. 2014; Leimgruber et al. 2018), only a few studies (e.g. Hansen 2018; Schröter 2018) in Singapore English have been conducted from a variationist sociolinguistic perspective. The currently available corpus of Colloquial Singapore English – the Singaporean component of the ICE – does not provide metadata about the speakers (Hansen 2018:89), which makes it impossible to analyze whether the production of CSE markers is correlated with gender, educational level, or ethnicity. Given that the unused data source from the Oral History Interviews provides precious individual metadata (e.g. gender, year of birth, interview date, educational level, and ethnicity), it allows for more variationist sociolinguistic insights into CSE.

Thus, I will approach CSE from a sociolinguistic point of view by exploiting the OHI data. For that purpose, three major questions will be addressed. The first concerns the general trend of CSE, i.e. whether CSE has been diachronically stable over the past century. The second question attempts to address the theories of language contact in the Singaporean case: how to explain with currently available language contact theories the differences and/or similarities of CSE with its Chinese substrates. The final question tries to unfold the variationist sociolinguistic puzzles of CSE. Each question will be broken down into quantifiable sub-questions by focusing on the CSE markers *already*, *also*, *ever*, and *one*. The research questions are formulated as follows:

- (1) What is the developmental trend of CSE during the 19th to the 20th century? Do the data in OHI provide empirical evidence to support the Dynamic Model, according to

which CSE has reached stage 4 (endonormative stabilization) and possibly stage 5 (differentiation)? That means: (i) Did the uses of prominent markers of CSE (*already*, *also*, *ever* and *one*) increase over time? (ii) Did these markers of CSE obey usage principles and distributions different from today?

- (2) To what extent are the usages of *already*, *also*, *ever*, and *one* related to their Chinese substrates? In other words, how similar/different are these CSE markers with/from their Chinese substrates in terms of functions and structures as well as frequencies of use?
- (3) Which social factors can explain the variation among individual speakers in using CSE markers? Particularly, are there correlations between the differences in the frequencies of the CSE markers (i.e. *already*, *also*, *ever*, and *one*) and the different social variables (i.e. gender, ethnic background, education level) of speakers of OHI?

Before going into detail about the methods and data employed in the current study in Chapter 7, we need to delve into theories regarding language contact in general and in Singapore English (see Chapter 4) as well as previous studies on CSE markers and its substrates (see Chapter 5 and Chapter 6, respectively). The subsequent section Chapter 3 introduces the historical background of Singapore, the people who came to populate this region, and its language-related policies. We will see that language-related policies in Singapore were not decided by one man's power (the late first Prime Minister Lee Kuan Yew 李光耀, 1923–2015) and did not arise suddenly when Singapore became an independent city-state in 1965, but have their roots in the precolonial history, reinforced by later historical events. This following chapter provides a sociohistorical account of Singapore, with several personal recollections chosen from the database of OHI.

3 A sociohistorical account of multilingualism in Singapore

In order to understand the origin of CSE, it is of fundamental importance to cast light on the past of Singapore and the people who came to populate this region. This chapter gives an account of the historical background which elucidates the language-related policy and the ethnic dynamics in Singapore. Besides summarizing historical facts and numbers, this chapter also offers excerpts of the local inhabitants drawn from the Oral History Interviews (OHI) for a more authentic and narrative sketch of the socio-historical picture from the precolonial era (1299–1818) to the modern-day Singapore (1965–).

Section 3.1 introduces the historical background divided in four different eras: (1) the precolonial era (1299–1818), (2) the colonial era (1819–1945), (3) the political change era (1945–1965), and (4) the independence era (1965–). Section 3.2 explores how immigration patterns in these different ages, the different ethnic communities (the Chinese, the Malays, the Indians, and the other ethnic communities), have shaped the linguistic ecology in Singapore. This is followed by Section 3.3, which highlights the ethnic diversity and linguistic complexity, the language-related policies in the past five decades, as well as their impact on language shift, as discussed in Section 3.4, which took place within the same time period in Singapore.

The bottom line of the discussion in this chapter is that the language-related policies, established in Singapore upon its independence as a city-state in the 1960s, did not appear suddenly, but have their historical roots dating back to the precolonial period (Siemund and Li 2020). Historical evidence found in OHI also reveals that Mandarin Chinese played a larger role in the linguistic ecology than many (Ansaldo 2004; Gupta 2001; Leimgruber 2013; Lim et al. 2010) have estimated before the establishments of the bilingual educational and the official language policies in the 1950s. It turns out that Mandarin started to gain significance in the linguistic ecology of Singapore in 1917 (after the Chinese revolution in 1911 in China). We will also see how the emergence and popularity of CSE are at odds with

the promotion of Standard English by the Singaporean government. The emergence of CSE is one consequence of a constant and dynamic language contact between different ethnic groups in Singapore. It serves as a convenient compromise to ease the tension between a pragmatic promotion of the two major languages English and Mandarin, and the ethnic identity and traditional value represented by the heritage dialects. This compromise is also reflected in the use of CSE, as we observe a mixed use of standard and substrate-influenced meanings within one single linguistic sign (e.g. *already*, *also*, *ever*, and *one*). These sociohistorical discussions are also highly relevant for Chapter 7, which will focus more on the linguistic perspectives.

3.1 Historical background

The National Library Board (2007) gives an overview of the history of Singapore by dividing it into five phases (see Figure 3.1), which are (i) classical emporium from 1299 to 1599, which marks the port-settlement of Temasek/Singapura as a major regional emporium in the maritime trade between the Indian Ocean and the South China Sea; (ii) maritime rivalry from 1600 to 1818, when Singapore became a strategic location in the maritime trade of Southeast Asia of internal rivalry among regional and European power; (iii) the colonial port city from 1819 to 1945, which started with the arrival of the British led by Sir Stamford Raffles and closed with the end of World War II; (iv) political change from 1946 to 1965, the period leading to the road of independence; and (v) national and global city from 1965 to present, with the commencement of Singapore as an independent city-state, separated from Malaysia on 9 August 1965.

As we can see in Figure 3.1, this classification is not based on evenly distributed time periods, but based on important milestones along the timeline. The fourth phase: political change just spans over 20 years, but involves the most dramatic change over the history of Singapore. Following Lim (2010), the period of classical emporium from 1299 to 1599 and the time of maritime rivalry from 1600 to 1818 are subsumed under the precolonial era. After giving a short introduction to the pre-colonial period, the following sections will focus on

the history starting from 1819 when Singapore became the colonial port to the independent era.

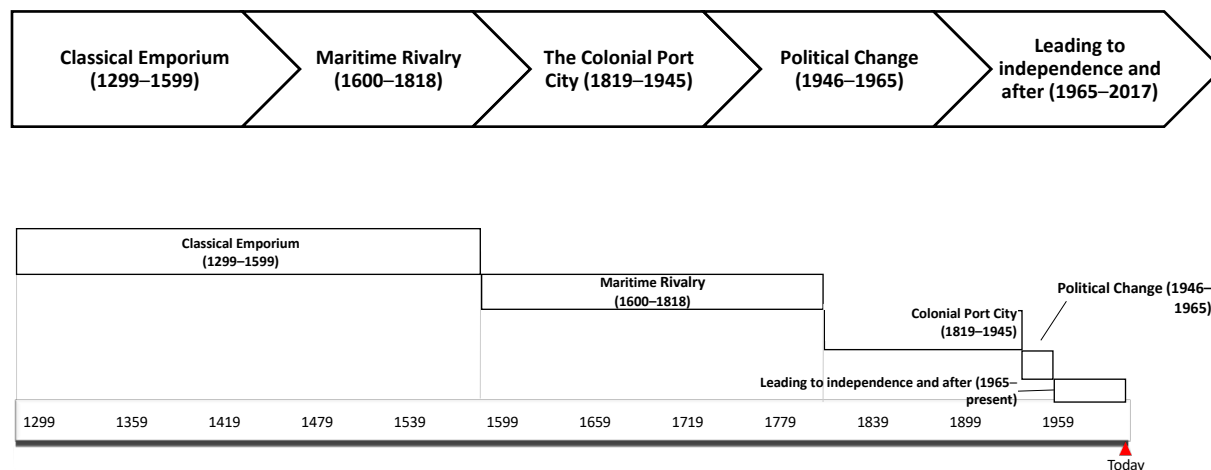


Figure 3.1: Representation of historical timeline of Singapore from 1299 to present
(data from National Library Board 2017)

3.1.1 Precolonial history (before 1819)

It is a misconception that Singapore was just a small sleepy fishing village before the arrival of the British Empire in the 19th century. For example, Schneider (2007) comments on Singapore before the arrival of Sir Thomas Stamford Raffles in 1819 as “little more than a jungle island with potential” (Schneider 2007:153). The island had in fact been a thriving center for maritime trade at least by the 3rd century as a strategic location positioned between the Malacca Straits and the South China sea (National Library Board 2017). This strategic importance of the location of Singapore in pre-colonial history is even evident in a Chinese account from the 3rd century, describing the island as *Pu Luo Chung* (蒲罗中 *púluózhōng*), which is considered as a transcription from the Malay name *Pulau Ujong* ‘island at the end’ (of the Malay Peninsula) (Asian Studies Archive 1994). By the 7th century, when a succession of maritime states arose throughout the Malay Archipelago, Singapore was probably one of the many trading outposts serving as an entrepôt and supply point which

attracted traders, sailors, and pirates from Arab, India, China, Malay and Siamese kingdoms (National Library Board 2017).

As Singapore occupies an excellent strategic location throughout the Malay Archipelago, different powers controlled it in succession, as partly suggested by the change of its name. A 14th-century Javanese chronicle referred to the island as *Temasek* (from Javanese *tumasik* ‘sea town’), which is transliterated by a Chinese account 岛夷志略 (*dǎoyìzhìlüè*, ‘Brief Description of the Barbarians of the Isles’) as 单马锡 (*dānmǎxī*) (see Miksic 2013:169–78). Back then, Singapore was an important island at the center of a thriving Malay kingdom, and appeared on the earliest Chinese maps of the region (Turnbull 1989). The ancient name of Singapore is *Singapura*, which is derived from Sanskrit *simha* ‘lion’ and *puram* ‘city’, lit. lion city, as noted by a 17th-century Malay archive (Turnbull 1996:22). Legend has it that the founding of the city of Singapura took place after a strange, lion-like beast had been sighted there by Prince Sri Tri Buana (also known as Sang Nila Utama), though there is no real historical evidence of this (National Library Board 2017). Singapura was then controlled by a succession of regional empires (the Srivijayathalassocracy, the Javanese Majaphahit Empire, the Thai Kingdom), and Malayan sultanates (Brown 1983:41; Turnbull 1996:4). Around the 15th century however, the Malay kingdom began to fall into decline and Malacca was taken by the Portuguese in 1511 (Wilkinson 1912:71–76). In 1613, the invading Portuguese destroyed a trading post at the mouth of the Temasek River, resulting in the Singapura island being largely abandoned (Turnbull 1996:4). However, the early 19th century saw the agricultural activities returning to Singapura and later in 1818 Singapura was settled by a Malay official of the Johore Sultanate and his followers, who shared the island with several hundred indigenous tribal people and Chinese planters (National Library Board 2017).

There is only limited research studying the demographic situation in the pre-colonial history of Singapore, though some records show that there were approximately 150 inhabitants at the time when Sir Thomas Raffles arrived in 1819 in Singapore (Newbold 1839:279). Among these, about 120 were Malays and the rest were Chinese (Saw 1969:37). The population rose very rapidly through a great influx of immigrants, and by the middle of

1819 the population had risen to about 5,000, as stated in a letter of Raffles to the Duchess of Somerset on 11 June 1819 (Saw 1969:37).

It was not until the mid-1830s that the Chinese had become the most numerous of Singapore's various ethnic groups (Saw 1969:38). They came from Malacca, Penang, Riau, and other parts of the Malay Archipelago (Saw 1969:38). However, before the 1830s, the Malays were the majority. According to Braddell, the first Attorney-General of the British Colony of Singapore, there were 4,724 Malays by the end of 1822 (Turnbull 1996:13). In comparison, the Chinese had a population of around 1,120, which was less than one fourth of that of the Malays during the same period (Turnbull 1996:13). The total population had reached between 10,000 and 12,000 at that time and the first official census of 1824 gave a similar number. According to this, Singapore had 11,000 inhabitants, and among these were 31% Chinese, 60.2% Malays, and 7.1% Indians (Leimgruber 2013:3).

English was not well presented before the British colonial influence, but a Malay pidgin called Pasar Malay (also known as 'Market' or 'Bazaar' Malay) was used as an important common language (Lim 2010:30). English only came into use under British colonial influence when it was fostered as a common tongue between the highly diverse residents of Singapore. Pasar Malay, on the other hand, was a Malay-lexified pidgin that had been used for international trade, politics, and inter-ethnic communication throughout the archipelago for several hundred years (Omar 1977, 1983). It is considered as the pre-colonial counterpart of CSE (Chew 2012:88). A distinctive variety of Pasar Malay is called Baba Malay, which is spoken by the group of Peranakans, a mixed ethnic community also known as the "Straits Chinese" or "Baba-Nyonya", who are the decedents of the Chinese settlers from the southern provinces and the locals in the Malay Archipelago (Lim 2010:30, also see Section 3.2.3 for a more detailed description of this distinctive ethnic group).

3.1.2 The colonial port city (1819–1945)

It is commonly agreed that the history of Modern Singapore began in 1819 when Sir Stamford Raffles arrived at the island and claimed it from the Johore Sultante for the British

East India Company (see Lim 2010; Leimgruber 2009; Schneider 2007). Again, the change of its name proves to be a pertinent tool to analyze the changes of powers in this area during the time frame from 1819 to 1945. Under the British rule, which succeeded in obtaining permission from the local officials to establish a trading post in 1819, the Lion City (Singapura) was renamed to a more Anglicized-sounding Singapore (Tarling 2012:25). The goal of the British was to establish a port city that would overturn the monopoly of the Dutch on trade in the region, so they allowed unrestricted immigration of labor and freedom from taxation upon commerce (Tregonning 1972, cited in Chew 2012:11). The establishment of a free trading port paid off, as it attracted the Babas, the Chetties, the Chulias, Burmese, Indians and Jawi Pekans who brought with them not only their vast experiences of European and Asian practices, but also knowledge of the English from the surrounding sister ports with them (Chew 2012:11–12). The idea of free passage and trade for all also brought a never-seen-before economic boom of the port city, which further attracted an increased number of traditional traders including Chinese, Malays, British Indians, Indonesians, and Arabs, as well as travelers, who arrived to seek their fortune (Sandhu and Mani 1993; Trocki 1979). Each new group of immigrants brought their own languages with them, all of which had an influence on the development of English spoken in Singapore today.

The significance of Singapore as an international port is reflected in its metaphoric names, given by scholarly circles as “powerful fortress” (Ross 1898:6), “New York of the Malay Peninsula” (Tregonning 1972:129) and “Mediterranean of Asia” (Widodo 2009:80). Indeed, in the following century from 1826 to 1941, Singapore was viewed as a vital nexus “belonging to more or less structured places in a networked geographical region” (Chew 2012:12). In 1832, Singapore became the center of government for Malacca, Penang and itself – the three areas of Straits Settlements established under the control of British India in 1826 (Heng and Aljunied 2009). Subsequently, the Straits Settlements became a Crown Colony in 1867. Even more significantly, with the advent of the steamship in the mid-1860s and the opening of the Suez Canal in 1869, Singapore became a major port for ships, loaded with tin, rubber, and crude oil, voyaging between Europe and East Asia (Heng and Aljunied 2009:52; Saw 1969:38). Before the close of the 19th century, Singapore experienced

unprecedented prosperity and trade grew eightfold between 1873 and 1913 (Saw 1969:38). The prosperity attracted more immigrants from areas around the region. By 1931, the population had reached 557,750 (Saw 1969:38).

Singapore was renamed *Syonan* (昭南島 zhāonán dǎo, also spelled as *Shōnan-to*) ‘Light of the South’ in 1942 by the Japanese after it was bombed by Japanese aircrafts on the 8 of December 1941 (LePoer 1991:4). This time, with the change of its name came not only the change of power, but also the sudden end of peace and prosperity for the city. It remained under Japanese occupation for the next three and a half years. For most of the population, the years under Japanese occupation were one of the most suffering periods. One of the reasons is the scarcity of supplies that often accompanies war, but for many it was the result of deliberately inhumane tactics used by the Japanese (Abshire 2011:83). This tragic episode in Singapore’s history still lingers in the memory of many, as can be derived from the data of OHI on the topic of Japanese Occupation.

The following are four excerpts drawn from OHI on the topic of the Japanese Occupation. The interviewee is Soon, Kim Seng (SKS), accession number 000543. He was born in 1924 in Burma, and worked in the Food Control Department, Syonan Municipality during Japanese Occupation. During the interview, we can see that the scarcity of food as well as the fear and anxiety experienced by the Singaporean people were the two central problems of the Singaporean population.

(1) TBL: So you were talking about material life. There were problems. What about mentally under Japanese Occupation

SKS: When a home has got shortage of daily necessities, that itself would cause mental depression. And when say, when a father had been taken away by the Japanese soldier, leaving poor mother to support the children, that itself is an unhappy situation. And if the brothers and sisters were too young to go and work, that is problem for the mother. So the mother would have to find means and ways to keep her family going from day to day. And imagine more than three years of Japanese Occupation.

(2) SKS: How people in those days solved their individual family problems? I can describe to you one of the scene[s] like an elderly old lady sweeping on the road. What was she sweeping? She was sweeping rice fallen from the moving lorries. This is a very rare sight. Before the war you don’t find people sweeping food from the road. And that itself would give you an idea of how difficult food problem was at that time.

- (3) Well, when you talked about happiness, probably after each meal is solved, probably there is a little bit of cheer because your stomach has filled up for that particular moment. And eventually when the time come for the next meal, well, you might be worrying again.

(OHI-000543-SKS)

Excerpts (1) (2) and (3) all illustrate the severe lack of food supply as a prevalent issue during Japanese Occupation. After a discussion on the scarcity of supplies, i.e. lack of charcoal and electricity supply in the family, the interviewer TBL directed the conversation towards how these problems about material life extended to the mental life of the people under Japanese Occupation. The interviewee SKS replied that what accompanied people besides their hunger for food was the fear and worry about the next day. Excerpt (4), on the other hand, expresses the fear towards the deliberately inhumane tactics of the Japanese soldiers.

- (4) In those days of early Japanese Occupation, people do not wish to go about freely because for fear of Japanese sentries, drunken Japanese soldiers. Especially women and young girls, they would definitely not dare to venture and go about alone unnecessarily. The Japanese sentry would bully people. When you passed by and if you don't bow they would call you up and they would slap you. Or they would even make you kneel down beside them for, say, half an hour or so, just simply to get some enjoyment out of it. So there are so many such cases happening all over Syonan-to at that time.

(OHI-000543-SKS)

It is perhaps not surprising that the Japanese language had barely any influences on the linguistic ecology of Singapore,⁴ though the Japanese considered the learning of the Japanese language as vital in inculcating the Nippon Spirit and culture among the people of Singapore (National Museum of Singapore 2019:3). The Japanese language was made compulsory for students in schools and adults who worked in the government agencies or Japanese corporations. Efforts were also made in publishing Chinese-Japanese textbooks, Japanese song books and Malay-Japanese dictionaries (National Museum of Singapore 2019:9–10). Most of these books taught basic level Japanese with the goal to make the Singaporean locals pick up basic Japanese in the shortest possible time. However, as evident

⁴ There are traces of Japanese in Hakka and Kristang (“Christian”) Creole though. Some contemporary Hakka usages in Taiwan have come from Japanese, along with Hokkien, and recently Mandarin Chinese (Chew 2012). Kristang has been influenced by Japanese along with Bengali, Ceylonese, and Timorese to a small extent, though the major influences have come from Malay.

in OHI, the majority of the Singaporeans did not make an effort to learn Japanese (see Interviews 000001 TMK, 000025 KRM, 000263 S, 000310 SGB, 000374 GWG).

3.1.3 Political change (1945–1965)

As demonstrated in Figure 3.1, the phase of Political Change covers the shortest timespan in the historical timeline of Singapore but encompasses the most dramatic political and demographic changes.

After World War II, the British forces returned in September 1945 and Singapore came under the British Military Administration. When the period of military administration ended in March 1946, the Straits Settlements were dissolved. On 1 April 1946, Singapore became a Crown Colony again, while Penang and Malacca became part of the Malayan Union in 1946, and later the Federation of Malaya in 1948 (LePoer 1991:44).

Post-war Singapore witnessed an acceleration in the growth rate of its population. The annual rate, at 4.5 percent during 1947–1957, was the highest growth rate since the 1840s (see Saw 1969:39). There are multiple reasons for this (i.e. rapid decline of mortality, high level of fertility, immigrants from the hinterland of Malaya), but the most important factor accounting for the rapid population growth was the vast number of immigrants streaming into the country, the main streams of which were from China, India, and the Netherlands East Indies (Saw 1969:40). Many arrived in Singapore with nothing more than the clothes on their backs and with the idea to return to their mother country after acquiring a fortune (Saw 1969:40). In the course of time, however, an increasing number of people built families. Thus, they remained permanently in Singapore, though throughout the 19th century transient population was a distinct characteristic of Singapore's population constellation (Saw 1969:40).

When the People's Action Party (PAP, formed in 1954 by Lee Kuan Yew) won the national election held on 30 May 1959, Singapore gained self-government with autonomy in all state matters except defense and foreign affairs, and Lee Kuan Yew became Singapore's first Prime Minister (LePoer 1991:51). In 1963, Singapore joined with the Federation of

Malaya, Sabah (North Borneo), and Sarawak in the Federation of Malaysia (LePoer 1991:54). The idea was to have central government responsibility for defense, foreign affairs, and internal security, but local autonomy in matters pertaining to education and labor (LePoer 1991:47). However, the merger proved to be short-lived. Singapore was separated from the rest of Malaysia on 9 August 1965, and became a sovereign, democratic, and independent nation (IBP 2015:31).

3.1.4 The independence era (1965–present)

Independent Singapore was admitted to the United Nations on 21 September 1965 and became a member of the Commonwealth of Nations on 15 October 1965 (IBP 2015:31). On 22 December 1965, it became a republic, with Yusof bin Ishak as the republic's first President and Lee Kuan Yew as Prime Minister (IBP 2015:32). Thereafter began Singapore's struggle to survive and prosper on its own, as Lee (2000) wrote in his autobiography:

All of a sudden, on 9 August 1965, we were out on our own as an independent nation. We had been asked to leave Malaysia and go our own way with no signposts to our next destination. We faced tremendous odds with an improbable chance of survival. Singapore was not a natural country but man-made, a trading post the British had developed into a nodal point in their worldwide maritime empire. We inherited the island without its hinterland, a heart without a body. (Lee 2000:3)

Singapore's strategy for survival and development was essentially to take advantage of its strategic location and the favorable world economy. The focus of the PAP's efforts was to make use of export-oriented manufacturing for economic development. Along with Taiwan, Hong Kong, and South Korea, Singapore became part of the elite groups of the regions called the Four Asian Tigers (also Four Asian Dragons, 亚洲四小龙 yàzhōu sì xiǎolóng, literally translated as 'four little dragons in Asia'). They all maintained exceptionally high growth rates (in excess of 7–8 percent a year) between the mid-1950s and early 1990s (Abshire 2011:133). Across the decades, the government shifted strategies, from export-oriented manufacturing to service-oriented and from value-adding to value-creation to stay with, or

ahead of the economic trendsetters (Abshire 2011:134; Tay and Soh 2015). The change of strategies has paid off: with strong footholds in manufacturing, especially in medical science, energy production, information technology, sustainable resource development, and international culture and the arts, Singapore is creating a vibrantly diverse center for innovation, and nurturing workforce skills and talent that are crucial to success in today's integrated global economy (Edb 2015).

The political leadership in Singapore has long argued that they adopt a “pragmatic” approach to government and denies the influence of ideology on its political decisions (Khong 1995:123–24). Ow Chin Hock, chairman of the feedback unit, argues, “The government is not shackled by any particular ideology or dogma” (Tan and Tan 2012:508). It is difficult to believe though, that there are no beliefs governing the actions of the political leadership of the city-state. Many researchers in East and Southeast Asian studies have argued that it is Confucianism that formed the central part of Lee Kuan Yew's political thought in the 1960s, with some elements tracing back to the 1950s (see Barr 2000; Tan 1989). This is not only reflected in the dominant role that the Singaporean government has played in its economic success (see Tan 1989:9),⁵ but also in shaping the cultural values of its citizens (Barr 2000:310). One of the examples is the “Asian values” argument. While the Singaporean government appreciated the western value that brings Singapore scientific and technological achievements, they were also afraid of the ideas and practices of excessive individualism that would be inimical to a disciplined, compliant, culturally conservative and racially sensitive society (Khong 1995:413; Tan 1989:14). The “Asian values”, according to Khong, is the alleged value of Chinese culture, which “bear an uncanny resemblance to Weber's Protestant work ethic” (Khong 1995:413). In the course of this debate, Confucianism has been singled out as exemplifying the Asian values, such as hard work, thrift, close family ties, filial devotion to parents and proper education for children as well

⁵ Tan pointed out that “Confucian belief emphasizes that the government is primarily for the good of the people and not the rulers. [...] The virtuous leader will be benevolent to his people and will win the mandate of heaven to rule. The people have an important role to play as subjects by responding to the good policies of the ruler” (1989:9).

as extension of peace and harmony for family to the community, society, and the world at large (Tan 1989:7–16).

Steps were taken to promote these Asian values, such as (i) the introduction of Confucian Ethics as one of six options in the “religious knowledge program” in secondary schools in 1982 (Tan 1989:5) and (ii) the establishment of the Institute of East Asian Philosophies within the National University of Singapore in 1983, which is a resource and research center of oriental philosophies (Tan 1989:5). In 1991, the government formulated and formalized the national ideology named “Shared Values” in a white paper which in many respects was a reiteration of the earlier Confucian values campaigns (Tan 2012). The Shared Values consist of the following five broad values (Tan 2012:450): (i) nation before community and society above self; (ii) family as the basic unit of society; (iii) community support and respect for the individual; (iv) consensus, not conflict; and (v) racial and religious harmony.

The goal of such actions is to hold the society together despite enormous differences in culture and language among different ethnic groups. One further step for creating a sense of national identity and consciousness among a disparate population of immigrants involves the language policies, including (i) the promotion of bilingualism, i.e. by adopting English as a neutral language for inter-ethnic communication and giving equal official status to Mandarin Chinese, Malay, and Tamil; (ii) the introduction of the Speak Mandarin Campaign. These language-related policies will be further discussed in Section 3.3.

3.2 Population, immigration, and languages

In language contact, the number of speakers in the respective linguistic groups, the relative social status of the groups involved as well as the relative prestige of the languages have been observed as the determinants of contact-induced change (Siemund 2008:4; Thomason and Kaufman 1991). In this section on population, immigration and languages, we will look at the population composition of different ethnic groups, their origins, as well as their languages. These social parameters will subsequently be compared with the top-down

language policies implemented by the Singaporean government, which, to a large extent, have resulted in a language shift in Singapore.

3.2.1 Overview of demographic trends

One salient characteristic of Singapore is the multi-ethnicity of the population, which are composed of three major groups besides other minorities that already existed in the early history of Singapore: the Chinese, the Malays, and the Indians. As mentioned earlier, the establishment of Singapore as a British trading post by Raffles in 1819 meant a rapidly expanding economy, which was coupled with a liberal, open-door immigration policy (Yeoh and Lin 2012). As a result, the population soared from around 150 inhabitants to more than half a million by the 1931 census (Saw 1969:37),⁶ before the Japanese Occupation (see Table 3.1). After World War II, the population continued to grow and tripled towards the end of the period of political change. Because of the stability of the society and immigration policies to attract foreign talents, by the 2019 census, the population had reached 5.7 million (Wong 2019:3), which is 10 times the population of 1931.

It was not until April 1871 that the first systematic census of Singapore was conducted. Successive censuses were carried out once every 10 years until 1931 when the census was interrupted due to World War II. The census resumed in 1947 and thereafter, censuses have been carried out at regular 10-year intervals (see Table 3.1). From 2005 on, there is also an annual series of population trends, which covers different aspects of demographic statistics in a single volume (Wong 2019).

⁶ According to Rahim (2010:4), there were 1,000 inhabitants.

	Census	Chinese	Malay (percent)	Indian	Others	Total pop. (x1 000)
	1824	31	60.2	7.1	1.7	10.68
	1830	39.4	45.9	11.5	3.2	16.63
	1836	45.9	41.7	9.9	2.6	29.98
	1840	50	37.3	9.5	3.1	35.39
	1849	52.9	32.2	11.9	3	52.89
Colonial Port (1819–1945)	1860	61.2	19.8	15.9	3.1	81.73
	1871	56.2	26.9	11.8	5	97.11
	1891	67.1	19.7	8.8	4.3	181.60
	1901	72.1	15.8	7.8	4.3	226.84
	1911	72.4	13.8	9.2	4.7	303.32
	1921	75.3	12.8	7.7	4.2	418.36
	1931	75.1	11.7	9.1	4.2	557.75
Political change (1946–1965)	1947	77.8	12.1	7.4	2.8	938.11
	1957	75.4	13.6	8.6	2.4	1445.93
The independence era (1965–)	1970	76.2	15	7	1.8	2074.51
	1980	76.9	14.6	6.4	2.1	2413.95
	1990	77.8	14	7.1	1.1	3047.10
	2000	76.8	13.9	7.9	1.4	4027.90
	2010	74.1	13.4	9.2	3.3	4401.40
	2015	74.3	13.3	9.1	3.2	5535.00
	2016	74.3	13.4	9.1	3.2	5607.28
2019	74.4	13.4	9.0	3.2	5703.60	

Table 3.1: Population and ethnic composition of Singapore from 1824 to 2019 (Saw 1969; Cheng 1985; Singstat 1990, 2000; Wong 2010, 2015, 2016, 2019)⁷

The ethnic composition has remained relatively stable since the 1900s. Towards 1901, the Chinese accounted for 72.1 percent of the total population, the Malays and Indians 15.8 and 7.8 percent respectively, and others made up 4.3 percent (Saw 1969:42). The latest data in 2019 shows a similar picture concerning the ethnic composition among Singapore residents: Chinese 74.4 percent, Malays 13.4 percent, Indians 9.0 percent and others 3.2 percent (Wong 2019). The following sections explore the background of each of these communities, which consist of further heterogeneous subgroups, with an emphasis on their diverse language profiles.

⁷ From 1990 on, the percentages of population by different ethnic groups only involve Singapore residents.

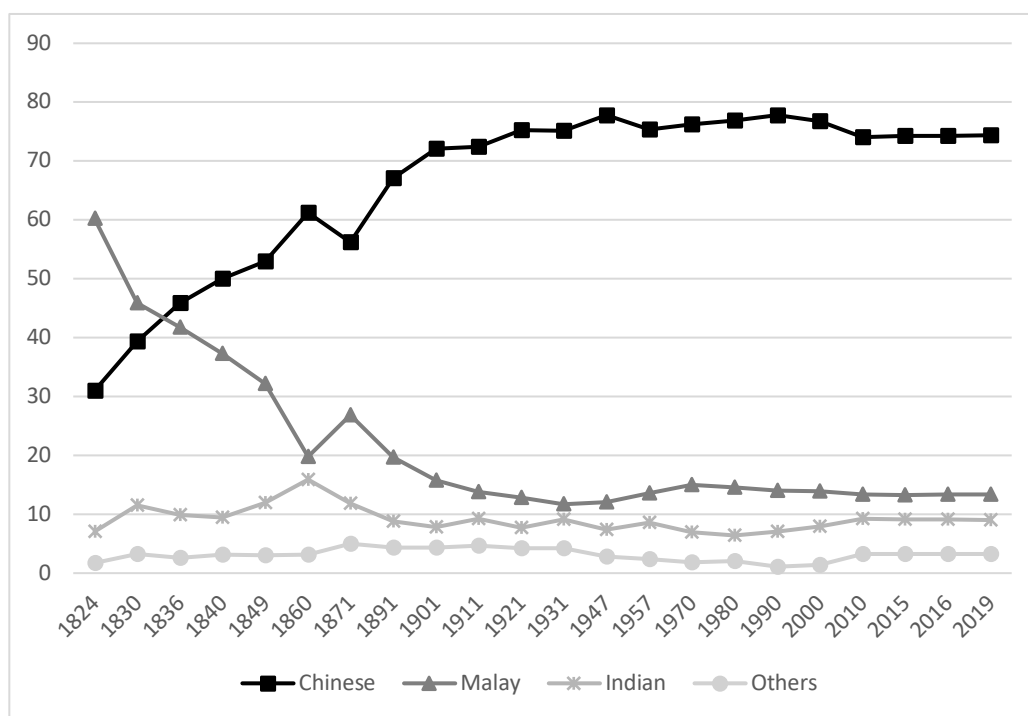


Figure 3.2: Ethnic composition of Singapore Population from 1824 to 2019 (Saw 1969; Cheng 1985; Singstat 1990, 2000; Wong 2010, 2015, 2019)

3.2.2 The Chinese communities and their languages

The growth of the Chinese communities was fueled by immigration that started soon after Raffles landed in Singapore in 1819. Turnbull (1989:5) estimated that Singapore had 30 Chinese and 120 Malays when Raffles arrived in Singapore in January 1819. By 1821, the Chinese population in Singapore was estimated to have increased to 1,150 (Saw 2012:8). The first population census was conducted in 1824 and recorded 3,317 Chinese, which constituted 31% of the total population 10,683 (Buckley 1984:154). The Chinese surpassed the Malay population in the 1830s and became the largest ethnic proportion of Singapore’s total population (see Figure 3.2). By the 1930s, the Chinese grew to make up 75% of the total population, and that proportion has remained more or less unchanged to this day. It must be noted the gender ratio remained highly skewed in favor of males until well into the 20th century (Crawford 2000:378–84).

The ancestral origins of the majority of the Chinese ethnic group in Singapore can be traced back to migrants from coastal southeastern China, mainly from today’s Fujian

Province (福建 fújiàn, formerly romanized as Foken, Fouken, Fukien, and Hokkien) and Guangdong Province (广东 guǎngdōng, romanized as Canton),⁸ areas of much linguistic and subcultural variation (see LePoer 1991:80; Lim 2010:23). In the southern regions of China, the southern Chinese dialects Hakka, Min, and Yue are spoken (see Figure 3.3).

To better understand dialects in China, it is helpful to draw a diagonal line from the Northeast around the capital Beijing down towards Southwestern China. The area to the North and West represents primarily the Mandarin-speaking area. The space of South and East, on the other hand, holds most of the Chinese dialects. These include provinces in central China, such as Sichuan (四川 sìchuān) and Hunan (湖南 húnán), those to the south such as Guangdong (广东 guǎngdōng) and Hainan (海南 hǎinán), as well as the Southeast, such as Fujian (福建 fújiàn), Jiangxi (江西 jiāngxī), Zhejiang (浙江 zhèjiāng), and Jiangsu (江苏 jiāngsū) (Gong, Chow, and Ahlstrom 2011:224). As shown in Figure 3.3, from North to South, the main southeastern dialects are Wu (吴 wú, spoken in the Shanghai region), Gan (赣 gàn, spoken in Jiangxi), Xiang (湘 xiāng, spoken in Hunan), Min (闽 mǐn, including the Hokkien variety spoken in Fujian and Hainan islands, and the Teochew variety spoken in the Chaoshan region of Guangdong), Hakka (客家 kèjiā, spoken primarily in southern China and Taiwan), and Yue (粤 yuè, Cantonese, spoken in Guangdong) (Gong et al. 2011:225).

⁸ In English, the term Canton is ambiguous. Traditionally it refers to the capital city Guangzhou of the Guangdong Province (LePoer 1991:80), but it also refers to the Guangdong Province (Tan 1990:9; Lim 2010:23).



Figure 3.3: Map of Chinese dialects in southern and eastern China (adapted from Gong et al. 2011:226)

The largest group among the Chinese communities in Singapore is the Hokkien, who surpassed the Teochew population in the early 19th century (LePoer 1991:79). In 1980, the Hokkien made up 43 percent of Singapore's Chinese population (LePoer 1991:80). They came from the area around the trading port of Xiamen (厦门 xiàmén, romanized as *Amoy* under the pronunciation of Hokkien) in southern Fujian Province. Even before the foundation of Singapore in 1819, the Hokkien traders and merchants had been active in Southeast Asia for centuries (LePoer 1991:80). They are said to be the most powerful community in the economy in Singapore and, therefore, the Hokkien dialect, mutually intelligible with Teochew, was the most frequently understood and spoken Chinese dialect in this period (Lim 2010:24).

Singapore Group	Alternate Names
Hokkien	Amoy, Fujian, Fukien, Xiamen, Hsia-men
Teochew	Chaozhou, Chaochou, Teochiu, Chao-chou, Shantou, Swatou, Santow
Cantonese	Canton, Guangzhou, Kuangchou
Hainanese	Hainanese, Hailam, Qiongzhou, Chiungchou
Hakka	Kejia, Kechia
Foochow	Fuzhou, Hokchiu
Henghua	Xinghua, Hsinghua
Hokchia	Fuqing, Fuching
Shanghainese	Sam Kiang, Sanjiang, Sanchiang

Table 3.2: Singapore Chinese speech groups and their alternate names (data based on Cheng 1985:15–23; W.K. Wong 2015)

The Guangdong Province, located in southern China on the South China Sea coast, harbors a heterogenous group, being the ancestral home not only of the Cantonese,⁹ but also the Teochews and Hakkas. The Teochew (sometime written *Teochiu*) was the largest group of the Chinese population in Singapore until the 1930s when they were outnumbered by the Hokkiens (LePoer 1991; Wong 2011). Today, they comprise 19.8 percent of the Chinese population, according to the latest General Household Survey (GHS) in 2015 (W. K. Wong 2015). Their home area is Chaozhou (潮州 *cháozhōu*, also written as Chiuchow, Chaochou) and Shantou (汕头 *shàntóu*, also Swatou and Santow). Chaozhou and Shantou are immediately south of the Hokkien-speaking area of Fujian, and, as mentioned earlier, both Teochew and Hokkien are closely related dialects of Min.

⁹ Like Canton, the term Cantonese is ambiguous. Here it adopts its traditional meaning and only refers to the group of people originated from the capital city Guangzhou, who speak Cantonese instead of other dialects (e.g. Teochew, Hakka) spoken in Guangdong Province, though it is used more often to refer to anyone coming from Guangdong Province today. The abbreviation of the Province is also 粵 *yuè*. The dialect Cantonese, however, is restricted to refer to the dialect that derives from Guangzhou, and its standard form (accent) is most predominantly used by local residents in Guangzhou, Hong Kong, and Macau. Residents around the three cities generally use Cantonese too, but in a slightly or vastly different accent varying from region to region.

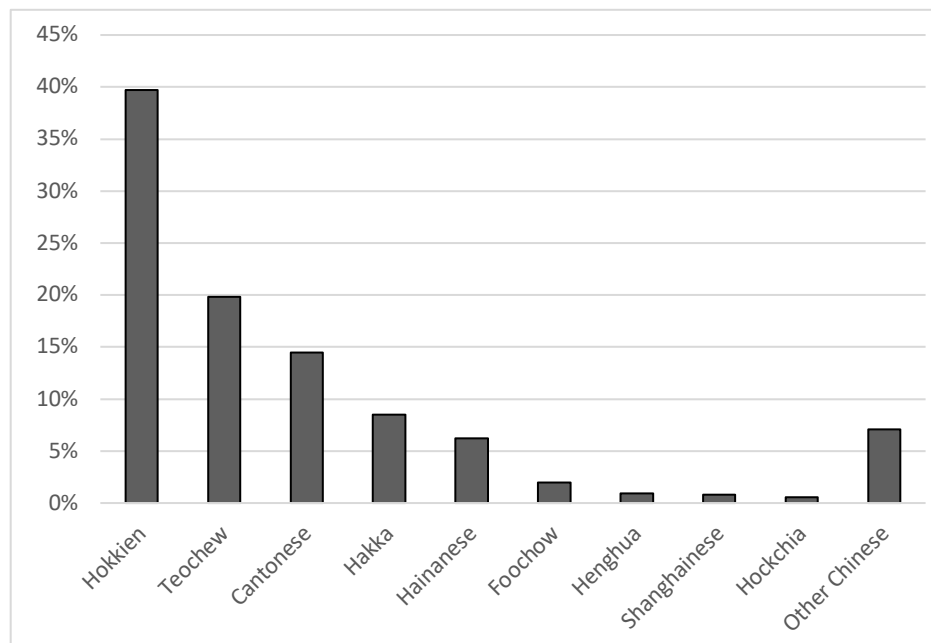


Figure 3.4: Chinese resident population by dialect group 2015 (Wong 2015)

The third most numerous group was, and still is the Cantonese group which makes up around 15 percent of the Chinese population today, according to the 2015 General Household Survey (LePoer 1991; W. K. Wong 2015). They are from the lowland of central Guangdong Province around the historically important port city Guangzhou (Canton), now the capital city of Guangdong Province. A group of the Macau people are also called the Cantonese, in reference to their overseas travels from the port of Macau before the opening of Hong Kong in 1842 (Clayton 2009:110–13). The first recorded instance of the arrival of a group of Cantonese to Singapore was in 1821 (Tan 1990:11–12).

There were sizable numbers of the Hakkas, who are a group of people scattered through the interior hills of southern China and generally considered migrants from northern China (Lim 2010:23). They are a subgroup of the Han Chinese people. However, unlike other Han Chinese groups, the Hakkas are not named after any geographical region. They are usually identified with people who speak the Hakka language or share at least some Hakka ancestry. The term Hakka (in Mandarin Chinese 客家 kèjiā) literally means ‘guest families’ or ‘strangers’. They are often compared with the Jews in the Western literature (see Erbaugh 1992 on the secret history of the Hakkas: The Chinese revolution as a Hakka enterprise). Though the particular dates and stages of migration have been a topic of debate

(Cohen 1968; Constable 1996; Kiang 1991), it is well accepted that the main body of those who later became the Hakka migrated from north central China around the Yellow River area, probably as early as the Qin dynasty (221 BC–207 BC) to escape from social unrest, upheaval, and invasions (Constable 1996:9). The first Prime Minister of Singapore Lee Kuan Yew himself was of Hakka and Chinese Peranakan descent (Martin 2015). Following Hakka, Hainanese ranks the fifth largest Chinese dialect group (6.24%) of the Chinese resident population (Wong 2015). For more on the Hainanese dialect group in Singapore, interested readers are referred to Lim (2010) and Tan (1990).

Additionally, there is a small number of the so-called “Three Rivers People” from northern and central China (Lim 2010:10). They are people who came from the provinces north of Guangdong and Fujian, more specifically, those provinces sharing the word “river” (江 jiāng) in their names – Jiangxi, Jiangsu, and Zhejiang. They would have already spoken one or some of the southern Mandarin dialects or the Wu dialect of Shanghai, Ningbo, and Hangzhou (LePoer 1991:81). There were not many of them until the 1970s, and they made up about only 1.7 percent of the Chinese population by 1980 (LePoer 1991:81; Lock 1982:302). They are represented by the Shanghainese dialect group in Singapore today, which comprises 0.77 of the Chinese resident population (Wong 2015).

People from the coastal counties of northern Fujian speak dialects that are quite distinct from those of southern Fujian. They are called Foochow (also Hokchiu), Henghua, and Hokchia, which make up 1.96 percent, 0.9 percent, and 0.58 percent of the Chinese resident population respectively (Wong 2015).

3.2.3 The Peranakan Chinese

Besides Mainland Chinese, there was a distinct and important group of Chinese from Malacca and Penang (Lim 2010:24). They are the descendants of the south Chinese immigrants to Malacca and Penang in the 18th and 19th century, mostly male immigrants, who often married local Malay or Indonesian women. They are often referred to as the Straits-born Chinese, or the Peranakan Chinese. This community often address themselves

as “Baba Nyonya”, “Baba” is the term for males and “Nyonya” is the term for females. According to Lim, although they only formed 9.5% of the Chinese population, they are considered the elite group and enjoy a high socio-economic status (2010:24). They are highly multilingual, commanding English at a high level while also speaking Malay and Chinese, and have good access to both education and trade (Lim 2010:25; Nathan 1922:77).

3.2.4 Multilingualism among the Chinese communities

The metadata in OHI also reveals that the Chinese communities consist of highly multilingual groups. Besides their mother tongue, which can be traced back to their ethnicity, they usually speak a couple of other Chinese dialects, specifically Cantonese and Hokkien. Also, they can speak Malay to some degree. On the other hand, back in the 1910s to 1960s, Malays and Indians could also speak fluent Cantonese and Hokkien. In the following excerpt from an interviewee with Chinese background (Hainanese speaker, accession number 002108), we can see how highly multilingual the Chinese communities as well as the Malay and Indian communities were, and how intense and complex the language contact situation was as they learned languages from each other:

So, I think one of the most beautiful thing[s] is because when you are a child, you don't really pay much regard to political tensions or colour or race. To us, everyone is the same. And the beautiful part is, you can find Malays who could speak very good Hokkien and they'll find Chinese who can speak very good Malay; Indians who are super fluent in Cantonese and Hokkien and then you wonder why is it this person dark and why are you not dark. But it doesn't matter, you see, the point is we are in the same kampong, I mean in the same precinct. So we'll go out and have fun together.

(OHI-002108-TSG)

Such a multilingual situation may not always be embraced by all individuals. Tension seems unavoidable between the promotion of one major language over the other due to pragmatic reasons (e.g. the majority of a population choose one language as the medium of instruction in a local school) and the ethnic identity represented by the heritage dialects. Here is an excerpt from a Hokkien speaker, giving us a narrative of an individual dialectal shift

(Hokkien – Cantonese – Mandarin) that he experienced from the time of kindergarten to high school.

So the interesting thing of course is that the students are mainly Cantonese. Medium of instruction mainly Cantonese. I'm a Hokkien. So it's a chicken going to a duck country, don't know what was going on. So initially, for this guy who has got all the attention at home, really it was not quite used to. So there was a lot of reluctance in attending school, missing classes and my parents, I suppose love me too much, allowed me to do it. So year one was... the greatest achievement is [I] managed to learn Cantonese. Children of that age really learn language very fast. I've got a few experiences with switching languages. The first time was of course, in Cantonese—from Hokkien speaking to Cantonese. That was in kindergarten. Now because I missed so many of the K1 classes, my parents were in some dilemma as to whether to allow me to repeat K1 or go to K2. You know, it was as bad as that. So apparently, they decided to let me stay in K1 one more year. So I'm like a slow developer. But the mistake of not going to K2 was that they started to learn Mandarin in K2. So I [was] stuck with Cantonese in K1. So when I went to Catholic High and everybody speaks in Mandarin, so again, I'm lost. This time it's a duck going to a goose country.

(OHI-003223-CSS)

In order to better understand this excerpt, we need to know the historical background of the shift of the medium of instruction in school in Singapore. Before the 1910s, Chinese schools designed their syllabi and chose their separate mother tongues as the medium of instruction (Chew 2012:29). Many Chinese schools were founded by the different dialect communities, with some well-known ones such as the Cantonese Yeung Ching School (养正学堂 yǎngzhèng xuétáng) founded in 1905, the Hakka's Yingxin School (应新学校 yīngxīn xuéxiào) and Kee Fatt school (启发学堂 qǐfā xuétáng) in 1906, and the Hainanese Yoke Eng School (育英学堂 yùyīng xuétáng) in 1910 (Peng 2012:448). However, from 1917, the medium of instruction began to shift to Mandarin after the success of the National Language Movement in China. Textbooks which were written in vernacular Mandarin gradually replaced regional language textbooks which were largely comprised of Chinese classics (Chew 2012:153).

The excerpt of interviewee CSS 003223 gives a narrative of what happened in this interesting historical setting, in which he experienced the shift of the medium of instruction from Cantonese to Mandarin. The first time when he attended kindergarten, he was required

to learn Cantonese despite being a Hokkien speaker, as the medium of instruction in kindergarten was Cantonese. The second language shift happened when he went to high school when Mandarin became the medium of instruction. There are two interesting metaphors he gave to describe how he felt with the switch of languages – “it is a chicken going to a duck country” and “it is a duck going to a goose country”. This comes from a well-known Cantonese idiom “Chicken talking with Duck” or what is better known as “Chicken and Duck Talk” in English (鸡同鸭讲 Mandarin jītóng yājiǎng, Cantonese gail tung4 aap3 gong2). It refers to a situation where two people talk in different languages (or at different wavelengths) and are not able to understand each other at all. The process may lead to misunderstanding and animosity. The second expression “it is a duck going to a goose country” is believed to be created by the interviewee himself based on an analogy of the idiom. Both sentences express the tension and anxiety he struggled with at the two times of language switch in school.

3.2.5 The Malay communities and their languages

The Malays have inhabited the area that is now Singapore since as early as the 13th century. Many of the indigenous Malays were living on the island under the Johor Sultanate before the arrival of Raffles, most of them came from the Malay Archipelago. These early immigrants were from the Malacca, Sumatra, and the Riau Islands, and later also from Java and Bawean Island, as well as Sulawesi and other islands (Lim 2010:23). Most of the Malays in Singapore today have their roots in Indonesia and Malaysia.

The Malays used to be the largest ethnic group (60% in 1824) in Singapore until the 1830s, their numerical dominance was swiftly surpassed by the Chinese (see Figure 3.2). According to the 2015 population trends, the Malay residents constitute 13.4% of the resident population (Wong 2016:5), making them the second largest ethnic group in Singapore.

Lim mentions that compared to the Chinese-speaking community, the Malay community is a fairly homogeneous group (2010:23), though this is only partly true. The

Malay category contains its own numerous minorities, including Orang Laut and Orang Seletar (originally boat- or shore-dwelling semi-nomads), groups of Indonesian origin such as the Minangkabau, Batak, Peninsular Malays and peoples of Borneo origin as well as Javanese, Bugis and Boyanese, who spoke other Austronesian languages such as Javanese, Buginese, Boyanese (Lim 2010:23). There was also the English-speaking Jawi-Peranakans and a small but economically important group of Arabs (Bloom 1986:353). Arabic is more common among the Muslim religious teachers and is the preferred language learned by the more religious Malay Muslims, who want to keep themselves connected to their religion and history (Aljamal 2015).

The Malay language (also called *Bahasa Melayu*) has been recognized as Singapore's national language since independence and has been mostly used within the Malay speech community. Virtually the entire Malay population speaks Malay and the status of the language is not threatened despite their minority status, as Malay is not only one of the official languages in Singapore but also the national language of Malaysia and Indonesia (Deterding 2007:4). Besides Malay, a Malay-lexified pidgin Pasar Malay (also known as 'Market' or 'Bazaar' Malay) mentioned earlier has served as the *lingua franca* in the region for centuries. It was also widely used in the East Indian Archipelago and was the basis of the colonial language used in Indonesia by the Dutch (see Bao and Aye 2010 on Bazaar Malay topics). The version of Bazaar Malay used in Chinese merchant communities and the Peranakans is called Baba Malay (Ansaldo, Lim, and Mufwene 2007). In the early 1970s, Bazaar Malay was still one of the most understood languages, only second to Hokkien, and widely used for inter-ethnic communication, with all Indians and 45% of the Chinese claiming to understand it (Lim 2010:27). However, its significance as a *lingua franca* gradually declined starting from the independent era of Singapore, except in the older generation and the lower social strata (Lim 2010:28).

3.2.6 The Indian communities and their languages

The term “Indian”, though convenient, is hardly accurate for the Indian ethnic group in Singapore. The Indian population, represented by 9.0 percent of the total Singaporean in Singapore (Wong 2019), is a heterogeneous group and differs in background, religion, and language. The only thing members in this category have in common are their origins from the British colonial empire which now make up Sri Lanka, Pakistan, Nepal, India, and Bangladesh according to the report of the World Directory of Minorities and Indigenous Peoples in 2017 (MRG 2017). Though Indian contact had been occurring since ancient times, it is considered that they exerted little influence on language contact in Singapore until the 19th century when the Straits Settlements were considered to form part of British India and, thus, Indian laborers and migrants were brought in large number to occupy various functions (Lim 2010:25). For example, Tamils tended to be employed on rubber plantations, though Ceylonese Tamils tended to work as clerks, civil servants in government departments and the Tamil Muslims, Sindhis, and Gujaratis were often small traders; Punjabi Sikhs would work in the army or police or as private security guards (Lim 2010:25). The population of the Indians peaked in the mid-1800s at 16 percent, but otherwise ranging from 7 percent to 9 percent from the 1900s onwards (see Figure 3.2).

3.2.7 Multilingualism among the Indian communities

In cultural and linguistic terms, the majority (over 54 percent) (Wong 2011:49) of Indians are of Tamil ancestry, which explains why Tamil is one of the country’s four official languages. In addition to Tamil, other dialect groups include Malayalee (7.48 percent), Hindi (3.71 percent), Sikh and Punjabi (5.26 percent),¹⁰ Gujarati (1.24 percent), and Sindhi (1.10 percent) among others (see Figure 3.5, data from Wong 2019). The various languages of the Indian ethnic group – Tamil, Punjabi, Malayalam, Hindi, Gujarati, Telugu, Bengali, and so

¹⁰ The 2015 population trends captured Sikh as a separate dialect category though Sikhs are a religious rather than linguistic group and virtually all of the Sikhs are Punjabis (see Wong 2015, Lim 2010).

on – have only been used in home domains of the respective communities (Lim 2010:26, 29).

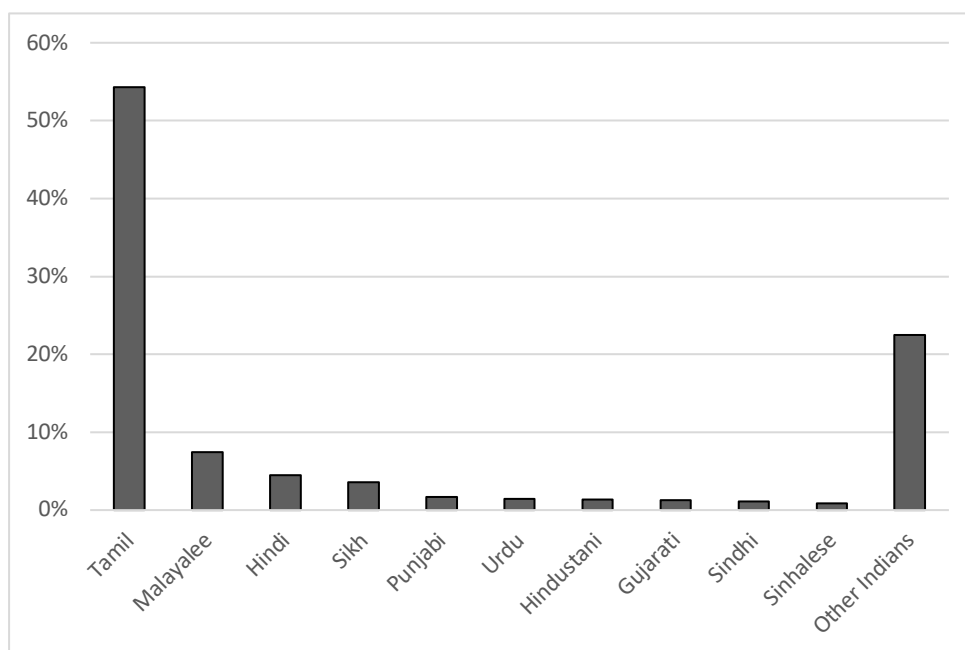


Figure 3.5: Indian resident population by dialect group (data from Wong 2015)

However, unlike the Chinese in Singapore who chose their mother tongue as the medium of instruction, many Indians strived for English education (Chew 2012:48). Even in the independent era with Tamil given the official language status, Tamil was not perceived as a prestige language with little functional use in the public domain (Saravanan 1993; Schiffman 2003). Children preferred to attend English-medium schools to learn Malay as a second language (Chew 2012:48). Even in the home domain, Tamil has been replaced by English in the community in inter-ethnic communication (Lim 2010:31). As exemplified by the following excerpt from Interviewee MS 001663, a shift of language repertoire happened across two generations, with the older one speaking a variety of Indian vernacular and Malay, while the younger one preferred English and Malay instead of the Indian vernaculars.

MS: My mother speaks Tamil, she speaks Hindi, Urdu and she can speak English and Malay. And my father also; my father speaks Tamil, Urdu, Hindi – Hindi and Urdu the same and of course when [he's] here, he learned Malay, he could speak and English. You see my parents were very good, they spoke all that languages.

And so ... But if you come from India, if you know Tamil, you know Hindi. So my parents spoke both the languages. So at home we speak all these languages. But we the younger ones, later, later on we never like to speak the other language, you know, because we go to school, we speak English and all that. So we speak English a lot amongst ourselves. And we read and then we speak Malay to the servants. And so we never got on to the mother tongue so much because it's so terrible. Like Tamil, you got to twist your tongue.

RZ: So you use more English and Malay?

MS: We use more English and Malay, the younger ones. But the older ones, parents all speak in Malay to the servants and they all use the mother tongue. My father speaks Tamil a lot with my mother. Yes, they converse in Tamil and with my eldest sister and brother, they all converse in Tamil. Only we younger ones because we were very young and we were going to school and it was so difficult for us to twist our tongue, you know. Eh, Tamil is no joke, the language.

(OHI-001663-MS)

Another excerpt from an interview with a British captain (Michael Gorrie, born in 1923, Pakistan, ethnicity Eurasian) indirectly reflects the multilingualism among the Indian speakers in the historical setting, as the British captain had to train not only in Tamil but also Urdu and Pashto (spelled in the interview as Pushtu) during his Malayan Civil Service (MCS). It also highlights that the Indian population did not all fit the rigid Tamil category which the language policies are based on.

IP: Would you have to learn the Indian language?

MG: Oh yes. But I already spoke fluent **Urdu** before then, because I had to do it during training as an officer cadet. We had six months' training, during which we did intensive language training because otherwise we could not communicate with our soldiers. In fact, many of the soldiers also had to learn the language because Urdu was not necessarily their mother tongue. They would have spoken different Indian languages. Some of them were akin to Urdu, others not. **Tamil**, for example, bears no relationship with Urdu whatsoever. But any Tamil serving in the Indian Army would know how to speak Urdu. We called it Hindustani in those days. And I also spoke some **Pushtu** (Pashto) which was the language of the Northwest Frontier, of the Pathans and the Afghan people. Very similar to Afghan language. So those were the two languages I had learnt whilst in the army. And I was demobilised with the rank of Captain at the age of 22 I think it was. Something like that. Twenty-three.

(OHI-001309-MG)

3.2.8 Eurasians, Arabs, and others

In addition to the big ethnic groups in Singapore (Chinese, Malays, and Indians), there have been many other minority groups such as the Eurasians, the Peranakans, the Arabs, and the Chindians (Chinese-Indian) who are well presented in the Oral History Interviews. Although they are subsumed at the end of the CIMO model of the ethnic composition in Singapore, as what is usually called the “Others” (Clammer 1988:96), they play a significant role in the emergence of CSE. Especially the Eurasians, though in small number in terms of population, are bundled with the Peranakan communities, which represent the genuinely Singaporean cultures of being both “local” and “international” at the same time (Clammer 1988:107). The following excerpt gives an impression of an interviewee with a Eurasian ethnic background in OHI:

I was born on the 12th of April 1925 in Singapore to Eurasian parents. On my father’s side, my great-grandfather was a Scot, Mr Palmer. And my great-grandmother was of Siamese descent. On my mother’s side, my great-grandfather was a German and my great-grandmother was of Burmese descent. My great-grandfather being Mr Alexander Fox. Of course I am a Singapore citizen. And at the time of this recording, I am 68 years old.

(OHI-001423-VP)

The origins of the Eurasians and the Arabs in Singapore can be traced back to the 16th century and the 18th century. Eurasians are, as implied by the name itself, persons with mixed European and Asian lineage, with the European part of their ancestry stemming from the Portuguese, Dutch, British, Danish, French, German, Italian or Spanish, while the Asian component of their ancestry is usually derived from the Chinese, Malays or Indians (Ho 2013b). They are believed to be Europeans traders, administrators and private individuals who traveled to Asia starting from the 16th century (Braga-Blake and Ebert-Oehlers 1992:25). Many of these Eurasians who settled in Singapore at different times were reported to be the descendants born in India, Malacca, and Macau, whose ancestry were the Portuguese who exerted their presence in India from 1505 and in Malacca from 1511 to 1641. Dutch settlements in Malacca, Ceylon, and the Dutch East Indies, as well as British colonies

in Penang, India, and Bencoolen (now the Indonesian city of Bengkulu) also gave rise to the Eurasian families in Singapore (Braga-Blake and Ebert-Oehlers 1992:28–30).

The Arabs are also a small but significant community in Singapore. The first Arabs were regarded to have arrived in Singapore in 1819, who were Arab merchants from Palembang (Ho 2013a). The majority of Arabs in Singapore are most likely to be descents of Hadhrami Arabs tracing their ancestry to the Hadhramaut region in the southern part of the Arabian Peninsula, now part of Yemen (Harasha 1996, cited in Ho 2013a:20). Many of them had been exposed to the Malay-Indonesian culture and custom before their arrival in Singapore with wealth made in the Dutch East Indies (present-day Indonesia). Besides the Eurasian and the Arab communities, the “Others” category also consists of small communities, as diverse as Filipino, Armenian, Jewish, Japanese, and people of Caucasian descent (Clammer 1988:97).

It is difficult to extract an accurate number of the Eurasian and Arab populations in Singapore because, as mentioned earlier, they find themselves under the category of “Others” as they usually identify themselves with the bigger blocs of ethnic groups, such as the Chinese, Malays, and Indians. Many Arabs, for example, officially registered themselves as Malays after Singapore became an independent state in 1965 (Mobini-Kesheh 1999:21). The major historical reason is that many of them had been exposed to Malay culture, custom, and Muslim religion before their arrival in Singapore with families in the Dutch East Indies (present-day Indonesia). The assimilation has been so pervasive that many of the younger generations of Arabs today are no longer fluent in Arabic or practice the traditions of their community (Ho 2013a). The second reason making their population count so difficult is that, traditionally, one’s community membership is only determined by one’s father’s ethnicity (Leimgruber 2013:11). For example, only persons whose fathers were of European origin or who had European surnames were considered Eurasian (Ho 2013b). Here, the traditional patrilineal idea of ethnicity and descent is at play. Nevertheless, some records show that 919 Arabs were living in Singapore in 1901, and 6,900 Eurasians in 1931 (Ho 2013a, 2013b), and by the 2010s, the number grew to about 17,000 Eurasians and 8,419 Arabs in Singapore (Anon 2012:12–13; Wong 2011:11).

Concerning language use, a general observation obtained in OHI is that the Eurasians behave similarly to the British, while the Arabs tend to show similarities with the Malays in their English proficiencies.¹¹ This is perhaps not of surprise since reports state that many of the Eurasians had become anglicized and identified more with the British than their Portuguese ancestors when they came to Singapore (see Ho 2013b). They adopted English as their first language while the Arabs were assimilated with the Malay-Indonesian culture (Ho 2013a, 2013b). Before the dominance of English, a Portuguese-based Creole language known as Kristang ‘Christian’, also a commonly used label that referred to the Eurasians, used to be the *lingua franca* of the Eurasians in Malacca and other Portuguese trading settlements (Chew 2012:121–22). It shares many similarities with Portuguese in terms of lexicon and pronunciation but has a grammar strongly influenced by Malay (Tessensohn 2001:128), and a number of older Eurasians still speak the language today (see Chew 2012; Ho 2013b). Arabic, on the other hand, is believed to be marginal or waning in the linguistic ecology from pre-colonial centuries to the late 1980s because of assimilation with the Malay-Indonesian culture (Lim 2010). However, Lim reports that it seems to come into prominence towards the start of the new millennium (2010:37–46). The reasons are new immigrants from the Middle East attracted by the more open immigration policies and efforts, made by the Arab Association of Singapore, to build a stronger sense of identity within the community (Ho 2013a).

The bottom line of the previous discussion is that the ethnic backgrounds of Singaporeans are very complex, and this diversity is also represented by interviewees of OHI in this study. Platt and Weber (1980) described the linguistic situation in Singapore before 1980 as “polyglossia” where an average Singaporean tended to have a linguistic repertoire consisting of six to eight language varieties. In the following sections, we will see how this diversity conflicts with the ideology of the top-down language policies implemented in the independent era of Singapore.

¹¹ There is only one Arab speaker in the selected interviews.

3.3 Language policies and language shift in Singapore

“Melting pot” does not come close to describing the rich cornucopia and complexity of Singapore’s language landscape. As pointed out by PM Lee (2000:4): “Singapore never had one common language. It was a polyglot community under colonial rule.” Obvious from the previous discussion, due to the influx of migrants, consisting of different ethnic groups originating mainly from Southeast Asia, China, India, and European countries who spoke a variety of languages and dialects, the language ecology of Singapore has been rich since its early days. In recent years, the increase in immigrants residing in Singapore has brought even greater richness to the sociolinguistic landscape of Singapore. How to manage the multilingualism in Singapore, uniting the multifarious cultural and linguistic groups has been the focus of governmental policies, and in the words of PM Lee, a “lifelong challenge” (Lee 2000).

Language policies, along with political upheaval, and migration patterns are major factors influencing the changes of linguistic ecology, and Singapore is a case in point (see Lim 2009:201). The bilingual policy has successfully transformed Singapore from an English-knowing to an English-as-first-language country, and the Speak Mandarin Campaign has effectively replaced the various Chinese dialects with Mandarin, especially among the younger generations of the Chinese community in Singapore. This next section explores the ideologies of different language policies in Singapore and to what extent these language policies relate to the language shift which Singaporean society has experienced and is still experiencing. It will be useful for a better understanding of the linguistic ecology in Singapore, especially of the study of CSE, the restructuring of which is, consequentially, influenced by new dominant languages.

3.3.1 Official languages and the concept of bilingualism in Singapore

The Republic of Singapore Independence Act of 1965 gave Malay, Mandarin, Tamil, and English the official language status. The policy corresponds to the previously mentioned ethnic categorization of the “CMIO” model, namely, Chinese, Malay, Indian, and Others.

The policy entails an emphasis on using English and the respective mother tongue languages of the three main ethnic groups (Chinese, Malay, and Indian). Of the four official languages, English was chosen as an official language in Singapore for pragmatic reasons. Policy makers (e.g. Lee Kuan Yew, Gopinathan) considered it as essential for economic survival, and an ethnically and culturally neutral language for different ethnic groups (see Lee 2000). It is, therefore, suitable to be the *de facto* official language and *de jure* working language of the country. The other three official languages, Mandarin, Malay, and Tamil, on the other hand, are seen as “mother tongues” of the Chinese, Malay, and Indian communities, respectively, and an essential link to traditional culture (see Pakir 2004). They were adopted by the government to give Singaporeans an anchor in their ethnic identities and traditional values to avoid excessive Westernization and to prevent deculturalization (Ng 2011:4).

Though sharing similar social, cultural, religious, and linguistic backgrounds, Singapore and Malaysia have developed in different directions with regard to language policies from the time Singapore gained independence in 1965. Malaysia chose Malay as the only official language while Singapore decided for a bilingual system with English as the co-official language along with Malay, Mandarin, and Tamil. The reasoning behind this language policy is explicitly expressed in Lee’s memoirs *My Lifelong Challenge: Singapore’s Bilingual Journey*:

If we were monolingual in our mother tongues, we would not make a living. Becoming monolingual in English would have been a setback. We would have lost our cultural identity, that quiet confidence about ourselves and our place in the world. (Lee 2000:181)

The aim of the official language policy in Singapore was to reduce language barriers and to promote better cross-ethnic communication amongst the three major ethnic groups as well as to prevent culture and heritage erosion of these groups. However, many scholars criticize that the “mother tongue” policy is an oversimplification of the differences among heterogeneous communities making each definable in terms of one single language (see Chew 2012; Ng 2011; Siemund and Li 2020). The community membership of Singaporean

citizens is determined by one's father's "race" (Leimgruber 2013:11),¹² and the mother tongue policy is closely linked to the ethnic policies. Before 2011, the decision of a child's official mother tongue was dependent on the ethnic group of the child's father, regardless of what first language the child may actually have spoken. For example, if a child spoke Cantonese, whose father is ethnically Cantonese Chinese and whose mother belongs to one of any of the other ethnicities, Mandarin would be regarded as the child's official "mother tongue" in Singapore. The categorization became less rigid in 2011 when the Parliament of Singapore defined mother tongue not just by the father's ethnicity of a child, but also home language or first language acquired by the child (see ICA 2011). However, the patrilineal ideal is still so prominent in the Singaporean context that a recent campaign with the aim to revitalize the use of dialects in Singaporean youths was named "My Father Tongue" campaign (Goh, Lim, and Cherie 2016).

3.3.2 The origins of language-related policies in Singapore

Many scholars assume (e.g. Alsagoff 2007; Bolton and Ng 2014; Cavallaro, Ng, and Seilhamer 2014; Lim et al. 2010) that language-related policies in Singapore arose suddenly since the independence of Singapore in 1965, and largely due to the political ideology of PM Lee. The pragmatic reasons masked behind the top-down language policies of the independent city-state seem straightforward: English is the language of technology and economic development, as well as inter-ethnic communication, while among the mother tongue languages, as is put by Chew (2012), "Mandarin is seen as a means to forge stronger business ties with China, and to a similar extent, the Malays and the Indians have been encouraged to engage the Islamic and Indian world respectively" (Chew 2012:173). However, the ideology of assigning one single language to one ethnic group did not appear suddenly, but has its historical roots in the precolonial period.

¹² In the Singaporean context, the term *race* is used in everyday discourse to denote ethnicity without the commonly negative connotation in the European context (see Leimgruber 2013:11). Hereafter, I will use the term ethnicity.

First of all, ethnic compartmentalization existed even before the colonial period under the administration plan of Sir Thomas Stamford Raffles in 1826, as efforts were made to assign each immigrant an area of settlement (LePoer 1991:15). This division resulted in residential enclaves such as Chinatown for Chinese, Geylang Serai for Malays, and Little India for Indians (Hodder 1953). The official division of Singapore's population into racial categories – Chinese, Malay, Indian, and Others – began in colonial times under British rule (Hirschman 1986). In addition, the ethnic groups occupied different economic sectors: Chinese in trade and commerce, Malays in agriculture, and Indians in administrative posts or plantation work (Hirschman 1986).

With regard to education, the British envisioned English as a common tongue between the highly diverse residents of the island, though they were quite relaxed about what ethnic languages the Singaporeans used. Raffles founded the “Institution” (subsequently referred to as Singapore Institution) on 1 April 1823, with the aim to establish Singapore as the center of higher education for the local aristocrats, professional teachers and government servants (Turnbull 2009:45). However, Raffles' vision was not shared by his successor, Dr John Crawfurd, who proposed to establish elementary education in Malay, Chinese varieties, and above all English (Beamish and Ferguson 1985:43). Private educational efforts, on the other hand, started early after Raffles' arrival in 1819. A group of European merchants raised money to complete the Institution founded by Raffles in 1835 (Prasad and Koh 2014). Reverend F.J. Darrah founded the Singapore Free School in 1834, and moved into the Institution's building (then operated as Singapore Institution Free School) in 1837 (National Library Board 2014). Different from Raffles' initial intension of establishing a college, it was established as an elementary school, which offered both English and vernacular education in Malay, several Indian languages and Chinese varieties (Cheeseman 1935:16). However, the Malay classes were closed in 1842 due to lack of enrollment, and English eventually became the medium of instruction of the school. In 1856, the Singapore Institution Free School was renamed Singapore Institution (Prasad and Koh 2014). It gradually turned into a high school and was renamed Raffles Institution in the 1860s (Cheeseman 1935:16). In the early decades of colonial rule, several English schools were

opened by Christian missionaries (Wasserfall 2019:18). After Singapore became a Crown Colony in 1867, education was transformed and became more systematic. Government schools were established in 1872 when the Education Department was established, and Inspector of Schools appointed (Wasserfall 2019:18). By 1902, English was adopted as the sole medium of instruction in the so-called “Branch English School” established by the government (Gwee and Doraisamy 1969:37). The emphasis on English education of the British and a lack of systematic vernacular school were also pointed out by PM Lee:

Singapore never had one common language. [...] The British left people to decide how to educate their children. The government provided a limited number of English language schools to train people to be clerks, storekeepers, draftsmen, and such subordinate workers, and Malay language primary schools for Malays. Indians ran their own Tamil and other Indian-language schools or classes. The Chinese set up schools financed by successful members of their community, to teach in Chinese. Because the different races were taught their own languages, their emotional attachment to their mother tongue was deep. They were like the 5 million people in Quebec tenaciously holding on to French in a continent of 300 million English speakers. (Lee 2000:146)

It is not accurate that the British government only provided Malay language primary school for the Malays though. Malay was the only official vernacular of Singapore besides English in the colonial period. A Malay high school was established in 1876 and was transformed into a Malay Training College for teachers two years later (Wasserfall 2019:18). Additionally, there is governmental coercion to bring Malays into the schools free of charge (Wasserfall 2019:18). Richard Olaf Winstedt, a Malay scholar was appointed as Director of Education of the Colony and the Federated Malay States in 1916 (Barrett 2004). From the 1920s on, secondary Malay education was provided for the Malays by the government (Gwee and Doraisamy 1969:105–9).

Different from the assumption that the preference for Mandarin in Singapore started after 1959 (e.g. Cavallaro and Ng 2014; Lim 2010), historical evidence shows that Mandarin began to gain significance after 1911. Traditionally, Chinese private schools or tutors in Singapore followed the classical education of China (Wasserfall 2019:19). Up to the early 20th century, Chinese schools in both China and Singapore were taught in dialects, for

example Hokkien societies set up schools to teach in Hokkien, Cantonese societies set up schools to teach in Cantonese (see Turnbull 2009:117; OHI-000564-HKW). The Xinhai Revolution in 1911 (辛亥革命 *xīnhài gé mìng*) led by Kang Youwei (康有为 *kāng yǒuwéi*) and the subsequent efforts in reforming the Chinese education according to Western and Japanese models had a long lasting effect on Chinese education both domestic and overseas (see Kaske 2008). According to Bao (2021:23), after Xinhai Revolution, China established Western-style schools with Mandarin as the medium of instruction.¹³

Although Mandarin was not the language of most of the immigrants in Singapore, it was considered a unifying factor by the Chinese leadership factions of both Singapore and China (LePoer 1991:33). Lim Boon Keng (林文庆 *lín wénqìng*), a third generation Peranakan, is an eminent figure in promoting the use of Mandarin in Singapore. Lim was hugely influenced by the reform movement in China during the late 19th century, and supported Sun Yat-sen and Kang Youwei, both of whom visited Singapore in the 1900s (Turnbull 2009:121). Turnbull (2009:117) commented that he exemplified the multifaceted cultural orientation of the Peranakans in Singapore. He spent a brief period at a Hokkien school and thereafter started his English education at the Government Cross Street, and later enrolled in Raffles Institution in 1879 (Tan 2008:109). He studied medicine at the University of Edinburgh with the Queen's Scholarship and graduated in 1892 (The Straits Times 2004). He is a staunch supporter of the British colonial government, pushing for social reforms such as female education and at the same time, a fervent advocate of Confucianism, promoting the use of Mandarin instead of dialects as medium of instruction in Singapore. He believed that English-educated Chinese should also be proficient in Mandarin (Song 1984:236; Turnbull 2009:121; Wang 2003:167–68). Therefore, together with Song Ong Siang (宋旺相 *sòng wàngxiāng*) and other members of the Straits Chinese community, he founded the Singapore Chinese Girls' School and started organizing Mandarin Classes in 1899 (Li

¹³ Mandarin, known as 官话 (*guānhuà*, literally 'language of officials') refers to different varieties in different times. For most of the Ming (1368–1644) and Qing dynasties (1644–1911), Mandarin was a koiné based on dialects spoken in the Nanjing area, but by the middle of the 19th century, the Beijing dialect had become more dominant (Coblin 2000:549–50). This led to the decision of the National Language Unification Commission to settle on the Beijing dialect as the standard national language in 1932 (Ramsey 1989). In 1949, the People's Republic of China retained this standard, calling it Putonghua (普通话 *pùtōnghuà* 'common speech') (Ramsey 1989:3–15).

1991:61–66). His efforts in spreading Mandarin was even named “the first Speak Mandarin Campaign of the Babas” (Rudolph 1993).

Between the World Wars (1919–41), the ties between the Straits-born Chinese and their homeland in China were strong, not only in the political domain and the economy but also considering educational exchange.¹⁴ For example, Chinese teachers and textbooks were sent by the Kuomintang (国民党 *guómíndǎng*, KMT, the Nationalist Party) to Singapore to encourage the use of Mandarin in Singapore’s Chinese schools. Since the establishment of the first Chinese secondary school in Singapore in 1919, a growing number of Chinese primary schools taught in Mandarin. Despite the discouragement from the colonial power to control the use of Mandarin, aware of the growing left-wing politics in Chinese schools in the late 1920s, Mandarin had become the medium of instruction in all of Singapore’s Chinese schools (LePoer 1991:33).

After the independence of Singapore, Lee Kuan Yew reinforced the choice of Mandarin over other Chinese dialects, extending its importance not just to the educational domain, but also to its use as a home language. He saw the use of different dialects as a problem, especially in the Singapore Armed Forces:

We were saddled with a hideous collection of dialects and languages and faced the prospect of going into battle without understanding each other in any of the four official languages. Many could only speak dialects, requiring special Hokkien-speaking platoons. The Chinese were speaking one of more than seven different Chinese dialects at home but learning Mandarin and English in school, neither of which they used at home. (Lee 2000:146)

Here, we see that the heritage dialects were considered by Lee (2000) as “obstacles” of nation building. The denial of heritage dialects reflects the ideology that languages perceived to be obstacles to economic development should have little place in the linguistic ecology of Singapore (Tan 2007). Therefore, he introduced the teaching of three mother tongues representing the three major ethnic groups into English schools as one of his measurements

¹⁴ Politically, there was an active Singapore branch of the Nationalist Party of China (KMT), boycott of Japanese goods, and mass support for Chinese nationalism (see LePoer 1991:33). Economically, there was heavy investment in Chinese industry and education by Singapore’s China-born businessmen, encouraged by Sun Yat-sen in the early 1920s (see LePoer 1991:33).

to mold Singapore into a nation. The task of unifying the highly heterogeneous group of people in Singapore with this highly top-down language policy was seen by Lee, in a speech delivered in 1984, as “extremely difficult, and there was no guarantee that we could succeed.” He did succeed, as we will later see in Section 3.4, the implementation of the language policies since the independent era has shaped the Singaporean population from rather diverse language backgrounds to an essentially bilingual model comprising English and one of the so-called “mother tongues” (i.e. Malay, Mandarin, and Tamil).

3.3.3 The status of English in Singapore

Being essential to Singapore’s independence from the early years, and crucial to science and technology, English is the language of technology and economic development. The policy of multilingualism was designed to serve the goal of establishing an equal status of all official languages (English, Mandarin, Malay, and Tamil). But in reality, English is still the most important language in Singapore in the field of government administration, business, law, banking, and accountancy. Besides, English is the official First Language (FL) in education, and the medium of instruction while the others have the Second Language status. Across ethnic groups, English serves as a neutral language, both ethnically and culturally, which can bridge cultural differences. For individuals, English is also crucial for one’s career path (Bokhorst-heng 1998). Besides the historical reasons mentioned earlier, the importance and dominance of English in Singapore may be partly due to language planning and partly because of the development of English into a lingua franca as a consequence of globalization. It is generally assumed that a negligence of English would result in marginalization of the country and a denial of the extensive resources available in English (Chew 1999).

Yet, the implementation of English as the medium of instruction is not without obstacles. During the earliest phase of the independent era, many Chinese-speaking parents could not understand why their children had to learn English under their own elected government, while even during the colonial era, English had not been compulsory and their children had been allowed to be educated completely in Chinese (see Lee 2000:148).

Therefore, during the early implementation of the policy, there were lots of protests, especially in the Chinese-speaking community, and there were cases of university students using Chinese in their English examination. PM Lee was well aware that the switch to English as the medium of instruction could not be made swiftly:

Not wanting to start a controversy over language, I introduced the teaching of three mother tongues, Mandarin, Malay, and Tamil, into English schools. This was well received by parents. To balance this, I introduced the teaching of English in Chinese, Malay, and Tamil schools. Malay and Indian parents welcomed this but increasing numbers preferred to send their children to English schools. A hard core of the Chinese-educated did not welcome what they saw as a move to make English the common working language, and they expressed their unhappiness in the Chinese newspapers. (Lee 2000:146)

However, evidence has shown that even back in 1965, the majority (57%) of primary students in Singapore were already enrolled in English-medium schools and the number had been increasing for some time while the enrollment in Malay-, Tamil-, and Chinese-medium schools had been declining (Chiew 1980; Khoo 1983). Although English had little basis in the Singaporean community, with virtually none of the Singapore population speaking English as a native language in 1965 (Arumainathan 1971), its high status, prestige, and capital have made it well accepted as the main medium of education in Singapore. The dominance of English and the language policies in favor of English as a common working language have resulted in a shift towards English, which has extended even beyond the administration and education level into the private domain.

3.3.4 Speak Mandarin Campaign

Over the past few decades, the Singapore government has actively encouraged the use of Standard (British) English over Colloquial Singapore English as well as Mandarin over Chinese dialects, i.e. Cantonese, Hokkien, Teochow, etc. The ideology of the Singapore government seems to be in line with one of the assumptions of language planning by Dixon (Dixon 2005:632) that a government should encourage the use of high-status languages. The

Speak Mandarin Campaign and the Speak Good English Movement are two very good examples.

The Speak Mandarin Campaign (SMC) is an initiative launched in 1979 by the government of Singapore to encourage the Singaporean Chinese population to speak Mandarin. The aim of the SMC is to promote the status of Mandarin as an intra-lingua franca of the Chinese both in school and at home, and to eliminate the use of Hokkien and other Chinese dialects in the mass media and educational domain. As mentioned before, these dialects were seen as “obstacles” of nation building. It was regarded as a burden for students to use Chinese dialects at home who had to master English and Mandarin at school. As expressed by PM Lee (2000:154): “It would make it easier for students to master English and Mandarin in school if they spoke Mandarin at home and were not burdened by dialects.” SMC has become an annual event promoting the use of Mandarin. Measurements have included, for example, discouraging the Chinese civil servant from using dialects during work, and producing cassette tapes of Mandarin lessons and distributing them via telephone to the public in the 1980s. A series of comic books called “Mr Kiasu learns Mandarin” was published for the public to learn Mandarin in a fun and light-hearted way. More recent measures were tied with the development of modern technology, such as organizing Mandarin Film Festivals and launching a bilingual language learning tool in the form of smartphone software. 50,000 commonly used business terms, in English and Mandarin, were programmed into the database of the software (Sim 2016).

The start of the Speak Mandarin Campaign during the 1970s had been regarded as a failure, but after a three-decade long campaign, it has proven to be an overwhelming success. In the 1970s, there were protests of dialect speakers, especially the elderly, complaining that Mandarin fails to correspond to the language profiles of local Chinese, and there was resistance of parents to register their children’s name under the Mandarin Pinyin version (see Gopinathan 1998). However, recent census statistics (Wong 2011) and scholarly research (Ng 2011, 2017) show an increase of Mandarin use at home and growing positive attitudes towards Mandarin. On the other hand, the use of Chinese dialects has decreased significantly. The number of speakers using Mandarin as a home language has increased remarkably from

less than 1 percent in 1957 to approximately 35 percent in 2015, while Chinese dialects have declined from around 75 percent to less than 15 percent in the same timeframe (Wong 2016). In Ng's (2017) study, the majority of respondents stated that they like speaking Mandarin and want their children to speak Mandarin, as well as agree that Mandarin is the mother tongue of Chinese Singaporeans. Besides, pragmatic reasons seem to reinforce the growing positive attitudes towards Mandarin as the majority of Singaporeans view Mandarin as an important language in trade and business with China (Ng 2017:36).

On the other hand, the regional Chinese dialects only play a marginal role today in Singapore, as one of the effects of SMC. As dialects were prohibited in the mass media, the former inter-lingua franca of the Chinese community, Hokkien, though still used in the 2011 parliamentary elections, was seen as only serving entertainment and comic purposes than playing a politically decisive role (see Chew 2012:174).

Additionally, the SMC has resulted in more Mandarin lexical influences in CSE. The occurrence of the Mandarin pragmatic particle *bah*, which was undocumented in CSE, is a case in point (Leimgruber 2015). Li et al. (2021) also found that there is an increased use of *ah* in CSE, which is assumed to be the result of Mandarin influence. Chew (2012:110) reported that Singapore English and Malaysian English shared common features before the 1980s, but they began to grow distinctively apart. Colloquial Singapore English began to borrow lexical items from Mandarin Chinese, e.g. *mah fan* 'bothersome' (麻烦 *máfan*), *gua gua jiao* 'make a big deal out of nothing' (呱呱叫 *guāguājiào*). Malaysian English, on the other hand, began to have Malay-exclusive borrowings such as *penghulu* 'the headman of the village', *bumiputra* 'the original inhabitant of the land', *lepak* 'someone who is idle and likes to waste time' and *lesen terbang* 'a driving license that is obtained illegally' (Low 2010). Phonological distinctions between CSE and Malaysian English were also discussed (see Low and Kuang 2016). For example, it is found that CSE speakers maintain some distinctions between the long and short vowel pairs (e.g. between /ɒ/ and /ɔ:/) whereas Malay English speakers tend to conflate the long and short vowel pairs (Tan and Low 2010). Chew (2012) attributes these distinctions to the separate educational and language policies of Singapore (i.e. the implementation of the SMC) and Malaysia (i.e. the widespread

establishment of Malay-medium schools). However, no evidence has been offered yet as to whether the phonological distinctions are due to the different phonological features between Mandarin Chinese and Malay.

3.3.5 Speak Good English Movement

The first Speak Good English Movement (SGEM) took place in 2000, two decades later than the SMC. The background of the SGEM provides a contrast between a local code of English that is gaining popularity among Singaporeans and a standard code of English that is narrowly defined and promoted by the government. The increasing research in the field of CSE and its increasing popularity in the media in the 1990s sparked debates on the use of CSE in schools and on standards of English in education (Bokhorst-Heng 2005:189; Gupta 2010:57–58; Wee 2018:37–38). The pronounced aim of the SGEM is to increase the use of Standard English and discourage the use of CSE among primary school children and in the media to enhance Singapore’s industrial and economic status in the region (Rubdy 2007; Tsui and Tollefson 2017). The movement was started with an official speech by the then Prime Minister, Goh Chok Tong, who considered CSE as an “enemy” of the economic development and “a corrupted form of English”. As expressed by Goh (2000, cited in Chew 2012:176), “investors will not come if their supervisors and managers can only guess what our workers are saying.” Many believe that CSE is a “less prestigious dialect associated with low social status” and an obstacle to the nation’s continued economic growth, which impedes Singapore’s place in the global marketplace (Rubdy 2007:308). Promotions of the SGEM events often put a negative connotation on the CSE speakers as being uneducated or uncultured (Hoon 2003).

The government’s stance against CSE had remained strong and unambiguous over the past three decades, though a recent report published in *The Straits Times* showed that the anti-CSE attitude of the government has gradually changed (Ang 2019). For policymakers in the past, a co-existence of CSE and Standard English was not “an option” (Rappa and Wee 2006:95), which is evident in a comment by the first Prime Minister on CSE as a

“handicap” (Tan 2007:88). This is also visible in a strong-worded opening statement in an open letter: “While Singlish might be a fascinating academic topic for linguists to write papers about, Singapore has no interest in becoming a curious zoo specimen to be dissected and described by scholars” (cited in Alsagoff 2015:127). However, the recent stance adopted by the government has grown closer to the diglossic view of linguistic scholars that CSE is a perfectly viable vernacular, existing side by side with the standard norm. Jason Leow, the Speak Good English Movement steering committee chairman of 2019, said in an interview with Ang (2019), a reporter of *The Straits Times*:

Singlish, text-speak, Net-speak, millennial-speak and Gen Z-speak all exist outside standard English. You don’t have to choose... but we do hope that people can tell the difference between Singlish and standard English. And our role is to have resources ready to enable that. Other than that, we don’t make a value judgment on who has good or bad English. (Leow Jason 2019, cited in Ang 2019)

The message seems to acknowledge CSE as a cultural marker for Singaporeans, and the approach of promoting Standard English has changed from banning CSE completely to raising awareness of the differences between Standard English and CSE among Singaporeans. Today, measures include the traditional methods, such as developing audio lessons and language tips, and organizing workshops and seminars, as well as presenting personal stories of the committee members who successfully acquired Standard English as multilingual speakers (see Ang 2019). The implementation of the SGEM has been very successful in promoting the use of English in general, because ever since its implementation, the literacy rate in English increased from 70.9 percent via 79.9 to 83.1 percent, and the proportion of bilingual speakers increased by 17.2 percent from 56 to 73.2 percent between 2000 and 2015 (see Table 3.3) (Wong 2011; 2016). A comparison of attitudes towards English and Mandarin by Ng (2014) shows that most Singaporeans consider English more important than Mandarin. The policies also induce a sharp increase in English as a home language (see Section 3.4).

	2000	2010	2015
General Literacy Rate	92.5	95.9	n.a.
Among Literate Resident Population			
Literate in English	70.9	79.9	83.1
Literate in Two or More Languages	56.0	70.5	73.2

Table 3.3: Literacy among resident population aged 15 years and over (data from Wong 2011; 2016)

Nevertheless, the initial goal of the SGEM, the removal of CSE, has remained unachieved. CSE, a hybrid language that used to be disapproved, has become a marker of the Singaporean identity. The Oxford English Dictionary added 19 CSE words to its lexicon in 2016, including *ang moh* ‘Caucasian’, *Shiok* ‘delicious or great’ and *lepak* ‘to relax’ (C. L.-L. Tan 2016). A popular novel titled *Sarong Party Girls* was written in CSE by Cheryl Lu-Lien Tan, a New York-based Singaporean journalist (C. L.-L. Tan 2016). CSE has become so prominent that it was included in advertising campaigns for SG50 (the celebration of the 50 years of Singapore) and on floats in the National Day Parade 2015 (T. Wong 2015). The reason behind the bottom-up acceptance of CSE is explainable. While it may seem unproblematic for the policymakers to promote the use of “correct” or “Standard English”, constructing a purist society where the whole society speaks “Standard English” is, if not impossible, idealistic, especially in a multilingual society like Singapore. Many linguistic scholars (e.g. Leimgruber 2013:250; Leimgruber et al. 2018; Lim et al. 2010; Siemund et al. 2014) predict that CSE is unlikely to disappear in the near future, given the increased number of migrants, inter-marriages, transnational trade and digital communications. Standard Singapore English may be the desired prescribed norm for the city-state, but CSE is still the language closer to the culture of most Singaporeans (Cavallaro and Ng 2009).

3.3.6 Other language campaigns

Besides the above-discussed language policies on the state governmental level, there are some other language campaigns organized by the Malay Language Council and Tamil

Language Council to promote the status of Malay and Tamil, respectively. The goal is to encourage the local communities to speak their mother tongues in their daily lives as well as to instill confidence and pride among the communities in speaking their mother tongues. The Malay Language Month and the Tamil Language Festival are two very good examples. The Malay Language Month (*Bulan Bahasa* in Malay) was initiated in 1982 by the Central Council of Malay Cultural Organizations in an attempt to address the decreasing proficiency standards among the younger generation of the Malay community (Malay Language Council 2019). Revived as a biennial event in 1988, it was well received by the local community and, later, became an annual event starting in 2010 (Malay Language Council 2019). Similarly, the Tamil Language Festival has been held annually since 2007 (Tamil Language Council 2020). However, the scale and effect of these campaigns cannot be compared with that of the SMC and SGEM, as evident in the census of the Department of Statistics 2015 that home language use of Malay and Tamil in the Malay and Indian communities has remained rather stable (see Wong 2016).

3.4 Language shift in Singapore

Language shift refers to the process whereby members of a community in which more than one language is spoken abandon their original vernacular language in favor of another (Thomason and Kaufman 1991; Mufwene 2008; Wendel and Heinrich 2012). We can clearly observe such a process happening in the city-state. Over the past five to six decades the Singaporean population has been shifting from a rather diverse language background to an essentially bilingual model comprised of English and one of the prescribed mother tongues. A schematic representation is shown in Figure 3.6:

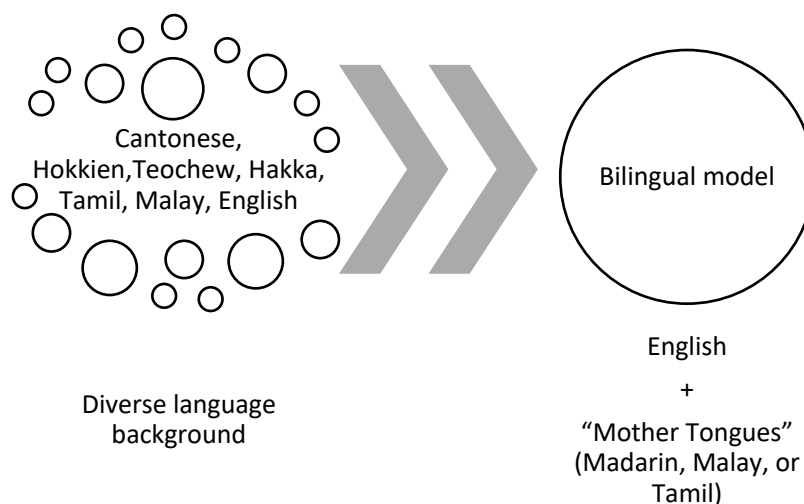


Figure 3.6: Diverse language background to English plus one of the Mother Tongues

Drawing data from Singapore Department of Statistics in 2015, we can see a language shift reflected in changes in home language use over the past six decades, as presented in Figure 3.7. In 1957, the census recorded only 1.8% Singaporeans using English as home language, but the proportion jumped to nearly 32% in 2015. Within the same period, the use of Mandarin in the Chinese community shows a similar increase, though such growth seems to stabilize in the past five years whereas English has become the main language at home. On the other hand, the use of Chinese vernacular languages continuously decreases, apparently being replaced by Mandarin, the official mother tongue of the Chinese-speaking community. By comparison, home language use of Malay and Tamil in the Malay and Indian communities remains rather stable.

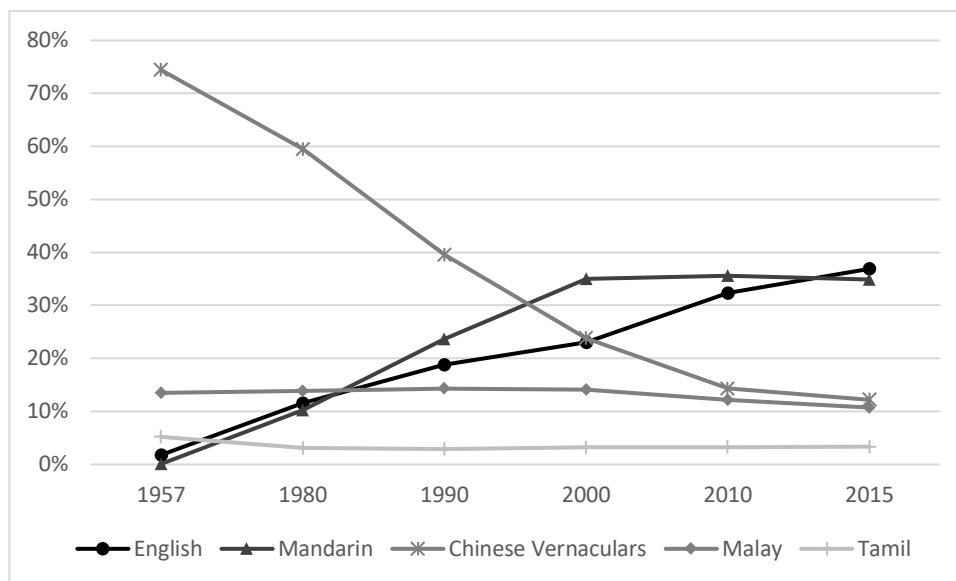


Figure 3.7: Changes in home language use over time (data from Cavallaro and Ng 2021; Wong 2011, 2016, 2019)

A similar picture emerges if age grading is considered, as demonstrated in Figure 3.8. Here, four age cohorts are distinguished in descending order to represent change in real time (over 65, 45–64, 25–44 and 5–24). With decreasing age (moving from the older to the younger generation from left to right), there is a strong increase in English competing with an almost parallel increase in Mandarin, though a slight decrease in the youngest group (see group 25–44 and group 5–24). Again, there is a sharp decrease in Chinese vernaculars. The Malay and Indian communities do not show such pronounced developments but rather equal percentages across the four age groups. What is worth noting is that all ethnic groups participate in the shift towards English as a home language, which is in line with the results of (Cavallaro and Serwe 2010), in that young Singaporeans aged 18–20 were more comfortable in using English in their daily lives.

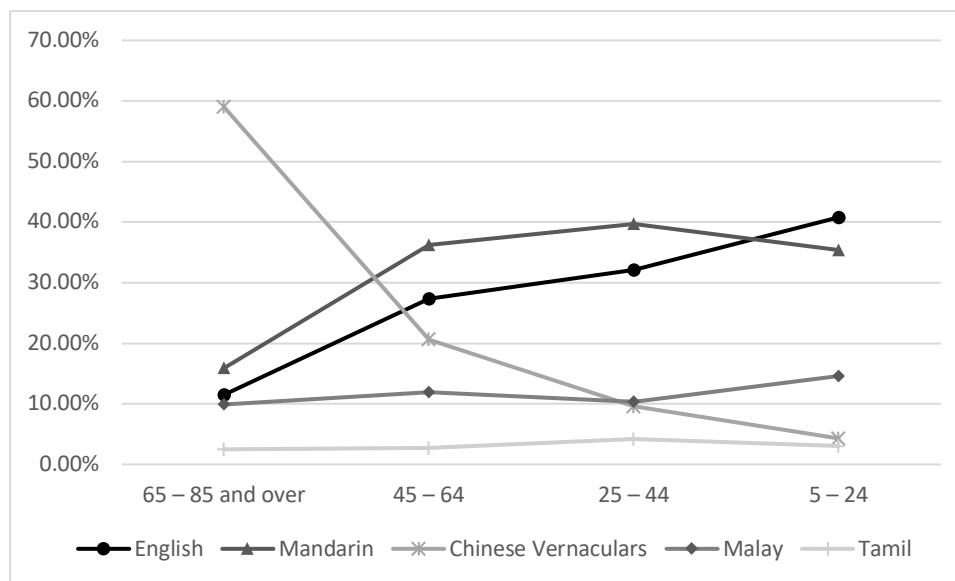


Figure 3.8: Differences in home language use according to age (data from Wong 2010, adapted from Siemund and Li 2020)

The causation of language shift has been explained on the basis of a variety of factors by sociolinguists. These include, for example, demographic variables such as birthplace, age, ethnicity, period of residence and education (Portes and Hao 1998) at the individual level as well as the size of the migrant population at the societal level (Holmes and Wilson 2017:55–74); linguistic variables, such as the typological similarity and difference between one’s language of origin and the target language (Siemund 2013); language variety (dialect or standard) and language planning (Fishman 1991; Hornberger 1988); pedagogical variables, such as the existence of language programs in the community and in the school, and the quality of teaching materials or curriculum (Krashen, Tse, and McQuillan 1998); and language attitudes and ideologies (Gardner 1985).¹⁵

In the case of Singapore, language policies, i.e. the promotion of English and Mandarin Chinese over regional dialects, played a decisive role in language shift over the past 60 years. Historically, Singapore is highly multilingual, as discussed in Section 3.2, but since the implementation of the language policies (official languages and bilingual education)

¹⁵ Other overlapping factors mentioned in scholarly circle include sociocultural and sociopolitical variables, such as reasons for migration, modes of acculturation such as social isolation, racism, employment channels (see Portes and Rumbaut 2001) as well as social network (see Gumperz 1982; Wei 1994).

in the independent era, Singapore has been molded into a bilingual society: English plus one of the three mother tongues. The effect seems to be reinforced by the Speak Mandarin Campaign and the Speak Good English Movement discussed earlier (see, amongst others, Bokhorst-Heng 1999; Rubdy 2007). Initiated in 1979, the Speak Mandarin Campaign can be regarded as responsible for the sharp decline in vernaculars and the concomitant increase in Mandarin in the Chinese community. The strong increase in English can be attributed to the bilingual policy and the role of English as a special official language. We can see that the Singaporean government has been actively promoting Standard English, and a good command of Mandarin, and the Singaporean people seem to follow their government's decisions. What should not be neglected, however, are the pragmatic reasons such as social mobility and economic success where "instrumental attitudes relating to the perceived utility of languages loom large" (Bolton and Ng 2014:315). English as a home language seems to gain even higher popularity among the younger generation, as they regard English essential for their success (Cavallaro and Serwe 2010). Curiously enough, the actual target of the Speak Good English Movement, namely, CSE, which the government frowns upon and considers broken English, seems unaffected and gains more popularity.

Traditionally, depending on the size of the migrant population, three classic scenarios of language shift can be identified: (i) migrant minorities, (ii) migrant majorities, and (iii) non-migrant communities (Holmes and Wilson 2017:55–74). In the first scenario, the migrant community belongs to the minority group. Due to pressure from the wider speech community, such as the pressure of institutional domains such as school and the media, economic reason of obtaining work, and the social status and prestige associated with the dominant language, the migrant community gradually adopts the language of the host country, which replaces their heritage language as the primary language of communication and socialization. An example of the scenario of migrant minorities is the Turkish immigrants recruited by the Federal Republic of Germany as "guest workers" in the 1950s, as a reaction to the labor shortage resulting from the economic boom of the post-war era (Kohlmeier and Schimany 2005:13). Although a considerable number of "guest workers" returned to their home countries, their population continued to increase as a consequence of

family reunions (Heckmann 2003). Among other qualifications such as basic skills and educational achievement, sufficient knowledge of the German language is considered as the fundamental precondition for their successful integration (see Hönekopp 2002; Suntum and Schlotböller 2002). Although the Turkish-German linguistic mixture has become a creative approach to both languages adopted by a group of native and migrant youngsters as a symbol of their socio-cultural identity, the Turkish language has gained only little appreciation (Kohlmeier and Schimany 2005:36). On the other hand, some studies on the educational standard of second and third generation immigrants of Turkish origin have shown that they have approached the level of German children (Gogolin 2000; Hunger and Thränhardt 2004).

In the scenario of migrant majorities, the migrant community becomes the dominant group – either economically or demographically – in the new land. Many former colonial powers such as Portugal, Spain, Britain, France, among others, fall into this category. Their languages were imposed on the various communities by becoming the official languages in the government and public sectors as well as in the educational system. However, linguistic subjugation and language shift do not always happen in this scenario, as it is not likely for a single alien and imported language to replace hundreds of indigenous vernacular languages, e.g. in India and Papua New Guinea.¹⁶ However, when multilingualism is not well-established, the indigenous language is often under threat. For example, in New Zealand, English has become a dominant language, and put the indigenous language Maori under threat (Holmes and Wilson 2017:57). The language shift followed a typical pattern over three generations: (i) monolingual in the indigenous language; (ii) bilingual in the indigenous language and the colonial language; and (iii) monolingual in the colonial language. For example, the indigenous group in New Zealand was monolingual in Maori in the late 19th century. However, their second generation was bilingual in Maori and English, and the majority of their third generation became monolingual in English in the second half of the 20th century (Holmes and Wilson 2017:57).

¹⁶ Studies on language shift in India (Laitin 1993) and Papua New Guinea (Kulick 2004) also show that English became the dominant language in both of these countries, but the existence of multilingualism prevents their indigenous language from endangerment.

The third scenario is non-migrant communities. In this case, we can see that language shift is not always the result of migration, but political, economic and social changes. One example Holmes and Wilson (2017:58) gave is Iran, where Azeri (or Azerbaijani), the largest minority language, though not in immediate danger, is put under threat by Farsi, the language of the largest and most powerful group, the Persians. As Iran struggles to achieve national unity, Farsi, the official language of the country is further promoted. It is taught in schools and dominates the public space, e.g. in Tabriz, the most populated city in northwestern Iran and Iran's historic Azerbaijan region (Swietochowski 1995), street signs are in the Farsi language, but not Azeri (Holmes and Wilson 2017:58). A recent study (Mirhosseini and Abazari 2016) on language attitudes among a group of 104 bilingual Farsi-Azerbaijani speakers in Tabriz shows that despite the positive emotions and feelings of the participants towards Azeri, there is hesitation and reservation in approving of its use in domains of education and new media.

We can conclude from the above discussion that language shift is to be understood, besides demographic factors and migration patterns, as part of broader socio-cultural, economic and political considerations (also see Milroy 1997; Blommaert 2010). Though sharing similar colonial histories and demographic compositions,¹⁷ Singapore finds itself in many ways different from other former British colonies in Asia (e.g. Malaysia, and Hong Kong) in language planning in the post-war decolonization process. Unlike Malaysia, which made *Bahasa Malayu* the only national and official language to build national identity in 1967, Singapore implemented four official languages, actively promoted the English plus mother tongue bilingual education, as well as the use of Mandarin Chinese and English. As a result, Malay became the *lingua franca* among the different ethnic groups in Malaysia while Singapore has moved from multilingualism to English-plus bilingualism and Mandarin as the pan-Chinese intra-ethnic language. Interestingly, despite the inhibition of English in the school system, English is still regarded as the language of prestige in today's

¹⁷ All three regions were colonized by the British in the 19th century and all for more than 100 years (Malaysia 131 years, Hong Kong 155 years, Singapore 133 years). In addition, the Chinese make up a significant part of their population (Malaysia 23.2%, Hong Kong 92%, Singapore 74.4%). Data are from the departments of statistics of Malaysia, Hong Kong and Singapore respectively (see Ho 2017; Tang 2017; Wong 2019).

Malaysia (Ng and Cavallaro 2019:48–50). Compared to Singapore and Malaysia, the majority of residents in Hong Kong during the colonial period were Chinese monolinguals – Cantonese, to be more specific – due to its proximity to Guangzhou. The colonial history placed an emphasis on English which led to a shift towards English-medium schooling (see Evans 2000:187). In 1995, a language policy commonly known as “biliterate and trilingual” (两文三语 liǎngwénsānyǔ) was announced, which aims to “develop a civil service which is ‘biliterate’ in English and Chinese and ‘trilingual’ in English, Cantonese and Putonghua” (Lau 1995:19, cited in Bolton 2012:187).¹⁸ However, despite the promotion of Mandarin Chinese as a medium of instruction in schools (Bolton 2012), Cantonese remains the medium of instruction and strongly entrenched as the language of identity, with 95% of the population speaking Cantonese (Ng and Cavallaro 2019:29–30). English is the official language, and continues to be a language of commerce, higher education, and the government administration and law institutes. Different from Hong Kong, the Singaporean population was linguistically diverse dating back to before the colonial period (see Section 3.2). And the series of language policies and language campaigns since its independence has transformed the linguistic profiles of the Singaporeans. We can also see that the force of English is strong in all three cases, which is probably not just because of the colonial past, but also due to the status of English as a global language in the new economy.

In the process of language shift, smaller minority languages and varieties such as Hokkien, Malayalam and Baba Malay are facing imminent endangerment in Singapore (see Lim 2010; Pillai 2018). Lim (2010) sketches a theoretical timeline of linguistic ecology in terms of the dominance of various languages in different ages in the Singapore speech community (see Figure 3.9). Four stages are differentiated: The first stage refers to the period roughly starting from pre-colonial to the post-independence years (mid-1970s), while the second stage is the age of the official languages corresponding to the independence era from the mid-1970s. The third stage begins around the 1980s and is named “the age of the global media” (Lim 2010:46), and the fourth stage, beginning in the new millennium (2000–), can

¹⁸ Here, “‘biliterate’ in Chinese and English” refers to written Chinese and English and Putonghua refers to Mandarin Chinese (Ng and Cavallaro 2019:29–30).

be seen as the expansion of the third era, where international trades attract talents and investors from all over the world, in particular from China, the Middle East, and India. Though it oversimplifies the prominence of different languages in different historical stages, it provides a good summary of the major players that have been, and still are in competition within the linguistic ecology.

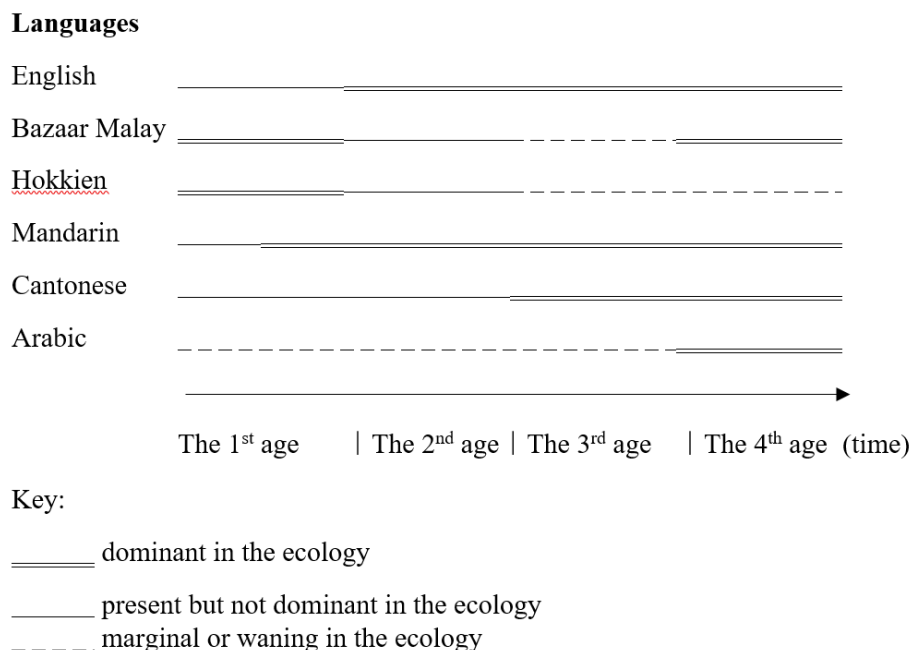


Figure 3.9: Representation of the relative significance of languages in the different ages of linguistic history in Singapore (Lim 2010:46)

According to this outline, Hokkien will only play a marginal and waning role in the linguistic ecology in the future. Lim (2010) is hopeful that languages such as Bazaar Malay and Arabic will revitalize due to reintroduction as third language choices in school for students. Cantonese, on the other hand, as one of the important Chinese vernaculars during the third stage, is predicted to continue its significance towards the fourth stage by Lim (2010) because of the popularity of Cantonese pop cultures and Hong Kong movies. However, Bazaar Malay, Arabic, and Cantonese are in reality, as presented in the above language shift and discussed in Section 3.2, on a continued decrease. And as indicated by the sharp decrease of the Chinese vernaculars as home languages presented in Figure 3.7 and Figure 3.8, this trend shows no indication of being reversed.

3.5 Summary

This chapter has provided an analytical account of the immigration patterns and language policies in different eras of Singapore's history (mainly from the early colonial age to the new millennium), and how they have influenced the linguistic ecology of Singapore. It offers a brief historical background of how Singapore, as a thriving island for maritime trade in the pre-colonial period, became an important colonial port in the 19th century, and gained independence and developed into a vibrantly diverse international city-state in the 20th century. It also highlights the immigration patterns in different eras with immigrants coming from China, Malay, India, and the Arab world, as well as the language use of these different communities. In addition, this chapter has introduced different language policies implemented since the independent era of Singapore and how they, as another important force additional to immigration patterns, supported, if not directly caused, language shift in Singapore. Moreover, all of these are of significance to understand the evolution of CSE in Singapore. Not only languages that were dominant in different eras in Singapore may have contributed to the emergence of CSE, social factors, such as economic values and prestige associated with the substrate languages, sense of ethnic group belongings, as well as sense of national identities, have helped to shape CSE as it is spoken today.

4 Colloquial Singapore English as a contact language

It is uncommon for linguists to use universal words like *everywhere*, *total* or *completely* when they make claims about language phenomena, but when it comes to the phenomenon of language contact itself, using these universal words seems to be a common practice. The German linguist Hugo Schuchardt (1884) declared more than a century ago that there is no completely non-mixed language.¹⁹ Likewise, Thomason (2001:8) argues that “language contact is everywhere: there is no evidence that any languages have developed in total isolation from other languages.” What they both underline is the ubiquity of language contact and the significance of foreign influence on languages.

Despite the omnipresence of language contact, the study of it, namely contact linguistics as a new branch of linguistic subjects, only established itself approximately seventy years ago with the earlier scholarly research dating back to the 1950s (e.g. Haugen 1950, 1953; Weinreich 1953). Since then various studies in language contact – both theoretical and empirical – have been conducted, and contributed to the field of contact linguistics and increased our knowledge as well as our understanding considerably (see among many Thomason and Kaufman 1991; Thomason 2001; Heine and Kuteva 2005; Siemund 2008; Matras 2009). The problem with most of these theories is that they presume that there is a source language and a recipient language while in many contact cases, however, situations are significantly more complex. Usually, in a multilingual society, e.g. a city-state such as Singapore, all languages and cultures present in the society are involved in language contact.

The growing interest in World Englishes has been accompanied by an impressive number of synchronic studies (Huber and Velupillai 2007; Jenkins 2009; Siemund 2013; Davies and Fuchs 2015). At the same time, diachronic investigations of postcolonial Englishes are still the exception. As introduced in Chapter 1, the Dynamic Model of

¹⁹ “Mit mehr Recht als Max Müller gesagt hat: ‘es gibt keine Mischsprache’, werden wir sagen können: ‘es gibt keine völlig ungemischte Sprache’. Wenn überall bei innigem Verkehr verschiedensprachiger Menschengruppen auch die Sprachen aufeinander wirken, so wird umgekehrt da wo eine physische Kreuzung, die ja den allerinnigsten Verkehr voraussetzt, nachgewiesen ist, auch eine Kreuzung der Sprachen sich vermuten lassen.” (Schuchardt, 1884: 5)

Schneider (2007) depicts a unified theoretical account of the development and evolution of postcolonial Englishes around the world. According to Schneider (2007), a diachronic process is shared by all postcolonial Englishes from the transplantation of English in a new land to the stabilization of a newly emerged variety. However, even such a diachronic model is supported by mainly using synchronic data.

Focusing on CSE, past studies on CSE have been trying to capture its emergence in relation to Standard English as well as its relationship with substratum influences (e.g. Platt 1975; Gupta 1989, 2001; Alsagoff 2007; Bao 2005; Leimgruber 2013, 2009). Although they offer interesting insights regarding the grammatical features of CSE, they often fail to recognize the role of individual speakers in influencing the outcomes of CSE. Besides, their theories could often not explain situations where speakers draw on both standard and non-standard features in a single utterance. Therefore, after introducing the previous models of CSE, this section also points out the necessity of new models with stronger explanatory power that put CSE in a wider context of linguistic studies, as well as of a focus on the importance of individual speakers in language change.

This chapter first discusses the theories which lay out the foundation of contact linguistics regarding the principles and mechanisms involved in language contact as well as their interaction, to set the stage for the discussions in the rest of this study. Section 4.1 discusses CSE in relation to the key concepts of language contact theories in general, for instance, the linguistic outcomes of language contact, such as “borrowing” and “interference” (Thomason 2001), “pattern replication” (Matras 2009), and “contact-induced grammaticalization” (Heine and Kuteva 2005). Section 4.2 examines the questions as to in which stage CSE situates in the Dynamic Model of the evolution of postcolonial Englishes (PCEs) (Schneider 2003, 2007). Section 4.3 highlights the complexity of a theoretical approach to CSE and presents various attempts to capture the variation observed in CSE, before Section 4.4 summarizes this chapter and gives a conclusion.

4.1 A contact perspective on CSE

The idea that contact languages are a well-defined linguistic phenomenon is controversial (see Thomason 1997:3). One definition of contact languages is that they are languages of wider communication, i.e. *lingua francas*, including not only pidgins and creoles, but also non-pidgin and non-creole languages (Thomason 2001:158). Under this definition, present-day English is also a contact language, which has borrowed thousands of new words and by some estimates up to 75% of its total vocabulary comes from other languages, mostly from French and Latin over the centuries (Crystal 2003:8). Many scholars (e.g. Bailey and Maroldt 1977; Poussa 1982; Warner 1982) even argue that (Middle) English is a creole, mainly because of the substantial changes between Old and Middle English. From a lexical point of view, English seems to be far more a Romance than a Germanic language (see Crystal 2003). Bailey and Maroldt (1977) claim that 40% of the English lexicon, semantics, phonology, and morphology are mixed. However, many others disagree, as lexical borrowings on their own do not make a creole (Thomason 2001; Thomason and Kaufman 1991). Moreover, Middle English did not form a “creole” overnight, rather it was a gradual process as more foreign words were introduced into the language, but creoles are considered to have developed by non-ordinary and unusual processes (Hock and Joseph 1996).

Studies on ELF (English as a *lingua franca*) generally define English as a contact language (e.g. Firth 1996; Jenkins 2009, 2014; Jenkins et al. 2011). Firth (1996:240) considers ELF as a “contact language between persons who share neither a common native tongue nor a common (national) culture, and for whom English is the chosen foreign language of communication”. Jenkins (2014:2) even extends its application to native English speakers, adding that ELF “refers, in a nutshell, to the world’s most extensive contemporary use of English, in essence, English when it is used as a contact language between people from different first languages (including native English speakers)”. English, as a global *lingua franca* today, continues to be in contact with many languages of the world via international trades and the Internet. Inevitably, it continues to change, taking in new loan words from other languages, and at the same time it is also changing other languages.

The definition of contact language by Thomason (1997, 2001) adopts a different scope. She does not consider (Standard British) English as a contact language and adds that

besides *lingua franca*, bilingual mixed languages are also contact languages, even though they are in-group languages, but not languages of wider communication. She defines contact languages on the basis of the types of historical connections to other languages:

[C]ontact language is any new language that arises in a contact situation. Linguistically, a contact language is identifiable by the fact that its lexicon and grammatical structures cannot all be traced back primarily to the same source language; they are therefore mixed languages in the technical historical linguistic sense: they did not arise primarily through descent with modification from a single earlier language. By definition, therefore, contact languages are not members of any language family and thus belong in no family tree – except perhaps as the ancestor of a language family: a contact language has no single parent language in the historical linguist's usual sense, but it may have descendants. (Thomason 2001:158)

Following Thomason (1997, 2001), there are three types of contact languages, namely (1) pidgins, (2) creoles and (3) bilingual mixed languages (also *bilingual mixtures* or *mix languages*) (Thomason 2001:203–17). Traditionally pidgins and creoles, which are regarded as the best-known contact languages, were paired with and contrasted to each other (Thomason 1997:1, 2001:187). Based on different contact situations, pidgins and creoles can be further classified into (a) pidgins and creoles with European lexicons and (b) pidgins and creoles with non-European lexicons. They are scattered along the routes that were followed by European powers engaged in trade and colonization, starting in the Age of Exploration. Bilingual mixed languages, unlike pidgins and creoles, are created by bilinguals. A scholarly well-known example of bilingual mixed languages is Michif, in which the verbs are based on Cree (a native American language of the Algonquian family) and the nouns come from French, with additional borrowings from English and indigenous languages of the Americas such as Ojibwe and Assiniboine (see Bakker and Papen 1997). It is spoken in North Dakota in the United States, and in the provinces of Saskatchewan, Manitoba and Alberta in Canada (Thomason 2001:11).

It is difficult, however, to locate CSE in the above types of contact languages. On the one hand, CSE resembles creoles in that its lexicon mainly comes from one source language, and it draws features from several languages it has contact with (e.g. several Chinese dialects, Malay, Tamil, etc.). However, the traditional view holds that pidgins and creoles are

generally “untutored” (Platt et al. 1984:2), namely, they are not the results of systematic second language acquisition in school, but originated in the street. CSE is believed to have emerged from the bilingual education system in Singapore (Platt 1975; Platt and Weber 1980; Pakir 2004; Tay 1982; Gupta 1994; Schneider 2003; Deterding 2007). On the other hand, CSE could not be a “mixed language”, if we follow Thomason (2001:197) that mixed languages arise in circumstances where imperfect learning plays no role. Yet, imperfect learning does play an important role in the forming of CSE, especially in the early stage of its formation. A solution will be to consider CSE as a type of New English, or “Postcolonial English”, using Schneider’s (2007) term. Gupta (2010) also suggests that Singapore English is a range of Englishes. This is not uncontroversial, because traditionally, as mentioned above, scholarly circles make a categorical distinction between New Englishes and pidgins and creoles: New Englishes are developed from the educational system but pidgins and creoles emerge from trade and colonization. However, this distinction is no longer upheld, as exemplified by Indian English, whose transmission is generally anything but “untutored” (Lange 2012:67).

More revolutionary proposals (Mufwene 2001; Ansaldo 2004; Ansaldo et al. 2007) suggest that there is no real dividing line between creolization and the forming of New Englishes. While Mufwene (2001) uses the model of language evolution to account for creolization, Ansaldo et al. (2007) apply the linguistic ecology concept of Mufwene (2001) to account for other varieties of English. A recent study of Ansaldo (2009) proposes three perspicuous features of CSE – copula deletion, predicative adjectives, and topic prominence – which are “selection[s] of non-English material from the multilingual pool of variables available in the linguistic ecology” (2009:142).

Ansaldo (2009) believes that it is equally misleading to treat CSE as an exclusively English or an exclusively Chinese variety in terms of language evolution. Following Mufwene (2008), he considers it unnecessary to establish *a priori* that a language cannot cross a certain level of mixedness in order to be classified as genetically related to one of its plausible ancestors. In Ansaldo’s (2009:138) words: “It is important to realize that normal transmission is untutored, creative and involves more than one language in most colonial

settings where AEVs [Asian English varieties] emerge.” CSE is usually considered as a type of English (a new variety of English or an Asian English) based on the recognition of the abundant lexical influence from English as well as the essential grammatical material from English (Pakir 1994; Schneider 2003). However, the features such as copula deletion, topic-comment patterns, reduplication, along with aspectual markers, and sentence-final particles (see Chapter 1.2) seem to suggest CSE is more of a language of the Sinitic type. It seems undisputable though to consider CSE as a contact variety, which is lexified in English but shows grammatical influences from Sinitic, Malay, and other local substrate languages it has had and still has contact with.

Many contact phenomena may occur before the emergence of a contact variety, including language shift (see Section 3.4) and language convergence (see Thomason 2001) at a macro-sociolinguistic level, as well as borrowing (i.e. lexical borrowing, grammatical and phonological borrowing) and code-switching, and shift-induced interference at a linguistic level (see Thomason 2001; Winford 2005). A common feature of borrowing and code-switching is that both the form and the meaning of a construction are introduced to the native or target language. In shift-induced interference, the meanings are often incorporated into the target language by using morphosyntactic materials from that target language.

4.2 Borrowing, copying, transfer, imposition, shift-induced interference or pattern replication?

Contact-induced changes in languages haven been classified into two broad categories: borrowing and shift-induced interference (Thomason and Kaufman 1991; Thomason 2001). These two categories are separated according to Thomason and Kaufman (1991) by a significant factor – imperfect learning. Borrowing refers to the incorporation of foreign elements into the speakers’ native language (L1 or some other primary language they speak fluently) when the process does not involve any effects of imperfect learning (Thomason and Kaufman 1991; Thomason 2001). Shift-induced interference according to Thomason (2001) occurs due to interferences of the speaker’s native language (L1) or some other

primary language on an L2 in the course of second language acquisition (SLA). The process usually involves shift of one speaker group to another group’s language, during which imperfect learning plays a role, and thus, results in changes in the Target Language (TL). Shift-induced interference is also called “substratum interference” in the context of creole formation, where a shifting population is socio-politically subordinate to the people whose language they are shifting to (Thomason and Kaufman 1991). An alternative term “transfer” is often applied in this sense in the context of second language acquisition (SLA) (see e.g. Weinreich 1953; Treffers-Daller 2009).

	Incorporation of foreign elements into the speakers’ native language	Interference of speakers’ native language (L1) or other primary language on an L2
Thomason and Kaufman (1991)	borrowing	shift-induced interference or interference through shift substratum interference
van Coetsem (1988); Winford (2005)	borrowing under recipient language agentivity	imposition under source language agentivity
Ross (1991)	Borrowing I	Imposition I
Johanson (2000, 2002)	code copying: code adoption	code copying: code imposition
Heine and Kuteva (2005)	transfer involving phonetic substance; transfer of form-function units	grammatical replication
Matras (2009)	matter replication	pattern replication
SLA studies (Weinreich 1953:1; Treffers-Daller 2009:58)	borrowing or code-switching	transfer

Table 4.1: Summary of different terminologies for borrowing and interference

Table 4.1 summarizes the existing terminologies based on the classification of borrowing and interference by Thomason and Kaufman (1991). Borrowing is also referred to in other scholarly work as “borrowing under recipient language agentivity” (van Coetsem 1988;

Winford 2005), “copying” (Johanson 2000, 2002), or “matter replication”, while shift-induced interference or substratum interference is also known as “imposition under source language agentivity” (van Coetsem 1988; Winford 2005), “pattern replication” (Matras 2009), or “transfer” in SLA studies.

However, these are overlapping terminologies in the literature and there is by no means any consensus on how one category can be distinguished from another. For example, Ross (1991) distinguishes two other kinds of contact situations in addition to the typical borrowing and shift-induced inference. Typical borrowing – “Borrowing I” in Ross (1991:126) – involves native speakers who import a word from a foreign language, while “Borrowing II” is the situation represented by non-native speakers who have “imported” words from their native language as a marker of their identity into the majority language. A good example of the second situation is Jewish Americans using Yiddish words in English (Haspelmath 2008). On the other hand, shift-induced interference could also happen among speakers who transfer syntactic features from a dominant language to their native language (so-called “Imposition II” in Ross 1991:126), in addition to the typical shift-induced interference mentioned above (“Imposition I”). Heine and Kuteva (2005:6) consider borrowing as “contact-induced transfer involving phonetic substance of some kind or another” (forms or form-meaning units), and distinguish it from “grammatical replication”, that is, transfer of grammatical meanings and functions as well as syntactic relations. This implies that they do not consider the transfer of structural patterns as a kind of borrowing. However, for Thomason (2001), borrowing starts out as lexical, but can also involve phonology, syntax, and morphology when the contact intensity between the two languages in contact increases. Moreover, Mufwene (1990:1) makes a distinction between transfer and substrate influence. In his word, “[t]ransfers apply putatively in the speech of multilingual speakers and/or at the stage of SLA; substrate influence is observed in a language as a relatively crystallized system. Once transfers have been replicated by different speakers, repeated by most of them, and established in the contact situation’s new linguistic system (even as variable features), they may be characterized genetically as substrate influence” (Mufwene 1990:2). This implies that Mufwene (1990) assumes transfer as the cause of

substrate influence. Clearly there is need to reconcile these different views and achieve a more systematic explanation of these terms. In the following sections, I will elaborate Thomason and Kaufman’s (1991) theories on borrowing and interference and present other classificatory explanations of other scholarly work, such as van Coetsem (1988, 2000), Johanson (2000, 2002), Heine and Kuteva (2005), and Matras (2009).

Borrowing vs. interference

Thomason and Kaufman’s (1991) treatment of borrowing and interference is innovative and influential, which aid researchers in describing contact-induced change in a more constructive way. I summarize the different features of borrowing and shift-induced interference described in Thomason and Kaufman (1991), and Thomason (2001) in Table 4.2.

Borrowing	Shift-induced interference
(i) L1 is the recipient language, speakers who borrow elements still retain their L1 (or the language that they speak fluently).	(i) L2 is the recipient language, L1 shifts to L2.
(ii) Imperfect learning is absent.	(ii) Imperfect learning plays a role.
(iii) It starts with the lexicon.	(iii) It starts with phonology and syntax, not with the lexicon.

Table 4.2: Borrowing vs. shift-induced interference

Borrowing is differentiated from interference in that (i) speakers who borrow elements still retain their L1. The features are introduced by the native speakers of the receiving languages; (ii) the borrowing process does not involve imperfect learning; (iii) borrowing starts with lexical items and as contact intensity increases, incursions into the phonology, morphology, and syntax occur as well.

Unlike borrowing, shift-induced interference occurs when (i) an L1 is not retained; (ii) imperfect learning plays a role in language shift which results in changes in the target

language; (iii) interference does not begin with the vocabulary, but with sound and syntax, and sometimes includes morphology as well. Most of the time, interference seems to require a more specific set of sociological conditions and contact intensity to occur than borrowing, although both happen in the case of bilingualism.

While the distinction between borrowing and interference certainly helps researchers in describing contact-induced change, the proposed distinctions are not without practical or theoretical problems. First, the structural boundaries between borrowing and interference are blurred, as pointed out by Myers-Scotton (2002:236). Further, borrowing and interference are not exclusively separated and can occur simultaneously, as stated by Thomason (2001) in the following paragraph:

A possible exception to this generalization might occur if the shifting group is a superstrate, a socioeconomically dominant group, rather than a substrate. The case of Norman French speakers shifting to English in England is the most famous example: hundreds and later thousands of loanwords poured into English as an indirect result of the Norman Conquest, and in fact the structural interference from French in English was much more modest. But that picture is complicated by the fact that English speakers were probably borrowing words from French speakers during the process of shift – that is, both borrowing and shift-induced interference were likely to have been occurring at the time.

Thomason (2001:75)

Secondly, borrowing and interference have been used interchangeably to refer both to the process of contact-induced change and to its results. The problem was even pointed out by Thomason and Kaufman (1991) themselves:

If we know that contact was intimate enough to make shift as well as borrowing possible, then there is no reason to suppose that one process operated to the exclusion of the other, barring established social or numerical asymmetry that would enable us to rule out one of the mechanisms.

Thomason and Kaufman (1991:69)

Moreover, the role of imperfect learning has been overemphasized in the distinction between borrowing and shift-induced interference. There are many exceptions of borrowing with presence of imperfect learning as well as interference without imperfect learning, especially in situations of bilingualism. Matras (2009:147) provides an example that young Kurdish

children in the outskirts of Diyarbakir in eastern Anatolia used to greet Western tourists with the word *okay* in the mid-1980s and later replicated the token in interaction with insiders in their native tongue. The problem also manifests itself in the categorization of borrowing or interference in CSE, since it is often not clear whether imperfect learning plays a role in transplanting a certain item in CSE. For example, it is difficult to tell whether a Mandarin Chinese-English bilingual speaker borrows topic-prominence structure from Mandarin Chinese or imposes such structure on English. Thomason (2001) explains those many exceptions by blurring the concept of imperfect learning:

It is important to keep in mind that imperfect learning in this context does not mean inability to learn, or even lack of sufficient access to the TL to permit full learning: learners must surely decide sometimes, consciously or unconsciously, to use features that are not used by native speakers of the TL.

Thomason (2001:74)

However, bilingual speakers may well be aware of the origin of a word or morpheme in a particular donor language, but this awareness may be blurred over time. Thus, bilingual speakers may borrow foreign words into an utterance as well as shift the meaning or structure of a certain pattern consciously or unconsciously. In such a case, imperfect learning plays no role in separating borrowing from interference.

A clearer classificatory framework to distinguish borrowing and contact-induced interference is offered by van Coetsem (1988, 2000). In his framework, the term transfer is used to refer to any kind of cross-linguistic influence, namely, both borrowing and interference (in Coetsem's (1988) term "imposition"). In both types of transfer, the material is transferred from a source language (SL) to the recipient language (RL). What separates borrowing from imposition is whether the "agent" (namely speaker) of the transfer is the SL speaker or the RL speaker. If the agent is the recipient language speaker, for example an English speaker using Chinese words while speaking English, the transfer of material (including structure) is borrowing (or "recipient language agentivity") (van Coetsem 1988:3). On the other hand, if the source language speaker is the agent, for example a Chinese speaker using Chinese articulatory habits while speaking English, the type of transfer is imposition

(or “source language agentivity”) (see van Coetsem 1998:3). The RL in these cases does not necessarily have to be the first or native language of the speaker, but the linguistically dominant language of the speaker, namely the language in which he is most proficient and most fluent (see van Coetsem 1995:70). This explanation solves the terminological issue of defining L1 and L2 among bilingual speakers in Thomason and Kaufmann (1991).

It is important to distinguish linguistic dominance from social dominance, which is related to the social or political status of a certain language. The socially dominant language is not necessarily the linguistically dominant language of the speaker. For example, an American-born Chinese-English speaker, who is most fluent in Chinese, using English-derived words while speaking Chinese is a case of borrowing (recipient language agentivity). A crucial factor in determining the extent of borrowing is the “stability gradient of language” (Coetsem 1988:25). This corresponds to the scale of borrowability that will be discussed in Section 4.1.2, in which certain linguistic domains, e.g. phonology, morphology (especially inflectional morphemes), syntax and semantics are more stable and thus less susceptible to change than others, e.g. lexicon and functional morphemes.

Johanson (1992, 2000, 2002, 2008) prefers the term “code copying” to borrowing and interference. In his “code-copying framework” first formulated in Johanson (1992), the setting involves contacts of Turkic languages, but in later version, he turned to more general aspects of language contact. For him, “code” stands for any linguistic elements. In other words, all levels of cross-linguistic influence are considered as copying. i.e. code A can copy lexical items, pattern, rule and meaning from code B. The code-copying framework makes a distinction between “code imposition” (copying from L1 to L2) and “code adoption” (copying from L2 to L1), which are more or less equivalent to borrowing and interference, respectively. However, according to Johanson (2008:62), copies are not transferred but have their own development trajectory in the new code, or are subject to code-internal development. They are not identical to that of the model language, even if a copy appears to possess all properties of the original.

There are four types of properties possessed by every linguistic item, including material (the phonetic shape of an item), semantic, combinational, and frequential. The

framework is not constraint-based as any kind of copying can happen, according to Johanson (2008). When all the properties of a linguistic item are copied, this is a case of “global copying” (which corresponds to borrowing or code-switching in other models). When only certain properties are copied, the type of copying is referred to as “selective copying”.

Heine and Kuteva (2005) use the term replication for any forms of transfer from a model language (M) to a replica language (R). They distinguish borrowing and interference from the structural point of view, i.e. borrowing is restricted to (i) the transfer of linguistic forms, i.e. sounds or combinations of sounds and (ii) linguistic form-meaning units or combination of form-meaning units. On the other hand, interference (or replication in their terms) is the asymmetrical transfer of (i) meanings (including grammatical meanings or functions) or combinations of meanings and (ii) syntactic relations, i.e. the order of meaningful elements (Heine and Kuteva 2005:2–3). In their monograph *Language Contact and Grammatical Change*, they mainly discuss the transfer of grammatical meaning (see Section 4.1.3 on contact-induced grammaticalization), which corresponds to the selective copying in Johanson (1992).

Matras (2010) comments that borrowing and copying are strange metaphors as the donor language never expects to receive its words back and the transferred items are not exactly identical to their sources. He therefore also uses the term “replication” for both borrowing and interference and distinguishes “matter replication” from “pattern replication”. The former is compatible with the notion of borrowing from a model language to a replica language, involving both form and function, and the latter is equivalent to what Heine and Kuteva (2005) refer to as replication, that is, the differentiated selection of word-forms and constructions (Matras 2010:235).

In sum, borrowing and interference are used in various ways in different contexts. Two main different approaches can be outlined. The first approach is often adopted in the context of SLA, while borrowing is referred to the cross-linguistic influences from a second language to a previously acquired language (typically one’s native language), interference is usually restricted to the influences of a native language or a previously learned language on the acquisition of a second language. The second approach, also the approach adopted by

more recent literature (e.g. Heine and Kuteva 2005; Matras 2010) on contact languages, uses the term borrowing for the incorporation of both form and function from a source language to a target language, and interference for a change in meaning and function but without involving the form of the original linguistic items. Before further illustrating the differences between borrowing and substratum interference in CSE, with examples found in the selective interviews of OHI in Section 4.1.3, we will take a look at some implicational hierarchies to capture the linguistic constraints on borrowing and interference in the following section.

4.3 Scales of borrowability vs. hierarchies of shift-induced interference

Although a well-established view in language contact studies is that any linguistic feature can be transferred from one language to another (Thomason and Kaufman 1991:14), not all linguistic features are equally likely or frequently subject to cross-linguistic influence. To this end, a number of implicational hierarchies have been proposed to capture which linguistic categories are more susceptible to contact-induced change (Field 2002; Haugen 1950; Matras 2009; Thomason and Kaufman 1991). Generally well-accepted predictions about borrowing is that vocabulary is borrowed before structure, unbound forms before bound forms, lexical items before grammatical items, semantically transparent forms before semantically opaque forms (e.g. Moravcsik 1978; Thomason and Kaufman 1991; Thomason 2001; van Hout and Muysken 1994). On the other hand, it is stated that shift-induced interference starts with phonology and syntax while lexical interference comes at the opposite end of the scale (Thomason and Kaufman 1991; Thomason 2001). However, it is also mentioned that sociolinguistic factors, rather than linguistic constraints are more crucial in predicting its outcomes (Thomason 2001; Thomason and Kaufman 1991:85; Treffers-Daller 1999:1; Siemund 2008:4). In the following sections, we will first take a look at some premises and assumptions on borrowability and then proceed to the hierarchies of shift-induced interference.

4.3.1 Scales of borrowability

The frequently cited scale of borrowability proposed by Thomason and Kaufmann (1991) and Thomason (2001) underlines contact intensity as the most crucial factor. Four different levels of borrowing are outlined in the borrowing scale, namely (i) casual contact, (ii) slightly more intense contact, (iii) more intense contact, and (iv) intense contact, as summarized in Figure 4.1. At the lexical level, non-basic vocabulary – most often nouns, verbs, adjectives, and adverbs, rather than basic vocabulary – are borrowed under casual contact conditions. As the intensity of contact increases, borrowing of function words (e.g. conjunctions and adverbial particles like “then”) occurs, followed by derivational affixes. At the structural level, structure borrowing is not likely to happen in the least intense contact situations, while more significant structural features, such as word order (e.g. SVO beginning to replace SOV or vice versa), the syntax of coordination and subordination, and morphological categories (flexional morphology and agglutinative morphology), are borrowed or replaced at a higher level of contact intensity.

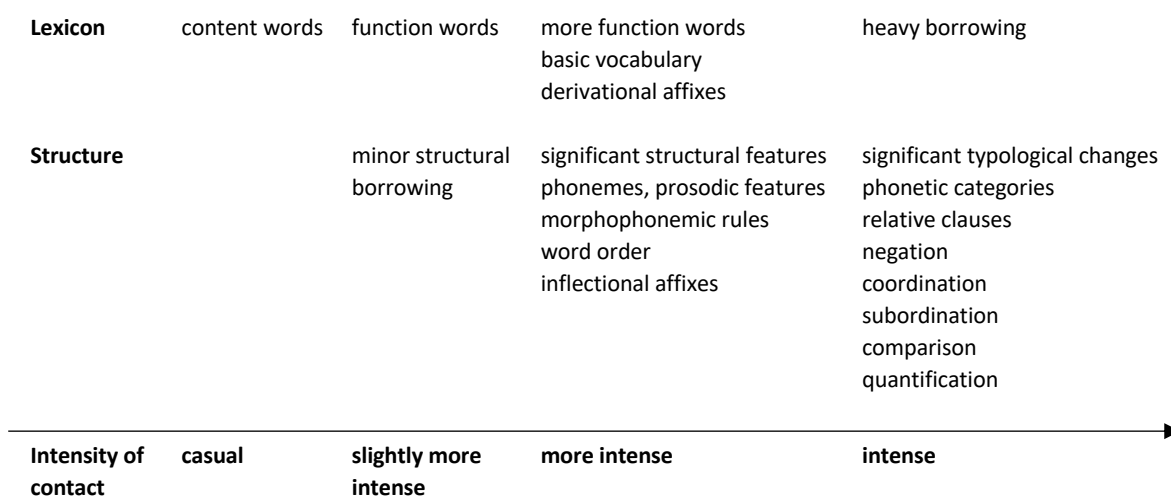


Figure 4.1: Representation of borrowing scale (Thomason 2001:70–71)

Some concepts in the model, e.g. intensity of contact and function words remain unclear though. Labeling the intensity of contact by degrees does not define the degrees, and in concrete language contact situations intensity of contact and cultural pressure are not strictly

linear (Matras 2009:157). Furthermore, the label “function word” is a rather broad concept, covering different categories such as discourse markers and definite articles, the first being extremely prone to be borrowed, the latter rather resistant (Matras 2009:157). Finally, the model offers no explanation as to what promotes lexical borrowing and structural borrowing in situations of different contact intensity. Although basic and non-basic vocabulary as well as intensity of contact are not well-defined categories in Thomason (2001), her borrowing scale provides some interesting insights in contact-induced change.

Besides Thomason (2001), there are several other attempts to capture borrowing hierarchies in language contact. Field (2002) proposes the scale on borrowability in (a) with regard to lexical items.

(a) Content item > function word > agglutinating affix > fusional affix
(Field 2002:38)

The scale can be interpreted in both temporal and quantitative ways. The temporal claim establishes that linguistic items that belong to the categories on the left side of the scale will be borrowed earlier than the one(s) on the right. That is to say that content items will be borrowed earlier than function words, and function words will be borrowed earlier than agglutinating affixes, and in turn, agglutinating affixes will be borrowed earlier than fusional affixes. The quantitative claim argues that language X will borrow from language Y a larger number of linguistic items that belong to the categories on the left side than the one(s) on the right. This claim entails that more content items will be borrowed than function words, more function words than agglutinating affixes, and so on.

The scale is to a large extent consistent with the previously proposed borrowability scales by Thomason (2001) (see above in Figure 4.1) and Whitney (1881) (with respect to borrowability of grammatical elements in (b)) in that content words are borrowed earlier than function words and affixes tend to occupy the end position of a borrowability scale, but Field (2002) adds that agglutinating affixes are borrowed earlier than fusional affixes. Content items are those words or expressions that have concrete and tangible referents, they are more salient and transparent. Therefore, from a cognitive perspective, they are easier to

be learned, and in turn, be borrowed (Field 2002:36). With respect to the ordering of agglutinating affix and fusional affix, Field explains that it is more likely that “segmentable affixes”, i.e. agglutinating affixes with one-to-one correspondences of form and meaning, will be borrowed than fusional affixes, which collapse a number of grammatical categories onto a single, often phonetically minimal form (2002:37).

(b)

Function words	Affixes
Preposition > Conjunctions > Pronouns	Derivational > Inflectional

(representation of borrowability of grammatical elements Whitney 1881:19–22)

Field (2002:35) notes that the scale mirrors the cline of lexicality-grammaticality, or the process of grammaticalization, i.e. a gradual historical development of a grammatical form from content word to inflectional affix (see Heine, Claudi, and Hünemeyer 1991; Hopper and Traugott 2006; Sapir 2008). From left to right in the scale of borrowability of Field (2002), we can also observe a gradual diminution of form, i.e. from stand-alone autonomous word to bound morpheme (or affix). Field (2002) suggests that such similarity between the scale of borrowability and the process of grammaticalization is not a mere coincidence. He claims that “the more structural (or grammaticalized) an element is, the less likely it will be borrowed from one language to another” (2002:35).

An important concept in Field’s (2002) borrowability scale is the Principle of System Compatibility (PSC). As languages have different morphological profiles, it is necessary to take into account the morphological type of the languages in contact (isolating, agglutinative, fusional/flectional, polysynthetic). According to this principle, “any form or form-meaning set is borrowable from a donor language if it conforms to the morphological possibilities of the recipient language with regard to morphological structure” (2002:41). Also presented is its corollary, the Principle of System Incompatibility (PSI): “No form or form-meaning set is borrowable from a donor language if it does not conform to the morphological possibilities of the recipient language with regard to morpheme types” (2002:42). These principles are based on scales of synthesis (the number of grammatical concepts expressed in a morphologically complex word) and fusion (the degree to which two or more of these

grammatical concepts are merged into one form). If the recipient language is fusional-synthetic, say for example Spanish,²⁰ it can borrow any foreign element, including free forms such as independent words and bound forms such as agglutinating and fusional affixes. If the recipient language is agglutinating-synthetic, e.g. Japanese,²¹ it can borrow free forms and agglutinating affixes, but rarely fusional affixes. However, if the recipient language is isolating-analytic, for instance, a variety of Chinese, it can borrow only free form, yet most bound forms are not borrowable. In other words, only morphologically compatible form-meaning sets can be borrowed into a language of a certain type, while non-compatible form-meaning sets are not borrowable in principle.

Field's (2002) empirical research was based on a corpus of 23,272 words of oral interviews collected from 1974 to 1984 by Kenneth and Jane Hill in the ethnographically distinct highland Malinche region of central Mexico. The language in focus in the study there is Malinche Mexicano (agglutinative), a variety of Modern Mexicano, which is said to have borrowed 60% of its lexical material in various registers from Spanish (see Field 2002:123–64). He examines Spanish borrowings in the portion of the corpus and compares the numbers of borrowings according to the categorization of different types of words in the borrowability scale. The result shows that the borrowing patterns of Malinche Mexicano accord with the borrowability scale (see Table 4.3).

Category	Number of types (tokens)
Content items (N, V, Adj, Adv)	767 (3,431)
Function words (particles)	46 (3,221)
Agglutinating-type inflectional affixes	1 (164)
Fusional-type inflectional affixes	0 (0)

Table 4.3: The occurrence of Spanish form-meaning sets in Mexicano (from Field 2002:142)

²⁰ Spanish is a good representative of fusional-synthetic languages, as Spanish verbs are nearly always inflected. Verbal suffixes in Spanish can provide multiple information, such as number, case and tense. For example, the Spanish verb *comer* ('to eat') has the first-person singular preterit tense form *comí* 'I ate'; the single suffix *-í* represents both the features of first-person singular agreement and preterit tense, instead of there being a separate affix for each feature (see Payne 2006:17).

²¹ Note that the agglutinative and fusional languages are two ends of a continuum, with various languages falling more toward one or the other end. Japanese is generally agglutinative, i.e. it generally has one grammatical category per affix, but with a few exceptions displaying fusion (see Narrog, Rhee, and Whitman 2018:180).

The number of the types of form-meaning sets that are borrowed is listed in the right column in the above table, while the number of their tokens is recorded in parenthesis. The information presented in Table 4.3 provides support for the observation that “the more structural (or grammaticalized) an element is, the less likely it is to be borrowed” (Field 2002:142). Intriguingly, the number of tokens of the function words, e.g. prepositions and conjunctions, that are borrowed is extraordinarily high (almost 1:1 with content items). He concludes that this reflects the complete integration of Spanish function words into Mexicano. Nevertheless, he expects the number of potential content items to grow as they belong to the parts of speech that readily accept new members.

Among scholarly circle, there is a consensus that nouns, verbs and adjectives are content items (see Haugen 1950; Whitney 1881; van Hout and Muysken 1994). It is disputed, however, whether adverbs should be classified as content words. The main reason is that there are various types of adverbs that serve a wide range of functions. Manner adverbs, i.e. adverbs that modify the head of predicate phrases (verbs), generally form an open class (Hengeveld 1992:71). An open class is a word class that commonly accept the addition of new words. It is a concept similar, though not identical to content words, which are lexical categories in the stricter sense (Carnie 2012:51). Other adverbs that modify larger constituents such as clauses and sentences, e.g. time and place adverbs, belong to function words or “closed classes”, which are less likely to take new members (Hengeveld 1992:71). Therefore, some scales (e.g. Field 2002; Onysko 2007) of borrowability categorize adverbs as content items, some identify them as function words (e.g. Haugen 1950, see (d)), while others distinguish manner adverbs (e.g. van Hout and Muysken 1994, see (e)) from sentence adverbs.

Apart from borrowability scales that generalize content items and grammatical borrowing (either function words or grammatical affixes), other proposals of borrowability scales focus on determining which part of speech is more likely to be borrowed. Again, Whitney’s (1881) is among the earliest proposals. According to Whitney (1881), nouns are

most easily borrowed, followed by adjectives, verbs, and adverbs, pronouns and others are collapsed in the category of other parts of speech.

(c) nouns > adjectives > verbs > other parts of speech (Whitney 1881:19–20)

Since then, other observations have been made on borrowability scales based on part of speech, which are more or less similar to Whitney's proposal. Haugen (1950), for example, elaborated on Whitney's scale based on the list of American English borrowings in Norwegian and Swedish:

(d) nouns > verbs > adjectives > adverbs > prepositions > interjections
(Haugen 1950:224)

Both (c) and (d) agree that nouns are more borrowable than other parts of speech. However, they disagree on the ordering with respect to verbs and adjectives. Similarly, van Hout and Muysken (1994) in (e) and Onysko (2007) in (f) provide different proposals which differ from each other in the position of the second grammatical category that follows nouns.

(e) nouns > verbs > adjectives > sentence adverbs > quantifiers > conjunctions > prepositions > interjections > negation > manner adverbs > greetings
(van Hout and Muysken 1994:42)

(f) nouns > adjectives > verbs > adverbs > others (Onysko 2007:131)

The position of adjectives or verbs in the scale of borrowability may be specific to certain structural properties of the languages. For example, some languages do not have adjectives, e.g. Cree and other Algonquian languages. Therefore, other parts of speech such as verbs or relative clauses are used as attributive modifiers of nouns, which results in an increase of other categories rather than adjectives (Nga 2017:45). A similar observation is made by Hekking and Bakker (1999, cited in Haspelmath 2008:50): Quechua, compared with Otomí, borrows more adjectives from Spanish, which may be due to the fact that Otomí lacks adjectives. Regarding verbs, Meillet (1921, cited in Thomason and Kaufman 1991:348) mentioned that it is difficult to borrow verbs from other languages into French as French has an elaborate

inflectional system. Others (e.g. Moravcsik 1975; Wichmann and Wohlgemuth 2008) note that it is more likely for a language to borrow nouns first and then to employ its own system of denominal verbalization than to borrow verbs directly.

A more recent proposal of Matras (2009) suggests a frequency-based hierarchy of different borrowed types based on a sample of 27 languages in contact from different parts of the world, as listed in (g). The scale gives a fine-grained chain including the different subcategories of function words (e.g. discourse markers, interjections, and particles) and affixes (e.g. derivational affixes and inflectional affixes).

- (g) nouns, conjunctions > verbs > discourse markers > adjectives > interjections > adverbs > other particles, adpositions > numerals > pronouns > derivational affixes > inflectional affixes (Matras 2009:157)

The frequency here does not refer to the quantity of tokens or types of borrowings in a specific language, but to the number of languages that show cases of borrowings in the relevant categories. Matras (2009) observes that while all languages in his sample display borrowing of nouns and conjunctions, the borrowing of inflectional affixes, however, is only found in the smallest number of languages, and pronouns are attested to be the least frequently borrowed word class. Yet, we can see that the scale of Matras (2007) is not identical to the previously proposed hierarchies. What is remarkably different is that conjunctions and discourse markers are placed on the highest position while content items like verbs and adjectives are scaled down in Matras (2007), yet van Hout and Muysken (1994) and Haugen (1950) all state that coordinating conjunctions and discourse markers (i.e. interjections and greetings in their scales) are less likely to be borrowed.

The expressions *already*, *also*, *ever*, and *one* in CSE, as observed in this study, are free grammatical elements, i.e. function words, which occupy mid-positions in the previously discussed borrowability scales. They are different from their counterparts in native Englishes, which occupy transitional positions between content items and function words: *Already*, *also*, and *ever* in standard varieties of English are used as adverbs, while *one* could serve different functions such as numeral and pronominal (see Chapter 5.4).

However, these expressions in CSE do not correspond to the profiles of borrowings, which are form-meaning sets obtained by a recipient language from a donor language. The contact-induced change in these expressions do not involve copying of forms from other languages, i.e. they retain the forms in English, but there is a shift in meaning and distribution, inspired by a different language or different languages. In addition, the formation of CSE is akin to the process of shift-induced interference (see Section 4.4). The hierarchies of shift-induced interference will be discussed in the following section.

4.3.2 Implicational hierarchies of shift-induced interference

Unlike borrowing, whose general directions of progression is commonly accepted,²² it is difficult to theorize a general hierarchy of shift-induced interference. Matras commented: “The initiation and successful propagation of new, replicated constructions [...] seems to take on a much more erratic course” (2007:243). Shift-induced interference, or pattern replication in Matras’ (2007:243) term, is in general a much more “volatile” and “opportunistic” strategy of speakers than borrowing. According to Matras (2007), successful pattern replications are those experiments of speakers’ innovative usages that are well-accepted by the interlocutors. As second-language learners of unbalanced bilinguals may not always succeed in conveying the intended meaning, i.e. sometimes their cross-linguistic selections of form-meaning sets (i.e. meanings originating from a source language that are mapped onto an already existing form in the target language) may sound strange or alien to the interlocutors. Therefore, a successful selection of a new construction depends to a large extent on the reactions of the interlocutors. If it is accepted by the interlocutors, and the communication is effective, it will be more likely to be used by the speaker again. Consequently, it will be replicated by others in the speech community and eventually result in language changes (Matras 2007:243–244). Romaine (1995:51) also identifies that part of the problem in discussing interference is dealing with it at both individual level and

²² e.g. content items before grammatical affixes, unbound morphemes before bound morphemes, and lexical borrowing before structural borrowing (see the section above).

communal level. While interference at the level of the individual may be sporadic and idiosyncratic, the effects of it at a communal level can be cumulative. Only those cumulative effects over time can lead to new norms. Thomason (2001) says that linguistic constraints on contact phenomena easily fail as both the direction and extent of linguistic interference is socially determined.

Nevertheless, there are several hierarchies of shift-induced interference proposed by scholars. A well-known implicational hierarchy is formulated in Thomason and Kaufman (1991:37) and Thomason (2001:75). They predict that shift-induced interference goes almost the opposite direction of borrowing: it does not start with lexical items. Instead, it starts with phonology and syntax, and may eventually include morphology:

(h) phonology > syntax > lexicon > morphology (Thomason 2001:75)

As presented in Table 4.4, the asymmetries between the mutual influences of the Germanic and the Romance varieties found in Treffers-Daller (1999) confirm Thomason's (2001) implicational hierarchy. In Treffer-Daller's (1999) study, the contrastive effect of borrowing and shift-interference is manifested in Brussel and Strasbourg: While lexical borrowings are mainly visible from the Romance varieties to the Germanic varieties, phonological and syntactical influences primarily originate from the Germanic varieties.

Direction of contact	Germanic > Romance	Romance > Germanic
Type of phenomenon	shift-induced change	borrowing
Phonology	strong	weak
Syntax	moderate-strong	weak
Lexicon (including derivational morphology)	moderate	very strong
Inflectional morphology	weak	weak

Table 4.4: Analysis of language contact in Brussels and Strasbourg (adapted from Treffer-Daller 1999:3)

The hypothesis of Treffer-Daller (1999) is twofold. First of all, she assumes that the language contact phenomena found in Brussels Dutch and in Alsatian as spoken in Strasbourg belong to the borrowing process, as defined by Thomason and Kaufman (1991). On the other hand,

the Germanic influences exerted on the French varieties found in these cities are affiliated with shift-induced interference. Secondly, aligned with the proposed direction of shift-induced interference by Thomason and Kaufman (1991), she predicts that shift-induced interference on the level of phonology and syntax will be more prominent in the French varieties, rated as strong and moderate-strong, respectively. Along the scale, interferences on lexical and morphological level in French are rated from moderate to weak. With regard to borrowing, while lexical influence from French is important in the Germanic varieties (Brussels Dutch and Alsatian as spoken in Strasbourg), lexical influence from the Germanic varieties in the French varieties is far less significant.

Many researchers believe that phonological interference is the most pervasive. Weinreich (1953), for example, proposes that the greater the differences of the phonemic systems between two languages are, the greater are the learning problems and the potential areas of interference.²³ He attributes interference to the mismatch of a phoneme of the secondary system with one in the primary system. When a bilingual reproduces a phoneme in the secondary system, he adapts it to the phonetic rules of the primary language (Weinreich 1953:14). He compares the sound systems of two languages in the Domleschg valley, canton of Grisons, namely, Schwyzertütsch (or Swiss German) and Romansh (also spelled Romantsch). One example he gives is that speakers of Romansh replicate their system of vowel length into Schwyzertütsch, where it is not necessary. He names such phenomena as over-differentiation, which refers to the result of imposition of phonological distinctions made in one language on sounds in the second one. Other examples that belong to over-differentiation are given by Romaine (1995:53) and Haugen (1953). Romaine (1995:53) claims that Italian-English bilinguals might pronounce double consonants in English words, e.g. *Patty*, as /patti/ according to the rules of Italian. Haugen (1953) studies Norwegian-English bilinguals in the United States and reports that they often substitute /z/ with /s/ in English due to the lack of the phoneme /z/ in Norwegian.

²³ Weinreich (1953) uses the term interference to denote both borrowing and shift-induced interference. Lexical borrowing is included in the section of lexical interference (see Weinreich 1953:56).

In addition to segmental phonology, more recent studies have documented changes in supra-segmental phonology such as stress, intonation, and tone due to contact (Burridge 2006; Hickey 1990, 2012 on prosodic traces of Irish in Irish English; Matisoff 1973 on Asian tonology). Mackey notes, “[o]f all phonological features, intonation is often the most persistent in interference and the most subtle in influence” (1962:48). A common reason adduced for the permeability of intonation to contact-induced change is that it is tightly bound to context-specific or discourse-related phenomena (Queen 2001). With respect to Asian Englishes, Ansaldo and Lim (2012) show that Asian Englishes are tonally far more complex, having added a number of suprasegmental features normally not found in standard varieties. For example, Singapore English and Hong Kong English show selected and restricted use of lexical tone in specific functional domains, i.e. discourse particles, or at word and phrase level but without contrastive meaning (Ansaldo and Lim 2012:202).

Besides phonology, syntax is regarded as a field that easily subjects to shift-interference. Heine (2008), for example, points out how the syntax of a language can be deeply influenced by contact. Matras (2009:244) suggests that interference begins with those that organize complex structures. He expects the structures of various types of clauses (e.g. complement clauses, adverbial clauses, and relative clauses), as well as the structure of coordination to be targeted first in the process of convergence. This is followed by phrases and words (see (i)). Here, convergence is understood as the enhancement of inherent structural similarities found between two linguistic systems (see Matras 2009:236). What is worth mentioning is that he uses the term convergence for both process (“progression of convergence” (2009:244)) and result (“convergence in the use of suffixes” (2009:250)).

(i) clause > phrase > word (Matras 2009:244)

To illustrate how the syntactic structure of one language can be remodeled on another in language contact, we can consider the following examples (13)–(15) of word order interference in CSE found in OHI. Instead of following the default SVO (subject-verb-object) word order of English or the subject-predicate structure (see Ziegeler 2021), the following sentences in CSE show resemblance to the preferred topic-prominent structure in Chinese.

- (13) And it's there that you mix with all sorts of people, coolies... Then I knew about management. Well, *management already I have*. [OHI-000526-LKS]
'I already had the knowledge of management.'
- (14) A: Did you bring up to the union about that?
B: *This one, long time we brought up to the union already*. But sometimes we think also-lah. If you want to survive, if your mind is still active, you must go for re-training. [OHI-001953-LAS]
'We already brought up this (topic) to the union a long time ago.'
- (15) That time the qualifications is not very high. I think *secondary school you can get in already*. [OHI-002206-MH]
'I think that you can get in secondary school.'

We can see that the direct objects *management* in (13), *this one* in (14) and *secondary school* (15) are fronted and become the topic. Alsagoff and Ho (1998:136) also found examples of topicalized structure along with a pro-dropped subject, e.g. *This book Ø sell already* (the pronoun *I* is dropped). Example (14) illustrates that almost any constituent can be fronted as a topic (Bao and Lye 2005). In this case uttered by Speaker B, the adverbial *long time* is fronted but placed after the direct object. In all cases, we see that the subject is not important enough to have the initial position and thus gives way to other constituents, i.e. objects and adverbials. As pointed by Bao and Lye (2005), topic-prominent structure of CSE is modeled directly on that of Chinese.

Taking an excursion to German influence in English, we can also find some typical examples of shift-induced interference in syntax. Wieden and Nemser (1991:360) provide examples produced by German-speaking Austrian students in their L2 English, as shown in (16)–(18). Interestingly, the argument structure of the verb parallels that of German rather than English.

- (16) Explain me something. (modeled on German: Erklär mir was.)
- (17) You just finished to eat. (modeled on German: Du hast gerade aufgehört zu essen.)
- (18) I would suggest him to go. (modeled on German: Ich empfehle ihm zu gehen.)

Curiously, Matras (2009) proposes that discourse should be placed even earlier than the position of syntax in an implicational hierarchy. The complete cline of interference of Matras (2009) is set forth as in (j). The reason he gives is that “bilinguals face strong pressures in copying with distinct procedures of organizing and managing the discourse and the arrangement of propositions in discourse. On this basis, we might expect the pressure to converge the inventory of constructions in the repertoire to begin with those that organize complex propositions” (Matras 2009:244).

(j) discourse > clause > phrase > word (Matras 2009:244)

To illustrate what is referred to as discourse, Matras (2009:27–28) gives the example of a four-year-old Hebrew-German bilingual child replicating the German second-person polite form *Sie* construction by using a Hebrew possessive construction in the third-person plural form. During a role-play as a customer, he addressed his father as a grocer:

- (19) a. Hebrew:
 yeš lahem tapuxím?
 there.is to.3PL apples
 [Intended meaning]: ‘Do you have apples?’
 [Actual meaning]: ‘Do they have apples?’
- b. German:
 Haben Sie Äpfel?
 have.3pl you.polite/3PL apples
 ‘Do you have apples?’

The German second-person polite form *Sie* shares the same form as its 3PL pronoun *sie*,²⁴ and has the same 3PL conjugation on the verb. At this stage, the child’s knowledge about the German politeness form is limited to this particular context *Haben Sie X*. In other words, he does not acquire the German politeness marker as such, but a construction linked to the speech activities in the context of grocery shopping (in a role play with his parents). Hebrew, on the other hand, lacks a politeness pronoun. Therefore, the child applies a Hebrew possessive construction in the 3PL, which is modeled on German.

Another example Matras (2009:28) gives to further explain what is meant by discourse organization is shown in (20). The Hebrew-German bilingual child, now seven years old, cried out loudly while watching a football match broadcast (in English) on television:

- (20) a. Child: *Penalty shot!*
 b. Father: me ha-nekudá ha-levaná [...]
 from the-spot the-white.SG.F
 ‘From the white spot [...]’
 c. Child: ze ma še *penalty shot.*
 that what REL *penalty shot*
 ‘That’s what a penalty shot is.’

Here in segment (c), the child uses a cleft construction which is not idiomatic in Hebrew. According to Matras (2009:28), the Hebrew cleft construction has the form {this + what +

²⁴ The slight difference between them is that the former is written with a capitalized S and the latter with an s in lowercase.

COMP + verb} as in *ze ma še aní amárti* (word for word literally ‘that what that I said’). This type of cleft construction in Hebrew is not compatible with present-tense existential predications. The child applies the structural features of Hebrew cleft constructions with an extension of the semantic function of the English left construction “that’s what a penalty shot is”. Matras (2009:28) explains that the interference is motivated by the need to employ a newly acquired semantic-pragmatic construction as an effective solution to an on-the-spot, immediate, local communication.

Some other scholars, e.g. Stolz and Stolz (1996) and Ross (2001) also point out the vulnerability of the discourse-organizational domain to language convergence in contact situations. An interesting case of this kind of interference is the use of discourse markers. Ross (2001:151) gives an example of the adoption of the discourse marker *aria* ‘all right’ to mark a change in the topic of conversation in Takia and Waskia, as well as in many languages along the north coast of Papua New Guinea. In addition, he observes that the same languages have adopted Tok Pisin conjunctions. He assumes that this kind of interference happens very early in the “metatypic” process as it makes no structural demands on the speaker.²⁵ Stolz and Stolz (1996:97) also mentioned the integration of Spanish discourse markers, e.g. *entonces* ‘(and) then’ and *pues* ‘after that’ into quite a variety of indigenous languages on both sides of the Pacific (e.g. Nahuatl, Zuctec, Zapotec, Totonac, etc.).

Such examples are also found in CSE, with respect to the use of various discourse markers (or pragmatic particles) from Sinitic languages. This is a well-known and much-documented feature of CSE (see Gupta 1992b; Ler 2006; Lim 2007; Wong 2005). In the following examples found in the database of OHI, the pragmatic particles serve various functions originating in the Sinitic languages Hokkien, Cantonese and Mandarin. While *ah*, *meh* and *leh* usually serve an interrogative or exclamatory function as shown in (21)–(23), *lah* and *lor* are used in declaratives to express persuasiveness and obviousness as shown in

²⁵ “Metatypic” is the adjective form of *metatypy* coined by Ross (1996), which is a process “whereby the language of a group of bi- or multilingual speakers is restructured on the model of a language they use to communicate with people outside their group. The process involves primarily (i) the reorganization of the language’s semantic patterns and “ways of saying things”; (ii) the restructuring of the patterns of linguistic materials i.e. sentences and clauses, phrases as well as words (Ross:2001:146).

(24)–(25) (see Gupta 1992b; Wee 2004a; Ler 2006 for a more comprehensive description of their functions).

- (21) a. They listen, *ah?* [OHI-003206-MA]
 ‘Did they listen?’ (interrogative)
 b. He was a young man that time *ah!* Ah, very young. Very smart.
 [OHI-003206-MA]
 ‘What a young man he was at that time! He was very young and very smart.’ (exclamatory)
- (22) What’s these coughs! No smoking *meh?* [OHI-001632-CHN]
 ‘How bad these coughs are. Is it true that you didn’t smoke?’ (interrogative)
- (23) My type of English maybe acceptable *leh?* [OHI-002206-MH]
 ‘Is my type of English acceptable?’ (tentative interrogative)
- (24) a. Of course the rope is very nice *lah*, it’s not just plain rope.
 [OHI-001970-NJK]
 ‘Of course this is a very nice rope, it is not just plain rope. (obviousness)
 b. Oh Ah Chek, give me *lah*, free. I got no money today. [OHI-002102-IT]
 ‘Oh Ah Chek, please give me this for free. I don’t have any money with me today.’ (persuasiveness)
- (25) If it’s raining then we have it in the school hall *lor*. [OHI-002827-EC]
 ‘If it rains, we will have it in the school hall.’ (obviousness)

Functionally comparable categories exist in English, e.g. *so, well, right, isn’t*, etc. (see e.g. Lenk 1998; Lewis 2014, 2018; Schweinberger 2012). However, unlike these expressions in English, pragmatic particles in CSE are monomorphemic units not bonded to any syntactic elements, and they occur in sentence-final or phrase-final position, mirroring that of their counterparts in Chinese vernaculars (Gupta 1992b; Lim 2007).

Some scholars (e.g. Lim 2007; Wee 2004a) in CSE also consider the use of these pragmatic particles as borrowing as they are form-meaning pairs incorporated in CSE. The sounds of the pragmatic particles in CSE matches that of its counterparts in Chinese. Phonologists in CSE even observe that the pragmatic particles in CSE are tone carriers, i.e. they carry with them the tonal systems found in their substrate languages (see Platt and Ho 1989; Lim 2007).

However, the crucial question for CSE is whether its formation is akin to the process of borrowing or to that of shift-induced interference. According to Thomason, the genesis of pidgin and creole is affiliated with shift-induced interference (Thomason 2001:285). If one agrees with Platt (1975) that CSE is akin to creole (i.e. in his term, CSE is a “creoloid”), then the emergence of CSE is associated with, though not identical, to shift-induced interference. Therefore, we could assume that changes of CSE follow the implicational hierarchy of shift-induced interference, that is, by introducing interference features starting with phonology, followed by syntax level, and then morphology and eventually the lexicon.

Let’s take a look at a few examples of lexical interference. Lexical items from Chinese, Malay and Tamil have been incorporated into CSE. For example, the word *ang moh* from Hokkien literally translates as ‘red hair’ and is used in CSE to describe Caucasians. This occurs in our selections of OHI 54 times. Four examples are listed in (26):

- (26) a. I think the Vietnam War also end. There was less *ang moh* already, less soldiers come in so the business also slowing down. And I think the government also want a different focus. [OHI-0002951-JY]
- b. Then like if the *ang moh* wins, the *ang moh* will get don’t know how many, how much... And these people always win the *ang moh*. But I think during those days, all these Caucasians also like. [OHI-002951-JY]
- c. Because we were, no doubt we was in a very Chinese [environment] but we were every influenced by the *ang moh* culture. Because during those times they were really, their technology everything is really more superior than us. [OHI-002951-JY]
- d. The *ang moh* doctors were all gone? All gone except for a few like Professor Ransom who remained. He was interned or what and then he came back. [OHI-002951-JY]

Leaving aside the cultural connotations of the term *ang moh*,²⁶ the information we can withdraw from the above examples is that, first of all, this word does not come from English, it sounds foreign, and second, as (26)b exemplified, the speaker clearly knows the alternative

²⁶ The term *ang moh* started as a racist and derogatory term in the 16th and 17th century, but is adopted today as a neutral term (The Straits Times 2007).

English word *Caucasians*. Additionally, it is quite clear that the noun is perfectly integrated into the English morphosyntax, even creating compounds such as *ang moh culture* in (26)c and *ang moh doctors* in (26)d. What is worth mentioning is that the tones originally carried by the Hokkien counterpart were omitted in the process, which is a common linguistic phenomenon called “phonemic assimilation” (Winford 2005). Clearly, this case represents Borrowing II (Ross 1991), where Singaporean speakers adopted the Hokkien word 红毛 *âng-mô* (their native language) into English (the majority language). It is different from the traditional sense of code-switching, in which a lexical item is not supposed to be syntactically or phonetically integrated into a recipient language (Treffers-Daller 1991:262).

The uses of *already* as a perfective aspect marker, *also* as an additive marker, *ever* as an experiential perfective aspect marker, and *one* in CSE do not belong to this type of lexical interference, as they replicate the grammatical meaning/functions from their substrate language(s), but not the form. We will discuss such phenomenon in more detail in Chapter 5.

4.4 Contact-induced Grammaticalization

Grammaticalization theory is concerned with the emergence and development of grammatical forms and constructions. According to Hopper and Traugott (2006:18), 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”. A good example is observed in the Chinese 过 *guò* ‘to pass’, which has experienced a grammaticalization process from a regular verb that means ‘to pass a place’, to an aspectual marker which emphasizes the experience associated with a completed event (Shi 2002:138) (see Section 6.3.2 for the grammaticalization of 过 *guò*).

Grammaticalization has typically been considered as a “language-internal” process. However, new findings from a number of fields, particularly in areal typology and creole studies, have suggested that grammaticalization could result from “external” factors such as geographical clustering and substrate influence (Ansaldo 1999; Heine and Kuteva 2003,

2005; Matthews and Yip 2009; Ziegeler 2014). Hopper and Traugott (2006:230) also acknowledge that “contact has been an important factor for most languages, and a strictly monogenetic view of grammaticalization is ultimately inappropriate”, emphasizing the important link between language contact and grammaticalization.

In Section 4.2 and Section 4.3, we discussed that language contact may lead to replication of material and functional properties of some linguistic units from one language into another language. Analyzing data from a wide range of languages, Heine and Kuteva (2003, 2005) propose that language contact may also lead to replication of the process of grammaticalization. This process is widely referred to as “contact-induced grammaticalization” (Heine and Kuteva 2005:13–21). The focus of Heine and Kuteva (2005) is on grammatical replication or the transfer of meaning/functions from one language to another, but not of form. They suggest that speakers of the “replica language” can identify grammaticalization processes that are assumed to have taken place in the “model language” and use available lexical material in the replica language to replicate that grammaticalization process (Heine and Kuteva 2005:13–21).

Heine and Kuteva (2003, 2005) argue that internal and contact-induced grammaticalization should not be regarded as mutually exclusive, and that contact-induced grammaticalization is not different in essence from internally motivated grammaticalization. In other words, the principles of grammaticalization are the same regardless of whether or not language contact is involved. The insight gained from viewing interference of grammatical meaning from the perspective of grammaticalization is that contact-induced change is not arbitrary, but tends to follow certain predictable pathways (Matras 2009:240). Similarly, Matthews and Yip (2009:373) view contact as a catalyst driving change along pathways of grammaticalization, and suggest that contact-induced grammaticalization is subject to the general constraints on grammaticalization. In their words: “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).

4.4.1 Ordinary contact-induced grammaticalization vs. replica grammaticalization

Depending on whether the developmental route taken in the replica language is the same as in the model language, Heine and Kuteva (2005:81–122) distinguish two types of contact-induced grammaticalization. If there exists no readily available pathway in the model language, the process is referred to as “ordinary contact-induced grammaticalization” whereas the other type is known as “replica grammaticalization”. To be more specific, ordinary contact-induced grammaticalization refers to changes brought about in the replica language by the existence of a category in the model language but proceeding in a way unrelated to the model language. This mechanism is outlined as involving four steps:

Ordinary contact-induced grammaticalization:

- (i) Speakers notice that in language **M** [the model language or source language] there is a grammatical category **Mx**;
- (ii) They create an equivalent category **Rx** in language **R** [the replica language or target language] on the basis of the use patterns available in **R**;
- (iii) To this end, they draw on universal strategies of grammaticalization, using construction **Ry** in order to develop **Rx**;
- (iv) They grammaticalize **Ry** to **Rx**.

(Heine and Kuteva 2005:81)

To put it another way, ordinary contact-induced grammaticalization is triggered by the need to replicate a grammatical function that exists in the model language. Speakers of the model language would expect the same grammatical distinctions in the replica language when they acquire the latter as a second language. Consequently, they would search for equivalents in the replica language to categories in the model language with which they are already familiar. Therefore, they map that grammatical function of the model language onto an existing lexeme in the replica language, resulting in its functional extension along a path of grammaticalization (Matras 2009:239). This process, according to Hickey (2010:155), is “an

unconscious one and persists even with speakers who have considerable target language proficiency”.

One possible candidate of ordinary contact-induced grammaticalization is the grammatical development of CSE *already* as a perfective marker (one focus of this study). It is observed that *already* occurs in CSE after uninflected verbs to mark perfective aspect without involving any inflectional morpheme (see e.g. Bao 2005; Platt and Weber 1980). This is consistent with the use of the perfective aspect markers *liau* in Hokkien in (27)b (Platt and Weber 1980; Ansaldo 2009), *jo2* in Cantonese in (27)c (Kwan-Terry 1989; Matthews and Yip 2009) and *le* in Mandarin in (27)d (Bao 2005). In this case, CSE speakers grammaticalize a lexical item which is referred to as a “phasal polarity expression” (van der Auwera 1998; see Section 5.1), to an aspect marker instead of using inflectional morphology to mark perfective aspect, following the aspect-marking strategy of the substrate languages.

- (27) a. CSE
I eat *already*.
- b. Hokkien Chinese
Gua tsiah png liau
1SG eat meal PFV
'I have eaten.'
- c. Cantonese Chinese
Ngo5 sik6 jo2 faan6
1SG eat PFV meal
'I have eaten.'
- d. Mandarin Chinese
Wǒ chī fàn le
1SG eat meal PFV
'I have eaten.'

The second type of contact-induced grammaticalization is replica grammaticalization. In this case, the process of grammaticalization in the replica language is not only triggered by the contact language, but also by the assumption of speakers of the replica language that a grammaticalization process has taken place in the model language (Heine and Kuteva 2003:539, 2005:92). Rather than drawing on universal strategies of grammaticalization (see

step (iii) above), speakers use an analogical formula, i.e. speakers replicate the original grammaticalization path in the model language, which is conceptually available to them in the replica language. Heine and Kuteva (2005) note, “knowledge of the cross-linguistic patterns of grammaticalization can be useful to reconstruct certain patterns of grammatical change, in that unusual patterns of grammaticalization found in neighboring languages are likely to be suggestive of language change induced by language contact” (2005:25). Replica grammaticalization is also described as a four-step process:

Replica grammaticalization:

- (i) Speakers notice that in language **M** there is a grammatical category **Mx**;
- (ii) They create an equivalent category **Rx** in language **R**, using material available in **R**;
- (iii) To this end, they replicate a grammaticalization process they assume to have taken place in language **M**, using an analogical formula of the kind [**My** > **Mx**]: [**Ry** > **Rx**];
- (iv) They grammaticalize **Ry** to **Rx**.

(Heine and Kuteva 2003:539, 2005:92)

Heine and Kuteva (2005:93) mention the example that the grammaticalization of the possessive verb *have* (= **Ry**) in CSE denotes existence (= **Rx**): *there have* ‘there is’. The grammaticalization is assumed to be based on a grammaticalization process from a possession verb as in 我有 *wǒ yǒu* ‘I have’ to an existential marker (= **My** > **Mx**) 那有 *nà yǒu* (lit.) ‘there have’, i.e. ‘there is’ that has taken place in Chinese.

One of the major research goals of this study is to find out whether the grammatical development of the aspectual markers *already* and *ever*, the additive marker *also*, and the emphatic marker *one* in CSE may be affected by the grammaticalization tendencies of Chinese.

4.4.2 Replica grammaticalization as recapitulation

Based on the concept of replica grammaticalization, Ziegeler (2014) reveals that there may exist another major group of linguistic settings that needs to be taken into account in a comprehensive study of grammatical replication, where earlier diachronic stages of a lexifier language can be recapitulated in the replicated grammaticalization process in a new variety of English (see the retentionist hypothesis in Pietsch 2009a; and the model of replica grammaticalization as recapitulation in Ziegeler 2014). Taking Singapore English as a case, Ziegeler (2014, 2016) demonstrated that some special uses in CSE (e.g. the predominance of habitual aspectual uses of the modal verb *will*) match with a similar predominance of habitual and/or generic uses of the predecessor of *will* in Old English texts. Therefore, she suggests a model of “replication by recapitulation”, proposing that some novel uses in CSE may have recapitulated the same pathways of grammaticalization observed in the lexifier model – (British) English. Replication in such cases is assisted by the identification of co-existing, lexical source meanings recoverable from the grammaticalized meanings in the lexifier (Ziegeler 2016:311).

One interesting case includes the accounts of the origins of the sentence-final particle *what* in CSE (Kuteva et al. 2018). Previous studies largely converge on interpreting the variant as a result of interference from the local Chinese substrates (e.g. Cantonese, Hokkien, Te Chew, see Smith 1985; Lim 2007). The sentence-final particle in CSE is referred to as “objection particle” (Kuteva et al. 2018:33). The primary function of it is to indicate that the speaker objects to something in the context:

- (28) CSE (Context: Discussion of a student who is going overseas for one month and missing classes) (Smith 1985:126)
A: He’ll never pass the third year.
B: It’s only for one month *what*.

In (28), B uses *what* to refute what is assumed in the context of A’s comment – that the student they are talking about will never pass the third year. Kuteva et al. (2018:33) proposes that this “weird” use of *what* is a result of a grammaticalization process which started in

colonial British English. To be more specific, they argue that the objection particle *what* is the fourth (final) stage of the grammaticalization of *what* which can be traced back to the interrogative use of *what* (stage 1) in the non-local, lexifier language, British English. The development of *what* is schematically presented in the following figure:²⁷

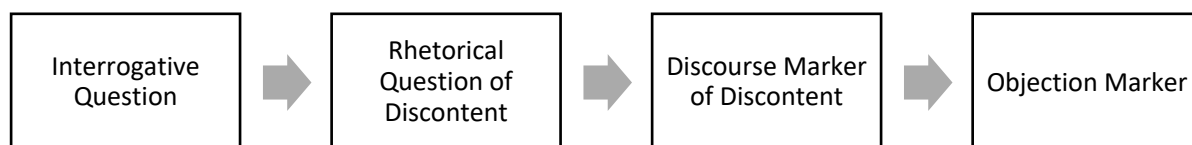


Figure 4.2: Schematic grammaticalization process of CSE objection particle *what* (adapted from Kuteva et al. 2018:33)

These functions of interrogative question, rhetorical question of discontent, discourse marker of discontent can be identified in the following examples from (29) to (32). The stage of interrogative *what* can be exemplified by a *wh*-question in (29):

- (29) Stage 1: Interrogative question
English: *What* do you think?

At this stage, *what* is used in clause-initial position to seek information. Kuteva et al. (2018:46) assumes that such use of *what* is universal since most languages are found to have interrogative clauses with the equivalent of a *wh*-word.

At the second stage, some interrogative clauses start to be used as rhetorical questions. Neither do these rhetorical questions intend to seek information, nor do they require an answer (Herring 1991; Rhee 2004). One of the strategies of rhetorical questions is to express discontent,²⁸ as exemplified in (30):

²⁷ Other pathways were also suggested in Kuteva et al. (2018), such as (i) interrogative question > rhetorical question of mitigation > discourse marker of mitigation and (ii) interrogative question > rhetorical question of agreement > discourse marker of agreement. However, they consider the one in Figure 4.2 as the most likely candidate for the grammaticalization process that gives rise to the objection marker in CSE.

²⁸ Other strategies listed in Kuteva et al. (2018:46) are: mitigating an assertion, expressing (or appealing for) agreement/solidarity. Only the strategy of expressing discontent is regarded as relevant for the grammaticalization of *what* as an objection marker.

- (30) Stage 2: Rhetorical question of discontent
British English (Kuteva 2018:53)
Context: Son was driving drunk and crashed into a tree.
Father: *What* the hell were you thinking?!

By the end of the 18th century, *what* had moved away from its original interrogative meaning in the direction of emotive meaning. Another significant change is that it started to be used sentence-finally. It is assumed that the syntactic change involves a step from the use of *what* as an independent expression that follows a preceding utterance to “a concatenation of that expression with a free-standing *what* (sometimes preceded also by [the discourse marker] *eh*)” (Kuteva et al. 2018:41–42). The following examples found in Kuteva et al. (2018) are texts found at the end of the 18th century, which shows that the function of sentence-final *what* as a marker of discontent in British English predates the formation of CSE.

- (31) Stage 3: Discourse marker of discontent
- a. British English (Lennox 1756, cited in Kuteva et al. 2018:42)
I have lost you then a second Time [sic] my dear Adelaida (cried I) and I have lost you for ever, *what*!
- b. British English (MacNally 1781, cited in Kuteva et al. 2018:43)
This is undoubtedly a predetermined scheme to affront magistracy, said my Lord Praetro—bring the rascals before me—*what*!

In (31)a, the speaker who produced the utterance grieves for Adelaida. The use of the sentence-final *what* emphasizes his grief intertwined with discontent with the situation in which the speaker finds himself. It is ambiguous though, whether *what* serves here as an emphatic marker or a discourse marker of discontent. Another example in (31)b from the end of the 18th century also demonstrates an emotional overload with strong discontent with the context situation. Kuteva et al. (2018) note that the functions of discontent and emphasis of sentence-final *what* are still to be recognized in the British English variety after the mid-19th century – this is the time when the British English variety is contemporaneous with the emerging CSE variety. However, such usages lost their popularity towards the beginning of the 20th century and eventually became archaic in the 21st century in British English.

On the other hand, CSE, as an emerging variety, seems to recapitulate these earlier but now archaic functions of *what* in British English. Curiously, it further developed such a grammaticalization tendency, reaching the most advanced stage of this development: objection marker. Besides the example listed earlier in (28), another example in (32) demonstrates such function. Here speaker B objects the presupposition that travelling to Singapore is expensive by using sentence-final *what*.

- (32) Stage 4: Objection marker
 CSE (Kuteva 2018:33)
 Context: Discussing the topic of A's brother travelling to Singapore.
 A: He probably hasn't got much money.
 B: But going to Singapore doesn't cost very much *what!*

However, the CSE markers²⁹ discussed in the related studies are more closely related to their local substrates. For example, the objection marker *what*, also spelled as *wat* (see Wee 2004b:1069), is more closely related to the Cantonese sentence-final pragmatic particle 𠵹 (pronounced as wo3 or wo5) both phonologically and semantically. Phonologically, Platt et al. (1989:219) point out that archaic British English *what* carries a rising intonation whereas CSE *what* has a lower pitch than the previous syllable. In Cantonese, the particle 𠵹 is pronounced as [wo3] or [wo5] (both at a low pitch) depending on the context. Semantically, the function of the Cantonese particle 𠵹 is (i) to remind or contradict the listener and to strengthen the current assertion from the previous one raised by the listener (as shown in (33)); or (ii) to mark reported speech, meaning 'accordingly' or 'I heard' (see (34)).³⁰ Both functions are relevant to the discussion in Wee (2004:1069) that *wat* in CSE can (i) carry the force of a contradiction to something that has previously been asserted and (ii) present a piece of information as being obvious.

²⁹ Other cases in the discussion of the model of replica grammaticalization include the progressive aspect, the habitual *will* and the use of *one* as a relative pronoun.

³⁰ 𠵹 translated from online Cantonese Dictionary:
<http://www.cantonese.asia/attachments/oncc/lessons/027.htm>

- (33) 係 啲! 唔 啱 啲!
hai6 wo3 ng4 ngaam1 wo3
yes PTC Not correct PTC
'Yes. This is not correct.'
- (34) 佢 話 係 噉 啲, 你 咪 信 佢 咯!
keoi5 waa6 hai6 gam2 wo5 nei5 miu1 seon3 kieo5 lok6
he said be this PTC you just believe him PTC
'He said it is not like this, you can just believe/trust him.'

In conclusion, the analysis of CSE markers recapitulating earlier stages of grammaticalization found in British English is tempting, yet it lacks systematic sampling and explanation to the crucial question as to what the mechanisms are that drive such recapitulation (e.g. motivations of CSE speakers recapitulating earlier functions in British English or contact scenarios between CSE speakers and British English speakers prior to the emergence of CSE). Obviously, more careful diachronic studies need to be conducted to find out the answers to such a question. The recapitulation of earlier stages of grammaticalization is not the focus of this study, as we examine more closely the relevance of CSE markers and their grammaticalization to the Chinese substrates (see the discussions in Chapter 9). Before reviewing previous CSE models on its formation as well as its complex relationship with Standard Singapore English and the local substrates, the following section takes a look at how social factors are involved in language contact studies, as they are the key factors to predict the outcomes of contact-induced grammatical change.

4.5 Social predictors vs. linguistic predictors of contact-induced change

Social factors, such as intensity of contact, number of speakers in the respective linguistic groups, socioeconomic status of the groups involved, imperfect learning and its absence, as well as speaker's attitudes are determinants in predicting the outcome of contact-induced change (see Thomason 2001:21,60; Thomason and Kaufman 1991:85). With "intensity of contact" Thomason (2001:66) refers to the duration of two communities with different languages involved staying in contact, and the degree of intensity of the social and linguistic contact between the groups. Broadly speaking, the higher the degree of bi- or

multilingualism in a community and the closer the contact, the stronger contact effects will be. As mentioned earlier, casual contact may result in light borrowing of content words, while intense contact may lead to morphosyntactic changes, and, eventually, may cause creolization or the emergence of mixed languages (see Section 4.3.1).

These social factors are considered more important than the linguistic factors, because they can override linguistic factors, pushing change in an opposite direction (Mufwene 2001; Siemund 2008; Thomason 2001). Although a lot of linguistic constraints on contact-induced change have been proposed, such as claims about unborrowable linguistic features, un-transferable syntax, universals of grammatical interference (see Harris and Campbell 1995, also see Section 4.3 on scales of borrowability and implicational hierarchies of interference), there are always counterexamples where these proposals fail. Contact linguists, who are interested in linguistic mechanisms, especially implicational hierarchies, are “required to filter out the noise produced by these social factors” (Siemund 2008:4). However, this is almost an impossible mission as there are no two snowflakes alike, and so do language contact situations differ in the world in terms of their social constellations.

Indeed, since the loci of language contact are individual speakers, it will be impossible to study language contact without considering speakers as the agents of language change, as well as the environment where they acquire the language and the network of communication in which they interact. Thomason (2001:61) also attributes part of the unpredictability of language change to speakers’ attitudes, labeling such factors as the “wild card”, as they violate most of the generally valid predictions about contact-induced change. These social factors are identified in Mufwene (2001:192) as the “external ecology” of the contact situation, which come together with internal ecology (i.e. structural characteristics) in the minds of the speakers, whose language use is simultaneously shaped by both (Mufwene 2001:195).

However, this is not to say that linguistic factors are not important. The linguistic factors play a prominent role in predicting the outcome of language contact, which include (i) universal markedness, (ii) the degree to which features are integrated into the linguistic system (see Section 4.3.1), and (iii) the typological distance between source and recipient

languages (Siemund 2008:4–6; Thomason 2001). For example, the general observation is that marked features are less likely to be learned by a shifting group or second language learners, and, therefore, are less likely to appear in the newly emerged varieties. In other words, an expectable result of contact-induced language change is a reduction of markedness. The lack of the third-person singular and the plural suffix in CSE (both in the form of *-s*) are two cases in point. The third point above suggests that typologically closely related languages are more likely to influence one another, while more distantly related languages are less susceptible to contact effects. Yet there are cases of interference between even the most divergent of systems (Burridge 2006:89). It is because that the linguistic factors are always intricately intertwined with social dynamics, which affects the direction and the extent of the interference.

The theoretical distinction between social factors and language internal factors helps to unravel and predict language changes in a contact situation. Therefore, in the case of the variables under study, this means that it is important to analyze the competition between substrate-influenced expressions and their functions as well as the extralinguistic factors.

4.6 Dynamic Model of the evolution of postcolonial Englishes

It is not until recently that models in research and theorizing on postcolonial Englishes have moved beyond synchronic cross-linguistic comparisons to identify historical developments (e.g. Schneider 2007; Mufwene 2008; Ziegeler 2016; Kuteva et al. 2018; Heine and Kuteva 2005; Pietsch 2009a). The Dynamic Model of Schneider (2007) is arguably one of the most promising and well-described theoretical frameworks to account for the evolution of English varieties in postcolonial contexts. It depicts a unified theoretical account of the development and evolution of postcolonial Englishes around the world. According to Schneider, a diachronic process is shared by all postcolonial Englishes from the transplantation of English to a new land to the renewed stabilization of a newly emerged variety. The diachronic process includes five stages based on changes in the identity constructions of the British settlers (STL) and the indigenous (IDG) population. These five stages (Schneider 2007) are:

(1) **foundation**: English is transplanted in a new land; (2) **exonormative stabilization**: Standard codes continue to serve as the norm with an increased number of English speakers from both the settler and the indigenous strands; (3) **nativization**: The new English variety diverges from its ancestral variety as a consequence of the prolonged contact between English and the indigenous languages; (4) **endonormative stabilization**: The innovative linguistic features of the local norms become increasingly accepted; and (5) **differentiation**: The local variety emerges.

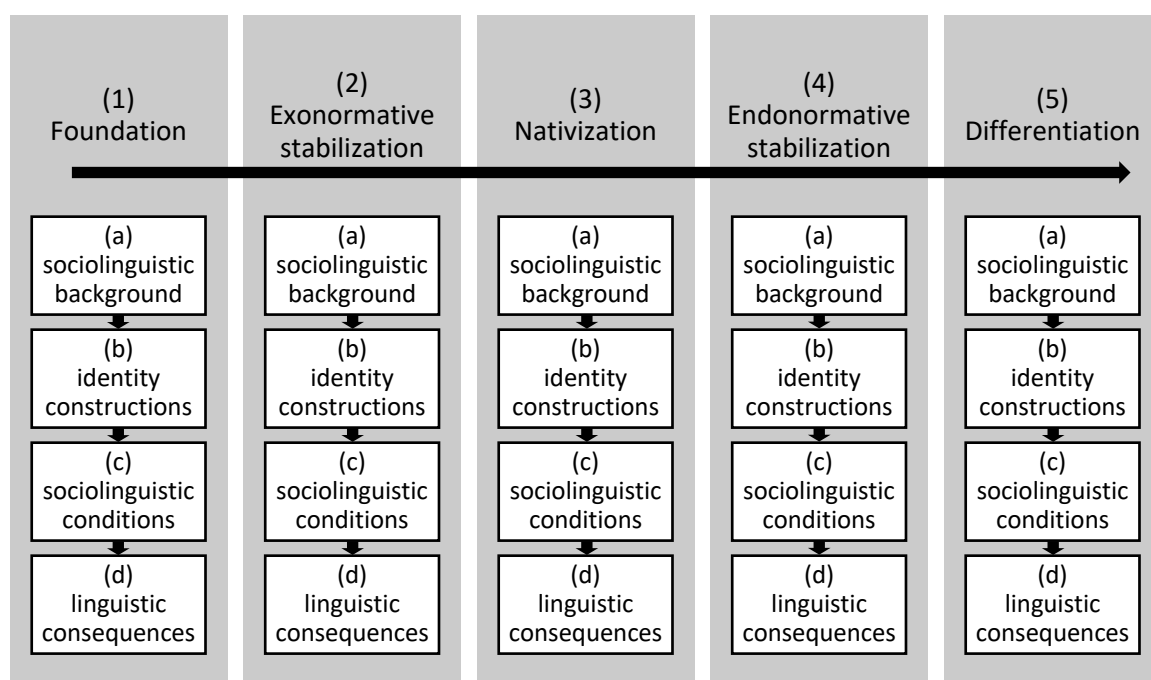


Figure 4.3: Five stages of Schneider's (2007) Dynamic Model

Each of the above five stages is characterized by four core parameters with a monidirectional causal relationship operating between them. These include: (a) sociolinguistic background (i.e. extralinguistic or socio-political background); (b) identity constructions; (c) sociolinguistic conditions (i.e. contact settings and participant's use of specific varieties; norm orientations and attitudes); and (d) linguistic consequences (e.g. structural changes on the levels of lexis, pronunciation, and grammar). We can see that the first three parameters all concern social factors, and only the last one is about linguistic outcomes. These parameters are not independent, but each has an effect on the next parameter. For example, the first parameter – socio-political background – shapes the identity constructions

(parameter b) of the two main parties STL and IDG involved in the process. In turn, these identities determine the sociolinguistic conditions (parameter c). In other words, dependent on their identity construction, speakers will decide on their use of specific varieties (British English or Singapore English), norm orientation (standard or non-standard), and attitudes towards different varieties. In the end, the linguistic consequences, such as structural changes on the levels of lexis, pronunciation, and grammar (parameter d) are the outcomes of the contact settings, language use, and language attitudes.

The first stage sets the foundation of the transplantation of English by the STL strand to a foreign land through the establishment of trading outposts and settlement emigration (see Schneider 2007:33). During the second stage, the colonial status stabilizes, resulting in the establishment of English as the language of administration, law, (higher) education, etc. Though the settlers start to accept the original norm and language contact expands, bilingualism exists mainly in the elite group of the IDG strand, looking towards the British norm which sets the standard of English usages. Lexical borrowing, especially terms of *fauna* and *flora* as well as cultural terms, can be observed. Moving on to stage 3, nativization, “a new identity reflecting the current state of affairs, combining the old and the new is in full swing” (Schneider 2007:40). This stage is often accompanied with political independence on the one hand, and with a close bond of cultural and psychological association with the mother country on the other hand. A good example of this stage is the form of the “Commonwealth of Nations” useful to both sides for maintaining mutual relationship. Linguistically, the newly emerging variety of English is characterized by heavy lexical borrowing, phonological innovations, and structural nativization. In this process, speakers function more as “language builders” (Heine and Kuteva 2005:35) than passive recipients of linguistic forms drawn from the input varieties (Schneider 2007:45). Innovative linguistic expressions are also used as markers of new national identity. It is noteworthy that high frequency lexical and grammatical innovations are likely to become firmly rooted, thus resulting in grammatical nativization.

Stage 4 – endonormative stabilization – is characterized by the change of orientation towards a local norm and the acceptance of a local model of English rather than looking

towards the old, i.e. British, norm codification by means of dictionaries and literary creativity (Schneider 2007:49). In this stage, a new language variety stabilizes due to growing acceptance of local norms and its role as an identity carrier, as well as positive attitudes towards it (Schneider 2007:49). The final stage – differentiation – is marked by a newly emerged language variety, accompanied with dialect birth (Schneider 2007:53–54). This stage is characterized by the acknowledgement of group specific identity in terms of ethnicity, regionality, and society (Schneider 2007:53). As members of the new nation see themselves as belonging to smaller locally defined communities, dense networks between members of these subgroups enforce group-specific norms of language use which in turn symbolize group identity. In other words, new varieties of the formerly new variety start to emerge, and these internal varieties (e.g. sociolects, ethnolects, or regiolects) represent group-specific identity.

The Dynamic Model has been the focus of considerable scholarly interest and has experienced slight modifications since its introduction in 2003 (see Evans 2009; Buschfeld 2013; Schneider 2014). However, one potential limitation of Schneider's model is his overreliance on secondary sources and expert opinion (e.g. Lim 2004 on Singapore English; , and mainly Bolton 2000a, 2000b on Hong Kong English). It has yet to be validated by evidence from detailed empirical research.

4.6.1 The Dynamic Model and Singapore English

Schneider (2007) suggests that the major Inner Circle varieties, e.g. American English, New Zealand English, Australian English (see Kachru 1986) have all reached the final stage, whereas the Outer Circle varieties, e.g. Singapore English, Hong Kong English (see Kachru 1985) have typically arrived at either stage 3 or 4.

According to Schneider (2007), Singapore English entered stage 1 in 1819, when Sir Stamford Raffles obtained the rights from the Sultan of Johor to establish a trading post in Singapore (2007:33;153). The year of 1867 marked the beginning of stage 2 when Singapore, as a part of the Straits Settlements, became a Crown Colony (Schneider 2007:153).

Singapore English entered the critical nativization stage in the late 1940s, marked by the return to colonial tradition after the Japanese occupation, although some traces of stage 2 are apparently still observable (Schneider 2007:157). Schneider (2007:160–161) observes that Singapore English has clearly manifested many characteristics of stage 4 by now, e.g. general acceptance of CSE as a code of Singaporean identity, use of CSE in creative literature, increasing similarity of Singaporean English as spoken by people of different ethnic backgrounds. He notes that it will very likely advance to stage 5 differentiation with its linguistic dynamics. Similarly, some scholars (e.g. Lim 2004; Low and Brown 2005) suggest that Singapore English is already in stage 5, as evident in the ethnic varieties within Singapore English and the codification of it, e.g. the compilation of systematic Singapore English dictionaries and comprehensive scholarly analysis of distinctive features of Singapore English (Schneider 2007:161).

It will be fascinating to see if such observations can be objectified by the diachronic data, as diachronic data, in Schneider's (2007:139) view, "constitutes an interesting test case for the predictive implications of the Dynamic Model and the inherent power of the developmental dynamism which it describes". The present study, therefore, attempts to capture the diachronic pathway(s) of CSE based on an analysis of primary data that covers the history of the presence of CSE in Singapore.

4.7 Models of CSE

4.7.1 The continuum hypothesis

Most of the earliest models of CSE discuss CSE in relation to Standard English. By the late 1970s to 1980s, there was a growing recognition that these features were natural results of the evolution of English in a multilingual setting. According to one of the earliest proposals in this period, the linguistic situation can be captured in terms of a post-creole continuum (Platt 1975). Within this proposal, the phenomena of lexical borrowing, phonological deviation, and syntactic manifestations in CSE were considered to be the outcomes of the development of English-medium education during the British colonial period. As put by

Platt (1978:30–31): “It developed in English medium schools by the transference of linguistic features, concepts and strategies from Southern Chinese Dialects, and to some extent Indian languages and Malay into a type of British English which was the variety taught.” As many children used this variety with their peers on the school playground, with their siblings and other relatives outside the classroom, and later on in the employment domain, CSE was developed in the local sociolinguistic landscape. Moreover, different varieties of CSE developed functional use, and were chosen for different situations, domains, and interlocutors, e.g. formal vs. less formal, with a government officer vs. with a friend, customer in a shop vs. a fellow sales assistant (Platt 1978:31).

According to Platt (1975), variation in English is constructed as a speech continuum and divided into basilectal, mesolectal, and acrolectal forms, ranging from most creole-like and localized to most standard-like and internationally intelligible, such as British Standard English (Gupta 1998). The speech continuum is subject to social stratification: speakers at the higher end of the social continuum have a wider range of available lects, while those at the lower end have a more restricted range of linguistic choices (see Figure 4.4).

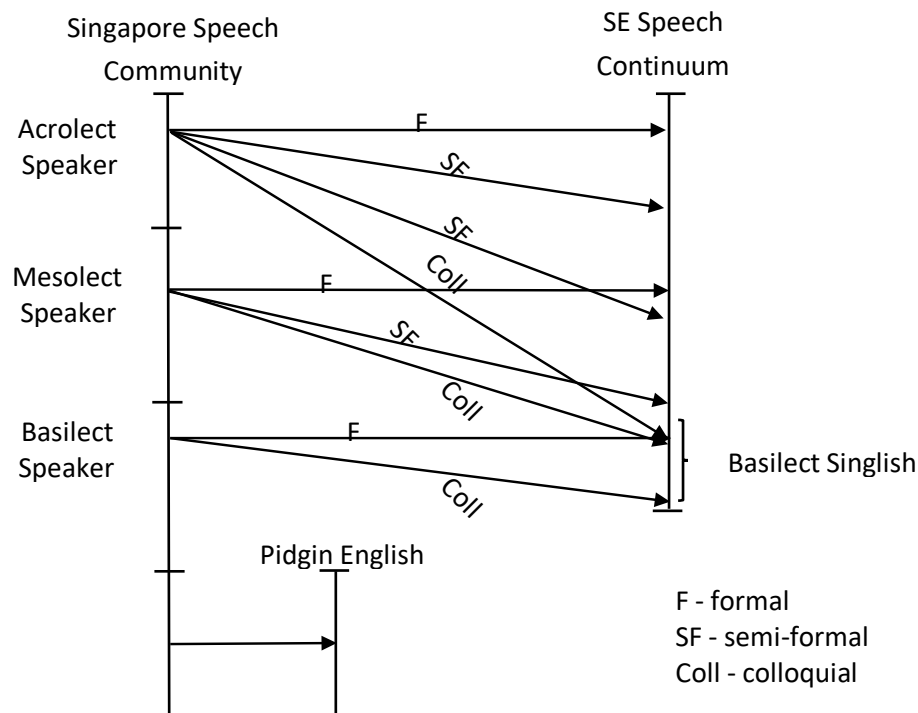


Figure 4.4: Relation between socio-economic factors and the range of sub-varieties of SE available to a speaker (adapted from Platt 1975:369)

Example (35) illustrates three different utterances expressing the same meaning according to their corresponding speech lects (Siemund and Li 2017:11).

- (35) a. Dis new phone si beh expensive leh. (basilect)
b. Dis new phone very expensive one. (mesolect)
c. This new phone is very expensive. (acrolect)

As the concept of a post-creole continuum suggests, the basilectal and mesolectal forms would move and evolve towards the more standard English norms. On the other hand, acrolect speakers can adjust to different situational settings. However, CSE has turned out to be surprisingly stable. Besides, not all speakers of Singapore English are proficient in the basilect forms as more recent studies report that English-educated Singaporeans struggle to understand CSE (Hussain 2006). Therefore, the theory is regarded as outdated, and we need a model that captures the co-existence of standard and colloquial forms (Leimgruber 2009).

4.7.2 Diglossia

Another prominent model to represent the diversity regarding CSE more adequately is *diglossia* (Gupta 1989, 2001). Applying Ferguson's use of diglossia, Gupta appoints the **(H) variety** to Standard Singapore English (SSE), a learned variety through formal education used in written and formal contexts, while the **(L) variety** refers to the local variety acquired natively and the normal code for communication in the community of CSE (Gupta 1994:7). Her observation is that CSE and SSE serve different functions, and that Singaporean speakers – as long as they are able to command either variety – choose to use them in different contexts. For example, it seems reasonable to assume that CSE is used as a home language and amongst friends in informal contexts, whereas SSE is used in more official contexts like administration, business, and education. Speakers can switch between (H) and (L) in the same context to achieve highlighted effects, such as irony, or when speaking to children (Gupta 1998).

The main issue with this model is that CSE is not a clearly defined norm and that, conversely, speakers tend to enrich Standard Singapore English with all sorts of elements

taken from CSE. Besides, the model cannot explain code-switching between (H) and (L) within single utterances or short stretches of discourse.

4.7.3 Cultural orientation and indexicality

Across speakers, considerable flexibility can be observed in the use of CSE elements and it stands to reason that these elements are contextually used for specific functions (Siemund and Li 2017:17). Such observations underlie the culture orientation model developed by Alsagoff (2007, 2010) and the model of social indexing as expounded in Leimgruber (2009, 2013). Alsagoff (2007) argues that the two roles of English, namely, as a global language and as an inter-ethnic *lingua franca*, are representative of and closely related to Singaporean macro-cultural perspectives and identity. The variation of English in Singapore is a reflection of the conflict between “being global” and “being local” (Alsagoff 2007:34). According to her model, Singaporean speakers can, consciously or unconsciously, use elements of CSE to signal their social positions and attitudes. Adopting CSE elements is associated with a local, intimate, and non-Western position, and, conversely, their avoidance is to be interpreted as more Western, international, and formal. Some of the more representative features of the socio-cultural values and practices related to the two poles are listed in the table below.

	International Singapore English Globalism	Local Singapore English Localism
a	Economic capital	Socio-cultural capital
b	Authority	Camaraderie
c	Formality	Informality
d	Distance	Closeness
e	Educational attainment	Community membership

Table 4.4: Features of the two orientations in the cultural orientation model (Alsagoff 2007:39)

Largely based on Alsagoff’s cultural orientation model, Leimgruber (2013) proposes an indexical field of CSE. Centered in the model are the two poles of cultural orientation: local

and global. The key words orientated towards localness are *closeness*, *friendly*, *informal*, *uneducated* whereas stances such as *educated*, *serious*, *distance*, and *formal* characterize the orientation towards globalness. Besides these typical features that can be easily categorized, stances such as *mocking* and *pretentious* fall somewhere between the two orientations (see Figure 4.5)



Figure 4.5: Indexical field of SE. Black = cultural orientation, grey = stances (Leimgruber 2013:244)

The selection of a code, either local or global, follows a three-step procedure within the model of indexicality (Leimgruber 2009:161). The speaker first identifies the formality of the conversation and decides the stances he/she wishes to take. As a sequential step, the speaker chooses the suitable feature(s) associated with the stance that he chooses and finally produces the utterance with the selected feature(s). As an example, Leimgruber (2013:105) mentions the use of discourse markers such as *lah* and *ah* (see (21)–(25) in Section 4.3.2). These features are chosen because the speaker identifies a stance of community membership and reckons localness as an appropriate index in the context.

In Leimgruber’s (2009; 2013) model of indexicality in Singapore English, he argues that there is no identifiable matrix language, i.e. Singaporean speakers create their utterances drawing from a pool of features (Leimgruber 2009:161). He adopts the diglossic term from Gupta (1989, 2009), with (H) and (L) variants referring to the grammatical features of SSE and CSE respectively. For example, the presence of a verbal inflexion is considered as a (H)

variant, while the absence of it is seen as a (L) variant (Leimgruber 2009:202–204).³¹ Interestingly, (H) and (L) variants can co-occur in single utterances or short stretches of discourse. Example (36) below contains the sequence L–H–L–H–H–L of grammatical features (L and H are marked by straight and wavy lines respectively), while example (37) contains H–L–H. Example (37) shows that CSE speakers can use two different codes (*hor fun*³² and *rice noodles*) to describe the same object. Copula *be* dropping (a feature in the (L)-code) occurs in group conversations with local peers while the (H)-code (the inflection on *means*) was applied when he turned to the interviewer to bring the concept *hor fun* out of its local meaning to reach the outsider's understanding (a more international context). This means that social indexing must be seen as a matter of degree.

- (36) No ah, that one is only for accommodation. But actually, if you go there, right, my brother say for just going there is actually five hundred bucks ah. (Leimgruber 2009:232)
- (37) We can eat *hor fun* there, I heard that the *hor fun* ___ quite famous. [to microphone] er *hor fun* means rice noodles. (Leimgruber 2009:217)

However, the theory is not without its limitations. First of all, as Leimgruber (2013:133) points out, while arguably solving the conundrum of synchronic variation in Singapore English, especially variation among different individual speakers as well as mix of (H)-codes and (L)-codes in a single utterance, indexicality is probably not suited to explain the changes in progress in the English spoken in Singapore. Trudgill (2004:156–57) remarks that linguistic expressions of identity are much more likely to be a result of language change (and language contact in the case of Singapore English) than its cause. Questions concerning how these features entered the feature pools of CSE in the first place, and whether their meanings and functions at the idiolectal and the populational level have changed over the past fifty to one hundred years since the formation of CSE, remain unanswered.

³¹ Other criteria include (i) inversion in interrogatives (absence of inversion L vs. presence of inversion H); (ii) the use of modal auxiliaries (except for *can*, other modals are H), (iii) copula *be* deletion (L), and (iv) the use of pragmatic particles (L) (Leimgruber 2009:202–208).

³² *Hor fun* (河粉 Cantonese: ho4 fan2; Mandarin: hé fěn) is a type of local Cantonese broad noodles.

4.7.4 Systemic transfer and lexifier filter

The theory of systemic transfer and lexifier filter is put forward by Bao (2005, 2015), which proposes a wholesale transfer of the aspectual system from Chinese to CSE. Bao's (2005) theory consists of two different levels: (i) the source language (the substrate language), which provides the entire grammatical system (e.g. aspectual system), and (ii) the grammatical system of the recipient language (the lexifier language), which functions as a gate keeper (filter) (Bao 2005:258). Both levels endeavor to achieve optimal effect of the linguistic outcomes.

(i) System transfer

An entire grammatical subsystem is involved in substratum transfer [The substrate language (Chinese) offers the entire aspectual system].

(ii) Lexifier filter

Morphosyntactic exponence of the transferred system conforms to the (surface) structural requirements of the lexical-source language [English acts as a filter and excludes those “inexpressible” features].

(Bao 2015:59)

In Bao (2005, 2015), he illustrates that the aspectual markers *already* and *ever* are heavily influenced by the aspectual markers *le* and *guo* in Chinese, where English would use either the perfect or the simple past (see Bao 2015:38–44, also see Section 5.1 and Section 5.3). Leimgruber (2013:81) points out that convincing as the theory of system transfer and lexifier filter might seem, it fails to explain the use of reduplicate use of verbs (VV) such as *read read* ‘to try to read’, *work work work* ‘to work continuously’ (Leimgruber 2009:178). Such reduplicate use of verbs is apparently influenced by the Chinese tentative aspect (or delimitative, see Leimgruber 2009:178, 2013:81), but is not filtered by structural requirements of English. However, the lack of productivity of the putative tentative aspect in CSE validates the effect of the lexifier filter (Bao 2015:64). Besides, English does not

completely rule out reduplication, as it also allows partial reduplication, namely a reduplication of part of the word, e.g. *willy-nilly*, *teeny-weeny* and *mumsie-wumsie* (Bao 2015:63).

Apart from the above-mentioned issue, recent research points out that Bao's (2015) theory oversimplifies the contact situation of CSE into contact between Mandarin Chinese and English. Ziegeler (2015:233) notes that as much as the parallels between Mandarin Chinese and English seem obvious, a complete relexification of Mandarin in CSE is not very likely the case. Historically motivated criticism (see Lim and Ansaldo 2016:126) also argues that the Malay varieties (i.e. Bazaar Malay and Baba Malay, see Chapter 3) are the principal historical substrate languages that give rise to CSE. However, there is no convincing evidence that grammatical restructuring in CSE can be attributed to the Malay varieties (Bao 2015:30). In addition, both varieties were already strongly influenced by Hokkien, which, together with Teochew and Cantonese, also directly functions as an important historical substrate of CSE (see Bao 2015:35; Gupta 1998:125). Furthermore, Mandarin Chinese has a more important role to play in the linguistic ecology of Singapore before the 1950s than many (e.g. Ansaldo 2004; Gupta 2001) have estimated (see Chapter 3). Another reason why Mandarin Chinese can be used to analyze the substratum influence of Chinese on CSE is the morphosyntactic similarities between Mandarin and the other Sinitic substrates (Bao 2015:20).

4.8 Linguistic ecology and feature pool model

The ecology theory of Mufwene (2001) is perhaps one of the most untraditional and adventurous theories of linguistic proposals in the 21st century, in which he adopts ideas drawn from population genetics and biology and applies them to describe the evolution of creoles and other contact languages. This model compares linguistic features to genes and languages to species (parasitic species, in particular) (Mufwene 2001:145). Mufwene (2001) argues that like genes, linguistic features are inherited, but can be transferred and changed;

and like parasitic species, language is dependent upon its hosts, “i.e. its speakers, on the society they form, and on the culture in which they live” (2001:16).

Mufwene (2001) regards language change, as well as the emergence of contact-induced varieties as part of a competition-and-selection process. The general idea is that speakers make selections from a linguistic pool consisting of the sum of individual forms and variants. While interacting with each other, they modify their selected utterances, i.e. idiolects, to accommodate to their interlocutors and/or to adjust to new communicative needs (Mufwene 2001:18). Some idiolectal outputs, which are added to the feature pool, may be repeated (and possibly modified) by the same or other speakers (Van Rooy 2010:9). Some features may be eliminated, others may be reduced or increased in number (Mufwene 2001:198). Such repetition “sets in motion constant competition-and-selection processes that bring about changes of all kinds”, which in turn cause a language to evolve into a new communal system (Mufwene 2001:12).

The position of Mufwene (2001) is consistent with Weinreich (1953) and Milroy (1997) that contact-induced change is speaker-based.³³ It is speakers who contribute features to a pool from which they select utterances and in turn these selections can affect the evolution of a language (Mufwene 2001:18). In Milroy’s (1997:311) words, “linguistic changes, whether their origins are internal to a variety or not, are passed from speaker to speaker in social interaction. As for language contact, it is not actually languages that are in contact, but the speakers of the languages.” Mufwene (2001:14) also emphasizes that the contact of the dialects “takes place in the mind of the speaker”. Another argument in favor of the language-as-parasitic-species analogy is that when the host population disappears, both languages and parasites vanish (Clements 2003:590).

However, although speakers play an essential role in language evolution, their selections and modifications of linguistic features proceed largely unconsciously, and are subject to certain selection principles and conditioning factors. These principles and factors are referred to as the ecology of the contact situation (Mufwene 2001:21–22). The ecology

³³ The term “speakers” here includes listeners and also the writers when literacy is present (Milroy 1997).

theory distinguishes internal and external factors, which correspond to the concepts of social predictors and linguistic constraints, respectively (as discussed earlier in Section 4.5). The internal ecology involves, for example, the language system, structural rules within a language and linguistic features within the system before the system reorganization (Mufwene 2001:22–23), structural characteristics of the languages that serve as input (Biewer 2015:82),³⁴ and “linguistic parallels” (Lim 2009:199) in form and function. External factors include the socio-historical background of the contact situation, the demographic make-up of the communities that are in contact and the power relations that hold them, language policies and language attitudes (Lim 2009:100,199). On top of the distinction between internal and external ecology, Mufwene also identifies frequency, salience, regularity, and transparency (Mufwene 2001:57) as well as cognitive factors (Mufwene 2001:31–32) as significant factors determining the outcome of the selection process. It is worth mentioning that while the theoretical distinction between internal ecology and external ecology helps to understand the forces driving linguistic change, they are in complementary relationship interacting with each other rather than being separate dimensions (Ansaldo 2009:112).

There are several asymmetries in the analogy between language and species. First of all, as pointed out by Mufwene (2001) himself, “linguistic features are transmitted not only vertically (from older to younger speakers) and horizontally (among peers), but also bidirectionally: children do in turn influence their parents’ linguistic behaviors, in some cases more so than their parents influence theirs” (Mufwene 2001:16). This mismatch is highlighted by Lang (2012:1) that “genes” in biological sense come from one or two parents, whereas the features of a new idiolect come from virtually all other idiolects present in the linguistic ecology. Secondly, a further asymmetry lies in that while biologists consider the mixing of species as rare (Lang 2012:16), Mufwene (2008:132) considers the mixing of languages as pervasive, and that “every language is naturally a hybrid of some sort” (Mufwene 2002:55–56). Another important difference between species evolution and

³⁴ Also “linguistic properties of the substrates” in Lim (2009:199).

language evolution is that it is nature and the specific environments that select the more suitable and adaptable genes for survival, while it is individual speakers that select features for their own idiolects in Mufwene's (2001) theory. Mufwene (2001:16) speak of an "intervention of will", which plays a significant role in language evolution. Speaker's linguistic behaviors, such as conscious decisions to speak similarly or differently from some other specific speakers for reasons of identity, can have an important impact on a communal language (Mufwene 2001:16).

4.8.1 The feature pool of CSE

Following Mufwene (2001), Ansaldo (2009) argues that the role of ecology should be fundamental to any area of language change, and ecological factors are the driving force in contact language formation (Ansaldo 2009:95,98). In the Asian context, Ansaldo (2019:4) proposes that the most important external factors include (i) size of the communities, (ii) type of intra- and intersocietal network, and (iii) multilingual practices. Lim (2009:199) adds (iv) language policies, (v) language attitudes and (vi) speaker's prestige as factors belonging to the external ecology. The internal ecology, on the other hand, is referred to as "typological matrix" (TM) in Ansaldo's (2009) monograph, which serves as a convenient abstraction for the purpose of function-form analysis (Ansaldo 2009:112). In the Singaporean ecology, Ansaldo (2009:113) considers a TM dominated by (a) CSE, (b) Standard English, and (c) Singapore Mandarin. Apart from these languages, (d) Hokkien, (e) Cantonese, and (f) other ethnic language still contribute linguistic features to the TM of CSE to some extent (see Lim 2007). Figure 4.6 illustrates the linguistic ecology of CSE, and summarizes its external and internal factors.

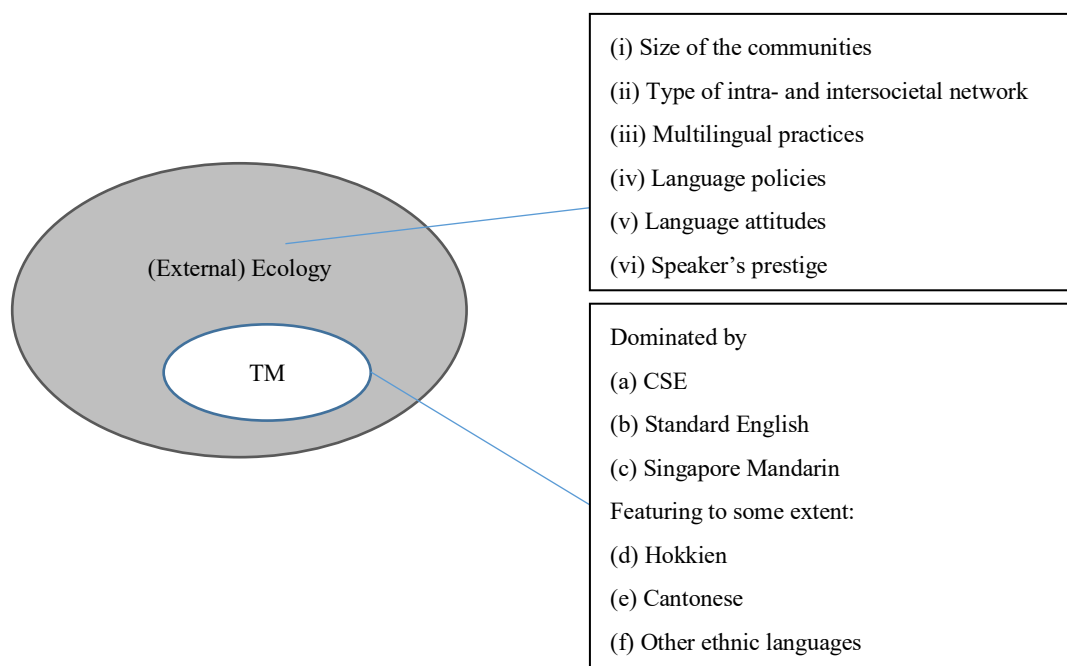


Figure 4.6: Internal and external ecology of CSE today (adapted from Ansaldo 2009:112; Lim 2009:199)

Ansaldo (2009) describes contact language formation as involving three processes, namely (i) selection of variant from a highly diverse and large pool of features; (ii) innovative replication and identical replication;³⁵ and (iii) propagation of variables in a population (Ansaldo 2009:110). Ansaldo (2009) argues that in the first process, selection and variation are caused by social and demographic factors. For example, linguistic features selected at different stages of the evolution of CSE reflects the dominance of different language(s) during the same time frame. The key question to understand the second process is why some features are more likely to be replicated. Following Croft (1995) and Mufwene (2001), Ansaldo (2009:111) sees that it depends largely on their typological and functional properties, i.e. internal ecology. He argues that a key factor behind the evolution of grammar is frequency, i.e. frequency of occurrence, and frequency of common structure. The third process concerns the stabilization of innovative features. Again, external ecology is fundamental, i.e. the characteristics of the community in which the propagation occurs, e.g. size, open, close, diffuse, stable, etc., will influence the stabilization of innovative features (Ansaldo 2009:11, also see Williams and Kerswill 1999; Kerswill and Williams 2000;

³⁵ Here, innovative replication refers to interference (i.e. mismatch of form and function) and identical replication corresponds to the notion of borrowing (see Ansaldo 2009:109). See also Section 4.3.

Milroy 2002). For example, stabilization of an innovative feature will take lesser time in small and tightly knit groups than large and diffuse communities (Ansaldo 2009:111).³⁶

Applying the feature pool concept lifts an obvious limitation of the traditional contrastive approach, which compares the contact variety to the putative lexifier – usually using Standard English as the target language (Bao 2005, 2015; Gupta 1989; Platt 1975). The shortcoming of these earlier contrastive approaches is, as a growing number of studies demonstrate, that no single language seems to constitute the exclusive target of CSE (Ansaldo 2004; Leimgruber 2013; Lim 2007; Ziegeler 2016, 2021). However, it is difficult to validate the feature pool concept and apply it in empirical research as it is very “metaphorical” in nature (Lange 2012:68). Another criticism on the model is that the feature pool appears “unstructured” (Lange 2012:67). The feature pool, according to Schneider’s (2007:22) interpretation, “consists of the sum total of the individual forms and variants that each of the speakers involved, with different language backgrounds and varying linguistic experiences, brings to the contact situation.” The questions as to how new features emerge, and how we can test the stabilization of an innovative feature that was selected from the feature pool remain unspecified.

Yet, a useful tool we can use is the measurement of frequency. Ansaldo (2009:116) shows that frequency is a fundamental factor behind contact-induced change in an evolutionary framework. Bybee (2007) and Haspelmath (2006) also confirm that frequency of usage is one of the most important factors responsible for the depth of entrenchment and the degree of abstract generalization of grammatical rules. Frequency is also a tool to test difference made through substrate influence (Biewer 2015:97). Tagliamonte (2012) also finds that frequency of innovative features identifies ethnicity.

In the case of the variables under study, this means that it is essential to study the language-internal competition among different functions of *already*, *also*, *ever*, and *one*, as well as language-external factors that affect this competition-and-selection process. In Chapter 3, I have described the socio-linguistic background of Singapore and showed that

³⁶ This does not mean that small groups experience frequent change. In Trudgill (2002), for example, small communities demonstrate very strong maintenance.

the Sinitic language are dominant in the ecology of the formation of CSE. Ansaldo (2009:144) also notes that “numerical and typological dominance mean that Sinitic and Malay variables are more frequent and salient and thus more readily available for selection and replication”. Therefore, I will analyze different functions of these variables (see Chapter 5) and offer an approximation of the diachronic development of CSE by counting the frequencies of the aspect marker *already* and *ever*, the additive adverb *also* as well as the emphatic marker *one* – plausible indicators of Chinese-influenced code (see Chapter 8 and Chapter 9).

4.9 Summary and conclusion

This chapter discusses the theoretical foundation of studying CSE. First of all, it highlights the complexity in the formation of CSE and in strictly defining it as a type of creole, mix-language, New English, or Chinese variety. The conclusion is that it is a contact variety, which is lexified in English, but shows lexical and grammatical influences from Sinitic, Malay, and other local substrate languages it had and has contact with. The grammatical restructuring, however, can be largely attributed to the Sinitic varieties.

Though various studies – both theoretical and empirical – of CSE have increased our understanding of the formation and variation of CSE considerably, the main problems with most of the theories (e.g. Platt 1975; Gupta 1989; Bao 2015) in accounting for CSE are that: (i) they presume there is a source language and a recipient language and (ii) they fail to take into consideration the extralinguistic factors such as the choices that individuals make in different communication contexts, and their social backgrounds (age, gender, educational level, ethnicity) in influencing the outcomes of CSE. As far as we can see, the Cultural Orientation (Asagoff 2007, 2010), the Indexicality (Leimgruber 2009, 2013) and the feature pool models (Mufwene 2009; Ansaldo 2009) seem to be superior in their explanatory power for CSE. They are more advanced as they consider, apart from typological issues, speakers as the locus in the forming of CSE, and, therefore, allow for more flexibility and intentionality in the linguistic behavior of the Singaporean speakers.

In light of the discussion of the Dynamic Model in accounting for the postcolonial Englishes by Schneider (2007), CSE is assumed to have reached stage 4: endonormative stabilization, in which Standard British English no longer serves as a norm of Singapore English, and the innovative linguistic features of the CSE become increasingly accepted. However, the theory is supported by mainly using largely synchronic data, and it will be fascinating to capture the evolutionary pathway of CSE by using a hitherto unused historical corpus, which also encompasses important metadata of individual speakers.

5 The use of *already*, *also*, *ever* and *one* in CSE

Having explored the theoretical context of contact linguistics in general and sociolinguistic typological models for CSE in Chapter 4, this chapter elaborates on the linguistic discussion on the use of *already*, *also*, *ever*, and *one* in CSE, which are the four linguistic variables used in the empirical part of this study. I argue that these four words differ significantly from native Englishes (BrE as a reference) in terms of their sentence position. *Already*, *also*, and *one* prefer phrase- and sentence-final position in CSE, while *ever* prefers a position directly before an actual verb in bare, non-inflectional form. The syntactic parallel patterns in their sentence positions and in negative sentences, in contrast to Mandarin Chinese, suggest a semantic extension or grammaticalization as they have expanded from their original usage to other grammatical domains, e.g. tense and aspect markers, along with focus markers. For example, CSE *already* has grammaticalized from a lexical item of what is known as a “Phasal Polarity (PhP) expression” (van der Auwera 1993; van Baar 1997) to an aspectual marker expressing the completive and inchoative aspect (Bao 2015). I propose that such semantic extension (or possible grammaticalization) is a result of contact-induced interference introduced by the Chinese substrates spoken in the local area. This chapter introduces the various meanings and functions of these four words in CSE and their preferred syntactic position.

In the following sections, I will describe the way *already*, *also*, *ever* and *one* are used in Colloquial Singapore English and how they differ from Standard English in terms of their syntax and semantics. Mandarin Chinese will be adopted as a representative for the analysis of the substrate languages. If not marked otherwise, examples of Mandarin are from personal knowledge as a native speaker. Some scholars on CSE argue that the inclusion of Mandarin among substrate languages of CSE is problematic (Ansaldo 2004; Lim 2007), as the principal historical substrates of CSE are varieties of Malay (i.e. Bazaar Malay and Baba Malay), Hokkien, Teochew and Cantonese. However, as discussed in Chapter 3, Mandarin Chinese has a more important role to play in the linguistic ecology of Singapore than many have estimated (e.g. Ansaldo 2004; Gupta 2001). In addition, despite the fact that CSE shows

lexical influences from Malay, there is no convincing evidence that grammatical restructuring in CSE is influenced by the Malay varieties (Bao 2015:30). In addition, both Malay varieties were already strongly influenced by Hokkien, which, together with Teochew and Cantonese, also directly functions as an important historical substrate of CSE (Bao 2015:35; Gupta 1998:125). Although phonological and lexical differences exist between Mandarin Chinese and other Sinitic substrates, they are morphologically and syntactically similar enough (Bao 2015:20; Hiramoto 2015) to yield interesting results concerning grammatical treatments of the expressions *already*, *also*, *ever*, and *one* in CSE.

5.1 Already

The adverb *already* with modified aspectual meanings in CSE is often considered as a result of contact-induced interference introduced by various Chinese vernaculars (e.g. Hokkien, Cantonese, Mandarin) (Platt and Weber 1980, Kwan-Terry, Bao 1995, 2005). Previous studies have shown that CSE *already* differs from *already* in Standard English in its ability to convey various types of aspectual meanings in CSE, such as (i) completive (or perfective to mark the completion of an event or action), (ii) inchoative (to mark the beginning of a state or event) or inceptive (to signal an event that just started) and (iii) prospective (which refers to actions that are about to start) (Platt and Weber 1980; Bao 1995, 2005; Fong 2005; Teo 2020). Table 5.1 offers an overview of different functions of *already* as described by various authors as well as the Chinese vernaculars used as reference languages in their studies. I will elaborate on each of these aspectual functions carried by *already* in the following subsections before capturing the grammatical differences between CSE *already* and *already* in Standard English. Furthermore, their differences in terms of sentence position as well as frequency of use will be highlighted.

	Platt and Weber (1980)	Kwan-Terry (1989)	Bao (1995, 2005)	Fong (2005)
Function of <i>already</i>	i. completive marker	i. perfect aspect marker	i. perfective aspect marker	i. reaching an endpoint
		ii. to signal change of a new state	ii. inchoative aspect marker iii. inceptive aspect marker	ii. beginning iii. an event in the immediate future (prospective)
Chinese vernacular(s) for analogy / of reference	Hokkien	Cantonese and Mandarin	Mandarin	N/A

Table 5.1: Overview of different functions of *already* as described by various authors

5.1.1 *Already* as an aspectual marker in CSE

Already as a completive/perfective marker

Platt and Weber (1980) are among the earliest scholars to shed light on the aspectual use of *already* in CSE. They consider *already* as marking the completive aspect, which is expressed in British English or American English by past tense or the present perfect. Consider the examples in (38):

- (38) a. I work about four months *already*. (Platt, Weber 1980:65–66)
‘I have worked for four months.’
- b. I stay in X seventeen year *already*.
‘I’ve lived in X for seventeen years.’
- c. I only went there once or twice *already*.
‘I’ve been there only once or twice.’

Compared to Standard English, CSE seldom employs inflectional morphology. As shown in (38)a and (38)b, *already* occurs after uninflected verbs to mark the completive/perfective without involving any inflectional morphemes. In (38)c, the simple past is used instead of the perfect. Besides, all the above examples with *already* show no emphasis that a previous

action has taken place, or a state has lasted for a long period of time, which would be the reading of *already* in Standard English (see Section 5.1.2). It is also important to note that CSE *already* prefers sentence-final position while Standard English *already* normally occurs in preverbal position (Bao 2005:260). Platt and Weber (1980:66) attribute this use of CSE *already* to the aspectual marker *liaú* in Hokkien Chinese (the counterpart of *le* in Mandarin Chinese). Like the CSE *already*, *liaú* functions as a completive marker, which expresses the completion of an action. It occurs in sentence-final position, as shown in (39).

- (39) Hokkien (Platt and Weber 1980:66)
 Gun thauke tng chhu *liaú*.
 our boss return home ASP
 ‘Our boss has returned home.’

Kwan-Terry (1989) notices that it is very common among school children in Singapore to use *already* to replace the aspectual markers (*-ed* or *have + -ed*) of native varieties of English. She examines the processes and strategies that Elvoo uses, a child from an ethnically Chinese middle-class family, while learning English and Cantonese simultaneously. As shown in (40), Elvoo’s utterances of *already* are built on the Cantonese aspectual marker *jo2* (左).

- (40) a. Cantonese (Kwan-Terry 1989:39, [glosses mine])
 我 食 个 cake 左
 Ngo5 sik6 go3 cake jo2. (3; 9)³⁷
 I eat CL cake ASP
 ‘I have eaten the cake.’
 b. CSE
 I eat the cake *already*. (3; 9)

The pair in (40) show that Elvoo considers the Cantonese aspectual marker *jo2* (左) and the English adverb *already* as parallels. Here, *already* expresses that the eating event is over, i.e. fully occurred before the moment of utterance. Interestingly, as shown in (40)a, instead of using the Cantonese completive marker *jo2* directly after the verb, Elvoo detaches it from

³⁷ (3;9) indicates that the utterance was produced by Elvoo when he was 3 years and 9 months old.

the verb *sig* (食 ‘to eat’) and places it sentence-finally, which is not idiomatic in local Cantonese. In the English utterance (40)b, he fails to conjugate the verb *eat* into its past tense form *ate* and instead uses *already* to express that particular grammatical meaning. Moreover, Elvoo uses *already* sentence-finally. The lack of inflectional markers in Elvoo’s English expressions, according to Kwan-Terry (1989), is due to Cantonese interference. Cantonese, like other Chinese vernaculars, does not possess overt inflectional morphology.³⁸ The parallel between *jo2* and *already* in form and function suggests a mutual influence between his English and Cantonese repertoires. In both (39) and (40), CSE speakers grammaticalize a lexical item to an aspect marker instead of using inflectional morphology to mark the completive/perfective aspect, following the aspect-marking strategy of the substrate languages.

Already as an inchoative marker

While completive/perfective *already* marks the completion of an action, inchoative *already* marks the beginning of an event (Kwan-Terry 1989:37; Bao 2005:241). Each of the examples (41)–(42) from Bao (1995) has two possible readings. With the completive interpretation, sentence (41) is interpreted as ‘My son has left for school’. However, the second reading is more frequent in which the sentence means ‘My son has started school’. We can see that in Standard English, the inchoative verb *start* is required to convey that particular meaning.

- (41) My son go to school *already*. (Bao, 1995:183)
 a. ‘My son has left for school.’ (completive)
 b. ‘My son has started school.’ (inchoative)
- (42) My baby speak *already*. (Bao, 1995:183)
 a. ‘My baby spoke.’ (completive)
 b. ‘My baby has started to speak.’ (inchoative)

³⁸ Some studies have shown that certain varieties of Cantonese, e.g. Taishan and Zhongshan use tone change as an inflectional or a derivational strategy (see Stump 2001).

Kwan-Terry (1989) notes *already* in collocation with an adjective³⁹ expresses the notion of a change to a new state or situation, as shown in (43). Again, parallel patterns can be observed in Cantonese and CSE.

- (43) a. CSE: The tongue red *already*, you see? (3;6)
‘The tongue has turned red, you see? (Kwan-Terry 1989:37)
- b. Cantonese: 熟 左 吗
Suk6 jo2 ma3 (3;8)
ripe ASP INT.PTC
‘Has it ripened?’ (Kwan-Terry 1989:40, [glosses mine])

In (43)a, *already* is used with the adjective *red* to signal the change to a new state, from the tongue being not red to red. Standard English *already* usually involves a “contrary to expectation” reading (Teo 2019), which assumes a proposition that is contrary to a hearer’s expectation. Yet, CSE *already* seems to signal the change itself. Thus, there are no conflicting expectations.

Bao (2005) further distinguishes the concept of “inceptive”, which means ‘just started’. While the function of inchoative is to mark a state transition, inceptive emphasizes that the start of a new state occurs shortly before the utterance, as exemplified in (44):

- (44) It rain *already*. (Bao 2005:241)
‘It has just started to rain.’

Both inchoative and inceptive will be referred to as inchoative in this study due to their overlapping temporal readings, i.e. the change to a new state is the prerequisite for the “just started” reading. For example, the sentence “My son go to school *already*” can be interpreted as ‘My son has just started school’ (the inceptive interpretation), which includes the reading of a change of state – ‘My son did not go to school before but he goes to school now’ (the inchoative interpretation). Apart from that, the subtle reading of “just started” is difficult to capture as it heavily depends on the context.

³⁹ Chinese adjectives can be used as static/stative verbs as exemplified in (43)b.

Already as a prospective marker

More recent studies have shown that *already* can also serve as a prospective aspectual marker, which denotes an action that will happen shortly in the immediate future (see Fong 2005; Teo 2020). Example (45) shows that *already* is used to indicate that the action of returning home will be completed in the near future, i.e. she is on her way home at speech time. It is worth noting that such an interpretation of *already* again depends on the context of the interlocution. Other aspectual readings, such as the completive and inchoative are possible.

- (45) She come home *already*. (Teo 2020:89)
 ‘She is on her way home.’

The completive, the inchoative and the prospective uses of *already* are not derived from English as English relies on the perfect and/or periphrastic expressions to express these aspectual meanings. It is commonly considered that these aspectual uses of *already* are remodeled after the Chinese substrates, which express precisely the corresponding meanings (see Bao 1995, 2005; Gupta 1994; Lim 2007; Platt and Weber 1980). Consider the following examples in Mandarin Chinese:

- (46) 我们 吃 了 饭
 wǒmen chī le fàn
 we eat ASP meal
 ‘We ate meal; we have finished eating.’ (completive)
- (47) a. 墙 白 了
 qiáng bái le
 wall white ASP
 ‘The wall has turned white, i.e. it is newly painted white.’ (inchoative)
- b. 我 儿子 上学 了
 wǒ érzi shàngxué le
 my son go school ASP
 ‘My son has started school.’ (inchoative)

- (48) a. 快 下 雨 了
 kuài xià yǔ le
 almost down rain ASP
 ‘It is about to rain.’ (prospective)
- b. 她 要 回家 了
 tā yào huíjiā le
 she will go home ASP
 ‘She will go home soon.’ (prospective)

Similar with the CSE *already*, the aspectual *le* directly following a verb gives rise to the completive interpretation, as shown in (46). In (47)a, *le* in collocation with the adjective predicate *bái* ‘white’ leads to the inchoative reading ‘the wall has turned white’, which corresponds to the inchoative *already* in CSE with an adjective in the earlier example (43). Example (47)b is almost a one-by-one direct translation of (41). Likewise, the prospective reading of the sentence-final *le* in (48) is identical to that of *already* in (45).

It should be emphasized that there are two different *le*-s in Chinese which serve different grammatical functions (Bao 2005; Chao 1968; Li and Thompson 1981), although they appear in the same form represented by the character 了. One is verbal *le* (V-*le*), which occurs directly after an actual verb; and the other is referred to as sentence-final *le* (S-*le*), which occurs sentence-finally (Bao 2005; Chao 1968; Soh and Gao 2006). While V-*le* marks the completive aspect, S-*le* is said to signal a “Current Relevant State” (Li and Thompson 1981), which corresponds to the English perfect. I will discuss the functions of these two different *le*-s as well as their different pathways of grammaticalization in Chapter 6. Regarding CSE *already*, Bao (2005:242) considers V-*le* responsible for the completive reading while S-*le* gives rise to the inchoative or prospective meaning, as exemplified in (49).

- (49) a. 下 了 雨
 xià le yǔ
 down ASP rain
 ‘It rained.’(completive)
- b. 下 雨 了
 xià yǔ le
 down rain ASP
 ‘It started to rain.’ (inceptive)

- c. 快 (要) 下 雨 了
 kuài (yào) xià yǔ le
 almost (will) down rain ASP
 ‘It is about to rain.’ (prospective)

However, CSE *already* seldom occurs in direct postverbal position. Examples like “I eat *already* lunch” are rarely attested in CSE. Bao (2005; 2015) attributes this to the force of the constraint of the lexifier filter on CSE (see Section 4.7.4), as the syntactic requirement of English does not allow direct postverbal position for *already*. Therefore, *already* is relocated in sentence-final position, which is compatible with both Standard English *already* and the Chinese sentence-final *le*.

Already in negative sentences

The CSE *already* can occur in negative sentences, but Standard English *already* seldom occurs in negative sentences (Bao and Hong 2006). As described in Huddleston and Pullum (2002:710), *already* is a “positively oriented polarity-sensitive item” (PPI) which characteristically occurs in positive clauses. Another adverb that belongs to this category is *still*, which is concerned with a continuing situation. Consider the following pairs:

- (50) a. *Mary isn’t *still* here.
 b. Mary isn’t here *anymore*.
 (51) a. *Mary *already* isn’t here.
 b. Mary isn’t here *yet*.

Examples (50)a and (51)a are grammatically not acceptable in Standard English, as Standard English normally does not allow using *already* and *still* in negative clauses. Instead, *anymore* and *yet* are used in these situations, which are the corresponding “negatively oriented polarity-sensitive items” (NPIs) (Huddleston and Pullum 2002:712). Example (50)b presupposes that “Mary was there” prior to a reference point and suggests that such a state no longer holds whereas Example (51)b suggests a common ground that “Mary is not here” and carries the implication that “she may be here in the near future”.

On the other hand, CSE *already* in negative sentences corresponds to the inchoative aspect, which expresses the change of state from a positive one to a negative one (Bao 2005:247). Consider Example (52), which suggests that Mary had too much food and cannot eat anymore at the moment.

- (52) CSE
Mary cannot eat *already*.
'Mary cannot eat anymore.'

Similarly, it is common for sentence-final *le* in Chinese to occur in negative clauses to mark a change of state, as illustrated in (53):

- (53) 我 找 不 到 钥 匙 了
wǒ zhǎo bú dào yàoshi le
I find NEG RVC key ASP
'I could not find the keys anymore.'

As shown in (53), Chinese sentence-final *le* can appear in negative clauses to mark a transition from a positive state of “being able to find the keys” to a negative state of “not being able to find the keys”. The parallel constructions of CSE *already* and sentence-final *le* in negative clauses suggest that sentence-final *le* is a possible source of linguistic interference.

Malay sudah

The aspectual use of *already* to mark the completive, inchoative, and prospective, however, is not an exclusive phenomenon of CSE. Such use of *already* is treated in Olsson (2013) and Dahl and Wälchli (2016) under the label “iamitive” (from Latin *iam* ‘already’). Besides Mandarin *le*, Indonesian/Malay *sudah*, Tai *léew*, Vietnamese *đã* and *rồi* are considered as iamitive. Iamitive and completive, inchoative, along with prospective as well as the perfects are overlapping grams. For a detailed description of these, interested readers are referred to

Dahl and Wälchli (2016). The following examples show the usages of Malay *sudah* ‘already’, which is a possible candidate responsible for the reconstruction of aspectual *already* in CSE.

- (54) Dia *sudah* membaca buku ini (Dahl and Wälchli 2016:327)
 he ASP read book this
 ‘He has read this book.’ (completive)
- (55) a. Maria *sudah* ada di sini (Dahl and Wälchli 2016:327)
 Maria ASP COP in here
 ‘Maria is already here.’ (inchoative)
- b. Iwan *sudah* bekerja (adapted from Grangé 2010:254)
 Iwan ASP work
 ‘Iwan already worked/Iwan has started to work.’ (completive/inchoative)
- (56) Budi *sudah* nak datang (Olsson 2013:22)
 Budi ASP want come
 ‘Budi is about to arrive.’ (prospective)

Examples (54)–(56) show that *sudah* is compatible with various aspectual meanings, from completive and inchoative to prospective, resembling its counterparts CSE *already* and Chinese *le*. In (54), *sudah* conveys the completive meaning when it occurs with the aktionsart achievements/accomplishments. When *sudah* co-occurs with states, as in (55)a, the inchoative meaning can be interpreted. Example (55)b is ambiguous as both the inchoative and completive readings are possible. Finally, *sudah* can also elicit the prospective reading when it co-occurs with volitional verbs like *nak* ‘want’, as exemplified in (56).

In the above examples, *sudah* occurs in preverbal position. Although it is possible to place *sudah* in sentence-final position, as shown in (57), the sentence is regarded as uncommon and even ungrammatical by native Malay speakers (see Teo 2019:99).

- (57) ?Perjalanan yang melelahkan itu berakhir *sudah*. (Grangé 2010:249)
 journey REL tiring this finish already
 ‘This exhausting journey is finally over.’

A further difference between CSE *already* and Malay *sudah* is that while CSE *already* occurs in negative clauses, Malay *sudah* rarely occurs in negative sentences. Negative

sentences with *sudah*, like Example (58) are rated as “rare” by Malay informants in Teo’s (2019:99) study.

- (58) ?Dia *sudah* tidak kaya. (Teo 2019:359)
he already NEG rich
‘He is not rich now.’

To sum up, apart from Chinese *le*, Malay *sudah* may be reconstructed for aspectual *already* in CSE. Like Chinese *le*, *sudah* can serve as a completive marker when occurring with achievements and accomplishment, as inchoative marker when occurring with states and as prospective marker in collocation with volitional verbs. However, unlike Chinese *le* and CSE *already*, *sudah/dah* occurs in preverbal position but rarely in sentence-final position. Moreover, *sudah/dah* is rarely compatible with negative sentences, while Chinese *le* and CSE *already* can be used in negative sentences to express the inchoative meaning. In short, compared with Malay *sudah*, Chinese *le* is less restricted in its syntactic environment. It will be interesting to see whether these differences may lead to variation of CSE *already* among speakers of different ethnic languages. Therefore, we will explore later how CSE speakers with a Chinese ethnic background and those with a Malay background differ in their use of CSE *already* (see Chapter 7 and Chapter 8).

In the following section, we will examine the similarities and differences between CSE *already* and Standard English *already*.

5.1.2 Phasal polarity expression *already*

Phasal Polarity (PhP) expression *already* refers to the adverbial use of *already* in Standard English (Auwera 1993). It is phasal because it typically involves reference points at two related phases; it is bipolar because it involves situations which are contrasted as opposites with different polarity values, i.e. one of the two situations in question holds (+) whereas the other does not (-) (Kramer 2018:1). Consider the following example:

(59) Mary is *already* in Singapore. (BrE, adapted from van Baar 1997:1)

As shown in (59), *already* signals a proposition that the state of Mary's being in Singapore holds at reference time. Furthermore, *already* implies a further reference point at a prior phase where such state is not the case, i.e. Mary was not in Singapore. In such cases, phasal polarity *already* is identical to the inchoative *already* in CSE, as exemplified in (60):

(60) CSE
Mary is in Singapore *already*.

Although the sentence position of *already* is different, the semantics of English *already* and that of CSE *already* is similar. However, besides marking the occurrence of a change of state, Standard English *already* is typically enriched by connotations of anteriority and counterfactual expectation. In some other literature, such connotations are referred to as “earlier than expected” (Michaelis 1992; Fong 2006:254; Bao 2015), and “contrary to expectation” (Soh 2009). As pointed out by Michaelis (1992:326), “*already* not only encodes the existence of a given state of affairs at reference time, but also presupposes the ‘anteriority’ of that state of affairs to an interval of a specific type.” Such examples are depicted in (61):

(61) a. When I finally arrived, John was *already* sleeping. (BrE)
b. Even though I arrived early, John was *already* sleeping. (BrE)

Example (61)a expresses that the speaker arrives too late to meet John, as the act of arriving happens in the post-time of the change of state from not sleeping to sleeping. In (61)b, the speaker fails to meet John even though the speaker arrived early. This example depicts an unexpected situation, as the state of John's sleeping began earlier than expected.

Following Löbner (1989), van der Auwera (1993) proposes that the following three points should be considered when describing the temporal uses of *already*:

- i. the time axis;
- ii. the obtainment of a positive state resulting from a change from a negative state; and

- iii. an alternative to the envisaged positive state (ii) obtaining at a point of time on the axis (i). (van der Auwera 1993:619)

These temporal readings sufficiently explain the intuitions behind the examples in (61). The requirement for the *already* reading is two phasal reference points, i.e. it involves a change of state from negative (with the state not holding at the reference point) to positive (the state holding at the reference point) in English. Apart from that, any additional connotations have the status of implicatures and can be cancelled. Consider another pair of examples in (62), where (62)a expresses a sense of earliness that (62)b clearly lacks (adapted from van der Auwera 1993:621).

- (62) a. I've met a girl who is only 13 years old but she is *already* married.
b. Suppose you want to marry a certain woman. You propose and you find out that she is *already* married. There is nothing necessarily early about this marriage. You simply come too late to have a chance [...]

In conclusion, the change-of-state reading of PhP *already* in Standard English typically involves two reference points: one when the situation happens, and there is an implicit contrast of another time point prior to the situation. On top of that, the reading of *already* in Standard English is often associated with the connotation of anteriority and counterfactual expectation. The following sections will focus on the differences of other aspectual readings between CSE *already* and its counterparts in Standard English.

5.1.3 Differences between *already* in Standard English and CSE *already*

In the above section, we have compared the inchoative use of CSE *already* with PhP *already* in Standard English and concluded that they are similar in terms of the change-of-state reading, except that *already* in Standard English is enriched by the connotation of anteriority and counterfactual expectation. However, up to this point, we have only compared the inchoative *already* that uses a stative predicate in the two varieties, i.e. *already* that occurs

with situation types of states. What about *already* that occurs with a dynamic predicate, i.e. in situation types such as activities, accomplishments, and achievements?

According to Fong (2005), CSE *already* is compatible with situations of any aktionsart, including states, activities, achievements, and accomplishments (see Vendler 1957; Comrie 1976; Shirai and Andersen 1995 on how to distinguish different aktionsarten). See the following examples in CSE adapted from Fong (2005:253):

- | | | | |
|------|--|-------|-------------------|
| (63) | a. She is in Singapore <i>already</i> . (state)
'She is <i>already</i> in Singapore.' | ————— | stative predicate |
| | b. She sing <i>already</i> . (activity)
'She has sung.' | | |
| | c. I paint a picture <i>already</i> . (accomplishment)
'I have painted a picture.' | | |
| | d. I recognize her face <i>already</i> . (achievement)
'I recognized her face.' | | |
| | | } ——— | dynamic predicate |

In (63)a, CSE *already* occurs with a stative predicate, and it gives rise to a change-of-state reading. In such case, it does not differ much from the English *already* (see Section 5.1.2 above). However, when it occurs with a dynamic predicate, i.e. verbs of activities, accomplishments, and achievements, it is compatible with all three aspectual readings, i.e. completive, inchoative, and prospective (Teo 2019:352). Yet, the translational equivalents of CSE *already* in Standard English are the perfect or simple past. Standard English *already* is equally compatible with these above aktionsarten, e.g. “I have finished the homework *already*”, though it normally involves two kinds of expectations: (i) that the state, activity, accomplishment, and achievement would begin or end at some point; (ii) that the transition would not take place as early as it actually did (Dahl and Wälchli 2016:317).

As mentioned earlier, *already* in Standard English requires a contrasting alternative to the envisaged positive state obtaining at a point of time (van der Auwera 1993:621). However, CSE *already*, be it inchoative, completive, or prospective, signals a change without another reference point. As illustrated by (64) and (65), there is no second reference point in these examples; *already* seems to signal the change itself. In both examples, there are no presuppositions involved; therefore, there are no conflicting expectations.

- (64) The tongue red *already*, you see? (Kwan-Terry 1989:37)
 ‘The tongue has turned/turned red./*The tongue was red.’
- (65) The patient eat food *already*. (Bao 1995:183)
 ‘The patient has started to eat food.’

The temporal schema proposed by Bao (2005:240) captures the aspectual readings of *already* in CSE quite accurately. This schema illustrates that the use of *already* marks the change from Not-P to P (it was not the case before, and it is the case now). R refers to reference time. What is more, sentences with *already* as an inchoative marker in CSE are not ambiguous. There are two entailments in sentence (66), namely (i) that Lily did not go to school in the past, and (ii) that she goes to school now. Without the use of inchoative *already*, a similar sentence in Standard English, as in (67), is ambiguous because the interlocutor can conclude from the sentence that either Lily is still in school, or she is no longer in school.

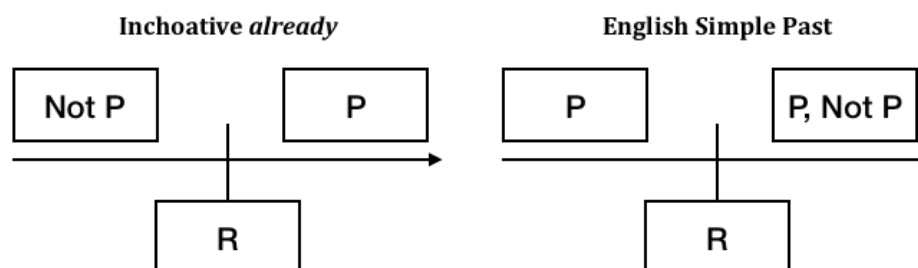


Figure 5.1: Inchoative already and the English simple past (adapted from Bao 2005:240)

- (66) CSE
 Lily go to school *already*.
 Means: ‘Lily did not go to school in the past.’ and ‘Lily goes to school now’.
- (67) Standard English (adapted from Bao 2005:240)
 Lily went to school.
 Means: ‘Lily is still in school.’ or ‘Lily is no longer in school.’⁴⁰

⁴⁰ Ziegeler (p.c. on 9 August 2019) kindly notes that English simple past on stative verbs is equally unambiguous, i.e. it carries a de-facto terminative presupposition. For example, the sentence “I was young” implies that “I am not young anymore”.

In sum, unlike *already* in CSE, which can be used to express the inchoative, completive, and the prospective aspect, Standard English relies on verbal morphology (*-ed*, *have + -ed*) to express the completive aspect, and does not use *already* to express prospectivity. In terms of the inchoative reading, CSE *already* is identical to Standard English *already* occurring with stative predicates, where *already* signals a change-of-state reading. Yet, Standard English *already* is typically enriched with the connotations of anteriority and counterfactual expectation, which is not limited to the change-of-state reading with stative predicate, but also concerns predicates that are dynamic. However, when CSE *already* occurs with a dynamic predicate, it is compatible with all three aspectual readings (inchoative, completive and prospective) depending on the context, but does not involve the above connotations of Standard English *already*.

5.1.4 The frequency of *already* and its preferred sentence position

Frequency difference of already between CSE and BrE

In the above section, we have examined the semantic differences between CSE *already* and Standard English *already*. Further differences between CSE *already* and Standard English *already* include the frequency of occurrences and their preferred syntactic frames.

Siemund and Li (2017) show that the frequency of *already* in CSE differs significantly from British English (BrE) with the ratio of *already* per thousand words in CSE being generally higher. They compare the occurrences of *already* in the Singaporean component and the British component of the ICE (ICE-SG vs. ICE-GB) corpora. The approach generalizes across standard and non-standard uses of these expressions, and does not consider their syntactic position.

Besides calculating the per-thousand-words ratio across the entire corpus, Siemund and Li (2017) also discovered individual differences in the usage of *already*. Table 5.2 compares the top 10 highest *already* ratios in ICE-SG with those in ICE-GB.⁴¹ The figures

⁴¹ Only the speaker contributions above one thousand words are included.

produced by the ten speakers with the highest ratios of *already* in ICE-SG are much higher than those in ICE-GB: The highest ratio of *already* appearing in ICE-SG is 10 ptw while the lowest is almost 2.9. In contrast, even the highest frequency of *already* produced in ICE-GB (3.7) would have only ranked the eighth on the list of ICE-SG.

The bottom line of the above discussion is that the per-thousand-words ratios of the prominent grammatical item *already* differ markedly in the varieties of Singapore and Great Britain, the values in CSE being generally higher. This suggests that higher ratios of *already* are indicative of higher substrate influence.

ICE-SG ID	already.ptw	ICE-GB ID	already.ptw
<icesg-s1a-007:1\$b>	10	<icegb-s2b-041:2\$a>	3.724
<icesg-s1b-057:1\$c>	6.293	<icegb-s1b-001:1\$a>	2.641
<icesg-s1a-051:1\$b>	6.162	<icegb-s1a-082:1\$a>	2.584
<icesg-s1a-020:1\$b>	5.709	<icegb-s1b-004:1\$a>	2.427
<icesg-s1a-042:1\$b>	5.239	<icegb-s2b-022:2\$a>	1.992
<icesg-s1a-049:1\$a>	3.937	<icegb-s1b-061:1\$b>	1.957
<icesg-s1b-034:1\$a>	3.726	<icegb-s1a-023:1\$b>	1.908
<icesg-s1a-054:1\$a>	3.626	<icegb-s2b-034:1\$a>	1.814
<icesg-s1a-091:1\$a>	3.509	<icegb-s1b-054:1\$b>	1.769
<icesg-s1a-013:1\$a>	2.874	<icegb-s1a-066:1\$b>	1.732

Table 5.2: The 10 speakers with highest ratios of *already* in ICE-SG and ICE-GB (adapted from Siemund and Li 2017:22–23)

Position of already

CSE *already* and Standard English *already* share the syntactic frames that they appear in, i.e. sentence-initial, pre-predicate (or sentence-medial),⁴² and sentence-final (Bao 2015). Consider the following examples drawn from OHI.

- (68) CSE *already* according to their sentence position
 a. Initial: *Already* there were controversies... [OHI-002307-VTA]

⁴² Pre-predicate position refers to the normal mid-position for adverbs (between the subject and the main verb, or after the modal verb, the first auxiliary verb or the copula *be* as a main verb.

- b. Medial: There was *already* a Chinese newspaper called...
[OHI-000064-SCY]
- c. Final: They can get married *already*. [OHI-002686-SMY]

However, the preferred position of CSE *already* is the sentence-final position (see Bao and Hong 2006), as shown in (68)c. Therefore, CSE *already* is also often referred to as one of the “sentence-final particles” or “sentence-final adverbs” (SFPs or SFAs), along with the focus marker *only* and the additive marker *also* (Hiramoto 2015:636). On the other hand, Standard English *already* occurs most frequently in the pre-predicate position, as in (68)b (see Brown 1999). Bao and Hong (2006) explore *already* according to these different positions in four major text categories in the ICE corpora. As displayed in Table 5.3, *already* in CSE appears remarkably more often in sentence-final position than in BrE. The differences are concentrated in the spoken registers, especially in the private dialogue category. Moreover, *already* occurring in negative sentences is only found in ICE-SG, but not in ICE-GB.

	Initial	Medial	Final	Other ⁴³	Total
PRIVATE DIALOGUE					
GB	0.02	0.35	0.07	0.00	0.44
SG	0.06	0.84	1.94	0.10	2.94
PUBLIC DIALOGUE					
GB	0.03	0.59	0.05	0.00	0.67
SG	0.01	1.11	0.18	0.03	1.33
MONOLOGUE					
GB	0.03	0.80	0.03	0.00	0.86
SG	0.03	0.62	0.12	0.00	0.77
WRITING					
GB	0.00	0.74	0.04	0.00	0.78
SG	0.04	0.78	0.03	0.00	0.85

Table 5.3: Frequency of *already* per text in ICE-GB and ICE-SG (Bao and Hong 2006:109)

Based on the above analysis, we can conclude that the most frequent syntactic position that *already* occurs in is different in CSE and BrE, which may be attributed to the substrate

⁴³ The category “Other” includes tokens of *already* found in negative sentences.

influences on CSE. Li (2014) reveals that CSE speakers with a Chinese ethnic background prefer *already* in sentence-final position while speakers with a British ethnic background mostly use *already* in sentence-medial position. A third group, which is referred to as “Others” in Li (2014), consists of speakers of other ethnicities such as Malay, Indian and Arab, who show an almost equal distribution of sentence-medial *already* and sentence-final *already*. A more recent study by Teo (2019) compares CSE *already* used by Chinese speakers with Malay speakers, and confirms that Chinese speakers tend to use *already* sentence-finally while Malay speakers do not have a preference for *already* appearing in either pre-predicate or sentence-final position. Additionally, *already* also occurs more frequently in negative contexts among the Chinese speakers than among the Malay speakers (Teo 2019:368). In conclusion, both studies support the assumption that CSE *already* is more likely to be influenced by the Sinitic substrates than other contact languages.

5.1.5 Summary

To sum up, there are three main differences between CSE *already* and Standard English *already*. First, CSE *already* is used to express various aspectual meanings, i.e. the completive, inchoative and prospective. Although Standard English *already* can also mark the inchoative aspect, it only does so when it occurs with stative predicates. Standard English *already* typically involves two contrasting temporal reference points. Secondly, CSE *already* can appear in negative environments, meaning ‘anymore’ or ‘any longer’, whereas Standard English *already* normally occurs in positive contexts, and cannot take a scope over the negative. Lastly, CSE *already* prefers sentence-final position while Standard English *already* typically occurs in sentence-medial (pre-predicate) position. The above sections have also highlighted a few studies related to the frequency of *already*, its preferred syntactic position in CSE and BrE, as well as the frequency of *already* in different contexts (written vs. spoken). It appears plausible to relate these frequency differences to substrate influence, with the higher ratios of *already* being indicative of higher substrate influence.

5.2 Also

5.2.1 Additive marker *also*

Compared with *already*, the semantics of the adverb *also* in CSE does not diverge significantly from Standard English. In both varieties, *also* means ‘in addition’, which is labeled as an “additive marker” or “additive focus marker” (Bao and Hong 2006:109; König 1991:48, 2017:33). The function of additive markers is to point to the existence of an alternative to the associate of the additive (Forker 2016:72). For example, *also* in (69) presupposes that “they brought something else” or “someone else brought his own boat”. However, Standard English *also* prefers sentence-medial position while CSE *also* is used more frequently in sentence-final position, as exemplified in (69). In Standard English, the more natural choices of additive markers in sentence-final position are *as well* and *too* (Brown 1999).

- (69) CSE
 They brought their own boats *also*. [OHI-000263-S]
 ‘They brought their own boats, too.’

Apart from the syntactic difference, Bao and Hong (2006) note that *also* carries a subtle grammatical function when used with *even* and universal quantifiers (e.g. *all*, *everything*, *everyone*, *every day*, etc.). It reinforces the universal meaning of universal quantifiers and the concessive meaning of *even*, as shown in the following examples from OHI:

- (70) Universal *also*:
 Everyone works as a team. Food, everyone eat *also*. [OHI-001953-LAS]
- (71) Concessive *also*:
 Even the holder *also*, they have different type of holder. [OHI-001953-LAS]

In OHI, I find cases where CSE *also* is used in negative contexts instead of *either*, as shown in (72). It seems that CSE *also* is not a polarity-sensitive item. However, in Standard English, *also* (or *too*) is substituted by *either* if it is negated. *Also* rarely occurs in negative contexts

(Bao and Hong 2006:110). It is worth mentioning that *too* can sometimes occur in negative contexts, but the proposition of the sentence is different, as illustrated in (73) (Fuchs 2012:41; König 1991:48, 2017:33).

- (72) CSE
 They don't have the time to have a drink *also*. [OHI-001953-LAS]
 So your lunch time wasn't fixed *also*? [OHI-001953-LAS]
- (73) Standard English (adapted from König 1991:48)
 a. I hope they did not lay off Paul, *either*.
 proposition: 'Somebody other than Paul was not laid off.'
 b. I hope they did not lay off Paul *too*.
 proposition: 'Somebody other than Paul was laid off.'

Bao and Hong (2006) postulate that CSE *also* is influenced by the Chinese additive markers 也 *yě* and 都 *dōu* 'also, all, both, too'. Like *also*, the additive markers *yě* and *dōu* reinforce the meaning of a universal quantifier or a concessive marker, as shown in (74) and (75). It should be noted that *dōu* can be used in concessive and universal situations in both Mandarin Chinese and Cantonese. However, while *yě* can occur in a concessive situation in Mandarin Chinese, it cannot be used in a universal situation. Apart from that, 也 *ya5* is not preferred in Cantonese, as shown in (75).

- (74) Mandarin Chinese
- a. 每 样 我 都 要
 měi yàng wǒ dōu yào
 every kind I ADD want
 'I want everything.'
- b. *每 样 我 也 要
 měi yàng wǒ yě yào
 every kind I ADD want
 (Intended meaning) 'I want everything.'
- c. (连) 皇帝 本人 也/都 亲 临 前线
 (lián) huángdì běnrén yě/dōu qīn lín qiánxiàn
 (even) emperor INT ADD INT come front line
 'Even the emperor himself went to the front line.'
 (adapted from Wang 2011:34)

- (75) Cantonese Chinese
- a. 样样野 我 都 要
 yeung6yeung6ye5 ngo5 dou1 yiu3
 everything I ADD want
 ‘I want everything.’
- b. 中文 都 唔 好
 zung1man2 dou1 m4 hou2
 Chinese ADD not good
 ‘Even his Chinese is not good.’
- c. ?中文 也 唔 好
 zung1man2 ya5 m4 hou2
 Chinese ADD not good
 ‘Even his Chinese is not good.’

The Chinese additive markers *yě* and *dōu* have two additive readings: (i) the simple additive reading, as in (74)a and (75)a; and (ii) the scalar additive reading, as exemplified in (74)b and (75)b (Forker 2016:74). In Standard English, while *also* (along with *too*, *as well*, and *either*) is responsible for the simple additive reading, *even* can be analyzed as an additive marker with an additional scalar component (König 1991:66–73). According to this analysis, the associate of the additive is the least likely candidate among the set of alternatives for which the proposition holds. For example, (74)c asserts that the emperor was the least likely person to go to the front line. The same applies to (75)b, which presupposes that Chinese is the least likely area in which he performs poorly. In both cases, the item marked by the scalar additive is characterized as unexpected or surprising, perhaps contradicting some previous expectations (Forker 2016:74).

The semantic functions of Chinese *yě* and *dōu* match that of CSE *also*, as they can give rise to both the additive reading and the additive scalar reading. Moreover, the use of CSE *also* mirrors that of its Chinese counterparts in negative contexts. However, neither *yě* nor *dōu* appears in sentence-final position. They only occur in preverbal position, irrespective of the sentence position of their focus (Paris 1989; Chao 1968:780).

5.2.2 The frequency of *also* and its preferred sentence position

Likely CSE *already*, there are significant differences between CSE *also* and Standard English *also* in terms of frequencies of occurrences and preferred sentence position (Bao 2005; Siemund and Li 2017). First of all, as with *already*, the ratio of *also* per thousand words is exceptionally higher in CSE than in British English. In Siemund and Li (2017:22), the ratio of *also* in ICE-SG (2.27 ptw) is more than twice as high as in ICE-GB (0.96 ptw). Secondly, differences regarding individual frequency can be observed. Table 5.4 shows that the range in the Singapore data extends from 6.6 to 10.7, while the range in the British data is notably lower in comparison to the Singapore data, ranging from 3.4 to 4.3. Here the list is restricted to the five speakers with the highest ratios, basing the calculation on contributions above one thousand words.

ICE-SG ID	also.ptw	ICE-GB ID	also.ptw
<icesg-s1b-057:1\$c>	10.699	<icegb-s1b-014:1\$a>	4.278
<icesg-s2b-032:1\$a>	7.66	<icegb-s1b-050:1\$b>	4.207
<icesg-s2a-044:1\$a>	6.778	<icegb-s1b-056:1\$b>	3.835
<icesg-s2a-024:1\$a>	6.7	<icegb-s2a-027:1\$a>	3.573
<icesg-s2b-001:1\$a>	6.645	<icegb-s1a-082:1\$a>	3.445

Table 5.4: The five speakers with highest *also*-ratios in ICE-SG and ICE-GB (adapted from Siemund and Li 2017:23)

As with *already*, CSE *also* prefers sentence-final position. Table 5.5 displays the distribution of *also* in different sentence positions (initial, medial, and final). The “Other” column shows the frequency of *also* with the universal and concessive interpretation as well as formulaic expressions as in *Also can*. While the position of *also* is predominantly in sentence-medial position in all examined registers, CSE *also* prefers sentence-final position in private dialogue. Only in more formal situation such as public dialogue, monologue and written register does CSE *also* appear more frequently in sentence-medial position.

	Initial	Medial	Final	Other	Total
PRIVATE DIALOGUE					
GB	0.24	0.77	0.00	0.00	1.01
SG	0.45	1.30	1.47	0.36	3.58
PUBLIC DIALOGUE					
GB	0.18	1.61	0.06	0.00	1.85
SG	0.33	4.39	0.28	0.03	5.03
MONOLOGUE					
GB	0.23	2.63	0.03	0.00	2.89
SG	0.29	4.70	0.06	0.01	5.06
WRITING					
GB	0.16	3.21	0.01	0.00	3.38
SG	0.29	4.76	0.01	0.00	5.06

Table 5.5: Frequency of *also* per text in ICE-GB and ICE-SG (Bao and Hong 2006:110)

It is worth mentioning that these findings on *also* echo research in Indian English (IndE). Fuchs (2012), for example, studies the distribution of *also* based on the Indian and British components of ICE, and finds that the usage of *also* significantly differs in IndE from BrE. Like CSE *also*, IndE *also* is often used in negative contexts, and has developed a new presentational usage, i.e. *also* marks its focus as new, un presupposed information. Besides, the frequency of these innovative features increases when the register moves from formal to informal, from written to spoken (Fuchs 2012:48).

5.2.3 Summary

To a very large extent, the differences between CSE *also* and Standard English *also* mirror the differences between CSE *already* and Standard English *already*. Semantically, CSE *also* has acquired subtle grammatical functions when used with universal quantifiers and the concessive adverb *even*, which is modeled on Chinese 也 *yě* and 都 *dōu*. The overall frequency results show a significantly higher ratio of *also* in CSE compared with BrE. Like CSE *already*, CSE *also* tends to have a higher frequency in sentence-final position than in

BrE, especially in private dialogue. Moreover, CSE *also* is often used in negative polarity contexts, which is rare in BrE.

Curiously, although *also* in CSE often occurs in sentence-final position, the presumed Chinese substrates *yě* and *dōu* show no preference in sentence-final position. Interestingly, *also* in IndE exhibits similar features.

5.3 Ever

The adverb *ever*, as will be described in the following sections, has the meaning ‘at least once’ in CSE, and can be used in an affirmative context. Its function is to mark the experiential aspect, which is expressed in Standard English in present perfect or simple past. However, the functions of *ever* in Standard English are more lexical than grammatical. Its core meaning is existential, which is largely restricted to negative polarity contexts (see Section 5.3.1). Again, similar to *already*, and *also*, it is observed that CSE *ever* replicates a grammatical function of the Chinese substrates, e.g. 过 *guò* ‘to pass’, which is an experiential perfect marker (Bao 2005; Ho and Wong 2001). In the next subsection, the use of *ever* in Standard English will be described; Section 5.3.2 will focus on the experiential marker *ever* in CSE in relation to the Chinese substrate *guò*, and Section 5.3.3 will discuss the possibility of CSE *ever* being derived from *never* by backformation. Section 5.3.4 will summarize the findings.

5.3.1 Standard English *ever*

Like *already*, Standard English *ever* belongs to one of the “polarity sensitive items” (Huddleston and Pullum 2002:60). Yet, unlike *already*, which is positively oriented (see Section 5.1.1, p. 147), *ever* is negatively oriented, i.e. it is one of the “negative polarity items” (NPIs). This means that *ever* often occurs in negative contexts and can be found in a range of other polarity contexts (e.g. if-questions), but not in simple affirmative sentences (see Israel 1998:30).

Among studies on Standard English *ever* in topics of polarity sensitivity (Israel 1995, 1998; Ladusaw 1980, 2008; Van der Wouden 2002), Israel (1998) focuses on a systematic construction of its meanings and distributions. He classifies *ever* into three basic types, namely (i) existential, meaning ‘even once’, or ‘at any time’, as in (76)a; (ii) universal, meaning ‘constantly’ or ‘at all times’, as instantiated in (76)b; and (iii) emphatic, which refers to those cases of *ever* in non-quantificational emphatic uses, as exemplified in (76)c (Israel 1998:29–34).

- (76) a. Existential: Glinda hasn’t *ever* robbed a liquor store. (Israel 1998:30)
 b. Universal: Let...nappy Ale be *ever* free to strangers that do come and go.
 [1686 Loyal Garland, ed. 5] (cited in Israel 1998:32)
 c. Emphatic: Was I *ever* shocked! (Israel 1998:29)

Existential *ever*, as exemplified by (76)a, accounts for 85% of the tokens found in Israel’s (1998) study. Therefore, Israel (1998) asserts that the existential meaning is the core meaning carried by *ever*, which may also give rise to the emphatic and universal readings. In (76)a, *ever* is existential because it “indicates a single, indefinite and temporally unspecified instance of a propositional relation” (Israel 1998:33). In other words, it carries the notion of something happening at least once at an unspecified time. Existential *ever* is a negative polarity item, as it is considered as ungrammatical in a simple affirmative sentence, e.g. **Glinda has ever robbed a liquor store*, if we negate (76)a. Yet, sentence like this is grammatical in CSE, as we will see later in Section 5.3.2.

The universal usage, meaning ‘constantly’ or ‘at all times’ is used in a positive environment, though this usage, as in (76)b is outdated. Only a few relics have survived, including its use as an adjectival or adverbial modifier, e.g. *ever increasing*, the comparative *ever closer*, and the continuous *ever since*. Israel (1998:33) reports that the universal usage accounts for just 1% of the total corpus.

The third usage, the emphatic usage of *ever*, is considered as an extension from its existential and universal uses, as both function as a quantificational adverb and add emphasis when occurring in a sentence (Israel 1998:34). The emphatic usage of *ever*, however, only

refers to those cases of *ever* in non-quantificational emphatic uses, such as in collocation with the degree modifier *so* (e.g. *ever so*), with inversion exclamatives, as in (76)c, and in rhetorical *wh*-questions (e.g. What *ever* did you expect?). Like the universal usage, the emphatic usage is unproductive, accounting for less than 1% in Israel's (1998) study.

Apart from the above three usages, *ever* can be used as a derivational suffix, which is attached to *wh*- words to form a series of free indefinite pro-forms such as *whatever*, *whoever*, *whenever*, etc. (Israel 1998:35).

To sum up, the core meaning of *ever* is its existential meaning, which is restricted to negative contexts meaning 'even once'. We can also observe from all of the above examples that *ever* takes sentence-medial position, except for the universal usage where *ever* occurs before an adjective modifier, the comparative, or the continuous preposition *since*, and for the emphatic usage before the emphatic focus.

5.3.2 The experiential aspectual marker *ever* in CSE

Different from Standard English *ever*, whose functions are lexical, meaning 'even once' or 'at any time', and which is limited to negative contexts, the CSE *ever* functions as a perfective marker "which emphasizes the experience associated with the completed event" (Bao 2005:243), and it is often used in affirmatives. Thus, the perfective marker *ever* is often referred to as the "experiential marker" (see Bao 2005; Leimgruber 2013). Consider the following examples from Hong and Wong (2001:81):

- (77) a. We *ever* come across a case, half year the new battery conk off *already*.
'We have come across a case where the new battery went dead in half a year.'
- b. I *ever* met some customer like that.
'I have met some customer like that.'
- c. This share *ever* hit forty dollars.
'This share was once forty dollars. /*This share has hit forty dollars.'

Bao (2005) observes that though *ever* can be translated into the English perfect, as exemplified by (77), there are differences between the English translations and the original

examples. While CSE *ever* denotes that the underlying “experience” that happened once in the past no longer happens at the time of the utterance, the English perfect does not implicate that contrast between the state or event in the past and that at the time of the utterance. For example, the implication of the CSE *ever* in (78)a implies that John used to love Mary, but he does not love Mary any longer.

- (78) CSE:
- a. John *ever* love Mary.
‘John has loved/loved Mary.’
 - b. The wall *ever* white.
‘The wall has been/was white.’

Bao (2005) suggests that there is no such implication in the English simple past/perfect interpretation. Therefore, while the English perfect and simple past are ambiguous in that the states or activities may have ended or have continued, the experiential *ever* in CSE is not equivocal. The following temporal schema presents the difference in terms of aspectual meanings:

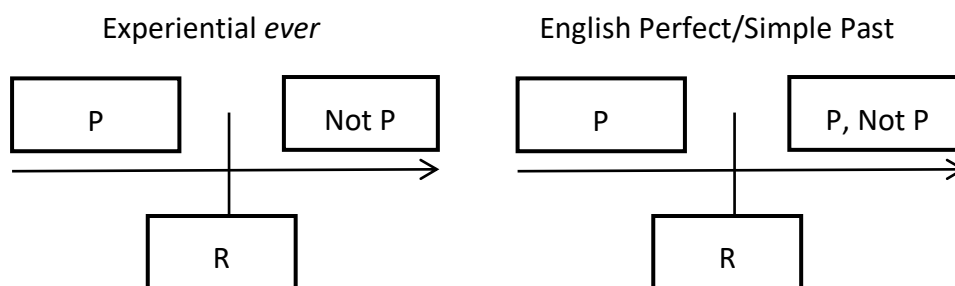


Figure 5.2: Experiential *ever* and the English simple past (adapted from Bao 2005:244)

As mentioned earlier, the inchoative *already* in CSE differs from the English simple past/perfect in that CSE *already* marks a new state. The experiential *ever*, on the other hand, marks that a state that happened in the past at some indefinite time fails to prolong to the present. The inchoative *already* and the experiential *ever* are different, yet complementary, in their aspectual interpretations. As described by Bao (2005:244), “*already* asserts the

existence, and *ever* the end, of a given state at the time of reference, or at the present time.”

Consider the following comparison between inchoative *already* and experiential *ever*:

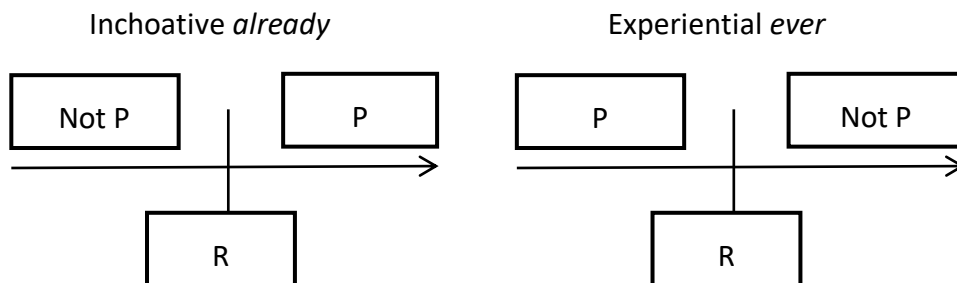


Figure 5.3: Temporal schema of inchoative *already* and experiential *ever*

Since experiential *ever* is used in CSE to refer to something in the past, it is not surprising that adverbials such as *before* or *last time* often co-occur with *ever*, as shown in (79):

- (79) CSE *ever* (Ho and Wong 2001:82)
- a. I *ever* try this type of fruit before.
‘I have tried this type of fruit before.’
 - b. I *ever* been out with her before.
‘I have been out with her before (but not anymore).’
 - c. Last time I *ever* bought something on sale.
‘I have bought something on sale in the past.’

As mentioned earlier, the use of experiential *ever* is often associated with the Chinese experiential aspect marker 过 *guò*. Li and Thompson describe the meaning of *guò* as to mark “an event [that] has been experienced with respect to some reference time” (1981:226). The reference time is usually left unspecified, and therefore, *guò* signals that the event has been experienced at least once in the indefinite past. Consider the following sentence in Mandarin Chinese:

- (80) 我 学 过 法语
wǒ xué guò fǎyǔ
I study Exp. French
‘I have studied French.’

Parallel to the experiential *ever* in CSE, sentence (80) implies that (i) the speaker learned French at least once in the past, and (ii) the speaker ceased to learn French after that reference point. Although CSE *ever* mirrors the aspectual meaning of the experiential *guò* in Mandarin Chinese, the position of *ever* remains the Standard English preverbal position, whereas Chinese *guò* is postverbal.

5.3.3 *Ever* in affirmative responses to polar interrogatives

Ziegeler (2015:152) argues that CSE *ever* could be derived from *never* by backformation, when CSE speakers misinterpreted the meaning of *never* as the emphatic minimizer ‘not once’ instead of ‘not any possible time’. This argument is evident in the discussion in Ho and Wong (2001), where the use of *ever* is found in affirmative responses to polar questions in CSE. They point out that while negative responses to polar interrogatives in CSE are similar to Standard English, their affirmative answers are very different. Compare the following examples:

- (81) Standard English (adapted from Ho and Wong 2001:80)
 A: Have you *ever* been to China?
 B1: (Yes,) I have. (affirmative)
 B2: No, never. / No, I haven't. (negative)
- (82) CSE (adapted from Ho and Wong 2001:80)
 a. A: You *ever* go to China (or not)?
 B1: (Yes,) *ever*. (affirmative)
 B2: No, never. / No, I haven't. (negative)
 b. A: Your husband *ever* bring fish home to eat or not?
 B: *Ever*.

In Standard English, the auxiliary *have* in the affirmative answer echoes the question “have you...”, as shown in (81). The affirmative answer in CSE, as exemplified in (82) seems to be derived from a backformation process: from the negative answer “No, never” to the affirmative answer “Yes, ever”. Note that *ever* in Standard English does not usually occur in positive sentences, so a complete answer to the if-interrogative in (81) would be “Yes, I

have been to China” (see Section 5.3.1, also Ho and Wong 2001:80). The affirmative response in (82)a omits the auxiliary *have* together with the subject *I* and replaces it with the adverb *ever*.

5.3.4 Summary

We can conclude from the above discussion that CSE *ever* differs from Standard English *ever* in three different ways. Firstly, CSE *ever* functions as an experiential aspectual marker, which corresponds to the English simple tense or present perfect, while the function of *ever* in English is more lexical than grammatical, and its core meaning is existential. Secondly, experiential *ever* in CSE is often used in positive affirmatives, whereas existential *ever* in Standard English is restricted in its negative contexts. The semantic difference between experiential *ever* in CSE and existential SE *ever* is rather subtle and is related to their contexts: CSE *ever* means that an event has been experienced at least once with respect to some reference time in positive affirmatives. In contrast to CSE *ever*, the main usage of Standard English *ever* is existential, meaning ‘even once’ or ‘at any time’. Existential *ever* is restricted in negative contexts, with the function of negating or questioning the existence of an event/activity.

Clearly, CSE *ever* has experienced a functional shift from a negative polarity item in SE to an experiential aspectual marker due to substrate interference from the Chinese *guò*, which exhibits the same aspectual meaning. It is noteworthy that the Chinese experiential marker *guò*, however, does not correspond to the universal *ever* meaning ‘always’, ‘constantly’ and ‘at all times’, nor to the emphatic *ever* before an emphatic focus. Nevertheless, it is also possible that *ever* in positive affirmatives is derived from the backformation of *never* – from *never* meaning ‘not once’ to *ever* meaning ‘once’.

5.4 One

Similar to *already*, *also*, and *ever*, the CSE *one* has experienced a functional extension as a result of the Chinese influence. Not only has the variant *one* in CSE inherited all the main functions of Standard English *one* (i.e. numeral and pronominal), but it also functions as a relative clause pronoun and an emphatic marker (Alsagoff and Ho 1998; Bao 2011; Gupta 1992b; Wong 2005).

5.4.1 Numeral and pronominal *one*

One in Standard English has two main functions, which are: (i) numeral, and (ii) pronominal (also *prop word*, see Jespersen 2003). The numeral-related function of *one* in CSE is identical to Standard English, as both can appear in the frames of [one + N]_{NP} and [one]_N, as exemplified in (83) (Teo 2014:847).

- (83) a. I have *one* hat.
b. *One* is a small number.

The pronominal function in Standard English, on the other hand, can be classified into two types: “anaphoric *one*” and “independent *one*” (Jespersen 2003:51–53). Consider the following examples:

- (84) Anaphoric *one*
A: Which hat do you want to wear tonight?
B1: The red *one*.
B2: The *one* over there. / The *one* that you like.
B3: That *one* in the box. / One of those nice *ones* that we bought in Singapore.
- (85) Independent *one*
My little *one* is sick again.

The function of anaphoric *one* is to replace a substantive which has just been mentioned before in order to avoid repetition. For example, in (84) in which hats are being discussed,

the red *one* means ‘the red *hat*’. In addition, *one* can be used between a definite article and a prepositional phrase, as exemplified by speaker B2 in (84), or between a demonstrative, and a phrase or a clause, as in B3. Independent *one*, on the other hand, can replace a substantive which is vague or general (Jespersen 2003:52). In (85), *one* does not require any further contexts, as we can assume that *one* refers to a child.

Not all languages have a form equivalent to English *one*. For example, German uses a nominalized adjective instead of the pronominal *one*, e.g. *the red one* is translated into *der Rote* in German (Sadock 1991:41) In Chinese, the particle *de* can serve the pronominal function like that of the English *one* (see Chapter 6.5). What is worth mentioning is that pronominal *one* in English can be omitted in some expressions, e.g. from *the red one* to *the red*, whereas pronominal *de* in Chinese cannot be omitted.

With respect to the pronominal *one* in CSE, it functions similarly to the pronominal *one* in Standard English, though slight differences have been observed. Like pronominal *one* in Standard English, it generally follows the frame of A-*one* (adjective-*one*, e.g. nice *one*). Yet, CSE *one* can include nominal words/phrases, following the frame of N-*one* (noun-*one*, e.g. silk *one* ‘a dress made of silk’), as exemplified in (86). The pronominal *one* in CSE also collocates with possessive pronouns, appearing in the frame of pronoun-*one* (e.g. my *one* ‘something that belongs to me’), which is ungrammatical in Standard English, as exemplified by the pair of sentences in (87) (Bao 2009:340).

- (86) CSE
That one is something like the yellow type-lah. That time is a short, yellow type *one*. [OHI-001953-LAS]
- (87) a. CSE: Last time she accommodates her sleeping times to Lucy *one*, you know. (Gupta 1992a:330)
b. Standard English: *Paul’s students are taller than John’s *ones*.
(Panagiotidis 2003:282)

The phrase *the yellow type one* in Example (86) from OHI is idiosyncratic because the modifier of *one* here is a noun phrase, yet English *one* is normally modified by an adjective (e.g. a new *one*, also see Panagiotidis 2003:282). In (87)a, *Lucy one* ‘Lucy’s’ refers to the

sleeping time of Lucy. Such a phrase is ungrammatical in Standard English as English does not tolerate possessive *one*. According to Panagiotidis (2003:282), the reason for a lack of expressions like *John's ones* as shown in (87)b in Standard English is due to “lexical economy” as the elliptical *John's* and alternative possessive pronouns like *his* are available.

In short, the main functions of *one* in Standard English are pronominal and numeral. These two functions are inherited by CSE, though slight differences have been observed, such as *one* modified by a noun phrase and the co-occurrence of the possessive pronoun and *one*. However, this is not the whole story of *one* in CSE. Apart from the pronominal and numeral functions, CSE *one* possesses the nominalizing and emphatic functions, as will be discussed in the following sections.

5.4.2 *One* as a nominalizer in CSE

Several related terms have been used to discuss the variant *one* in CSE, such as “relative clause pronoun” (Alsagoff and Ho 1998), “reifier” (Gil 2003), and “nominalizer” (Teo 2014; Wee and Ansaldo 2004). It is difficult to agree on a term because the variant *one* in CSE overlaps with the pronominal function of *one* in Standard English, yet it occurs in a different structure, as exemplified in (88).

- (88) CSE (Alsagoff 1995:85)
- a. The boy pinch my sister *one* very naughty.
‘the boy who pinches my sister is very naughty.’
 - b. The cake John buy *one* always very nice to eat.
‘The cakes that John buys are always very delicious.’

From the above examples, we can observe that CSE *one* follows a relative clause, whereas relative clauses in Standard English are introduced by relative pronouns such as *who* and *that*, as illustrated by the Standard English translations in (88). Therefore, Alsagoff and Ho (1998) define the variant *one* in CSE as a relative pronoun, even though it is structurally different from relative pronouns in Standard English.

To avoid syntactic characterization, Gil (2003:480) creates the term “reifier” to describe CSE *one*.⁴⁴ Gil (2003) claims that reifier *one* is used together with different constructions such as possessives, property words, locative expressions, and event expressions to form various NPs. Consider the following examples from Gil (2003:480):

- (89) a. Ah Chew buy Jamil *one*. ‘Ah Chew bought the one from Jamil.’
 b. Ah Chew buy expensive *one*. ‘Ah Chew bought the expensive one.’
 c. Ah Chew buy in Jurong *one*. ‘Ah Chew bought the one in Jurong.’⁴⁵
 d. Ah Chew buy yesterday Lisa choose *one*. ‘Ah Chew bought the *one* that Lisa chose yesterday.’

However, the term “reifier” does not suggest much about the functions of CSE *one*. More recent studies consider *one* as a “nominalizer”, whose function is to convert a non-nominal phrase into a noun phrase (Teo 2014:840; Wee and Ansaldo 2004:69). The term “nominalizer” comes from the functional descriptions of the Chinese counterpart 的 *de* (Liu 2003:74). Therefore, applying the term “nominalizer” not only solves the issue of the functional descriptions of CSE *one*, but also links it to the source of cross-linguistic influence from the Chinese nominalizer *de*. The examples in (89) find their ready one-to-one translations in Chinese, as shown in (90):

- (90) a. 阿初 买 吉姆 的
 Ā chū mǎi jímǔ de
 Ah Chew buy Jamil NOM
 ‘Ah Chew bought Jamil’s.’
 b. 阿初 买 贵 的
 Ā chū mǎi guì de
 Ah Chew buy expensive NOM
 ‘Ah Chew bought the expensive one.’

⁴⁴ The term “reifier” comes from “reification”, which is a calque from *Verdinglichung* in German (Brown 1987:199).

⁴⁵ Jurong /dʒu:rɒŋ/ (Chinese: 裕廊 Yùláng, Malay: Jurong, Tamil: ஜூரோங்) is a geographical region located at the south-westernmost point of the West Region of Singapore.

- c. 阿初 买 在 裕廊 的
 Ā chū mǎi zài yùláng de
 Ah Chew buy in Jurong NOM
 ‘Ah Chew bought the one in Jurong.’
- d. 阿初 买 昨天 丽萨 选 的
 Ā chū mǎi zuótiān lì sà xuǎn de
 Ah Chew buy yesterday lisa choose NOM
 ‘Ah Chew bought the *one* that Lisa chose yesterday.’

Examples in (90) illustrate that *de* can convert any phrase (e.g. determiner phrase, adjective phrase, prepositional phrase) or clause that it attaches to into a noun phrase. As manifested by (90)a and (90)b, there is an overlap of Chinese *de* as a nominalizer with the pronominal function of *one* in English. As mentioned earlier, “Jamil *one*” would be considered as ungrammatical in Standard English, as English does not allow the co-occurrence of the possessive pronoun and *one*. Yet, 贵的 *guì de* in (90)b, translated as ‘an expensive one’ or ‘the expensive one’ corresponds to the pronominal *one* in Standard English. However, the other usages of *de* as a nominalizer, as shown in (90)c and (90)d do not mirror that of Standard English *one*. In (90)c, *de* converts the prepositional phrase *in Jurong* into a noun phrase *the one in Jurong*. And in (90)d, it converts the non-nominal clause *yesterday Lisa choose* into a noun phrase *the one that Lisa chose yesterday*. Clearly, the term nominalizer unifies both the pronominal and the relative clause marking functions of CSE *one*.

Instead of giving a functional term for the variant *one* in CSE, Bao (2009) analyzes CSE *one* within a structural framework. He gives the nominalizer *one* in CSE the structural frame of XP-*one*, and includes prepositional phrases (PP), verb phrases (VP), and clauses (S) under the label XP. The main reason to group all phrasal categories under the same umbrella is to avoid the problem of multiple analyses (Bao 2009:341). Consider the following examples:

- (91) a. Showing in Cathay *one*
 (i) [NP[VP showing in Cathay] one]
 (ii) [NP[S showing in Cathay] one]

- b. Those wear black *one*
(i) those [NP [VP wear black] one]
(ii) [NP [S those wear black] one]

As shown in (91), there are different ways to analyze the structure of a given string, yet *one* transforms any given string into an NP. In (91)a, for example, *showing in Cathay* can be considered either as a VP or as a clause. Likewise, *those wear black* in (91)b can be grouped into a clause, which is nominalized by *one*. Another way to analyze it is to group the verb *wear* and the noun *black* into a VP, which joins *one* and transforms into an NP.

5.4.3 Emphatic *one*

Besides functioning as a nominalizer, *one* manifests a pragmatically oriented function. As exemplified in (92), *one* expresses a pragmatic or “interactional” meaning that cannot be communicated otherwise (Wong 2005:251).

- (92) CSE emphatic *one* (Lim 2004:69)
a. You always late *one*!
b. The car very expensive *one*!
c. The coffee nicer than the tea *one*!

One in such usage is referred to as “singulative *one*” (Lim 2004), or “emphatic *one*” (Bao 2009). Other terms such as “assertive particle” (Gupta 1992), “pragmatic particle” (Wong 2005), and “contrastive focus marker” (Teo 2014) were also used to analyze the pragmatic functions of *one*. Following Bao (2009), I adopt the term of “emphatic *one*”, which considers independent *one* as a marker of emphasis (Bao 2009:340). In (92), *one* emphasizes an entity (e.g. *late*, *expensive*, and *nice*) as particularly salient in a category. Unlike nominalizer *one*, emphatic *one* does not form an NP with any preceding words or phrases. Therefore, emphatic *one* can be omitted without making the sentence structurally incomplete. In other words, emphatic *one* is non-obligatory from a syntactic perspective.

The surface structure of emphatic *one* resembles that of the nominalizer *one*. Thus, the categorical difference needs to be resolved through intonation and context. Bao (2009)

gives emphatic *one* the same structural frame XP-ONE that he assigns to the nominalizer *one*. Yet, Bao (2009) underlines that while nominalizer *one* forms an NP with the preceding modifier, the emphatic *one* is attached to a phrase or a sentence, e.g. [_s [_s The car very expensive] one] ‘The car is very EXPENSIVE!’.

One as a contrastive focus marker

Though describing *one* as an emphatic marker generalizes the function of *one*, it does not explain why *one* has its pragmatic function in the first place (Teo 2014). Teo (2014:848) argues that *one* serves as a “contrastive focus marker”. He claims that emphatic *one* exhibits the property of a contrastive focus marker, which is associated with the notion of “exhaustibility” and “exclusivity” (Teo 2014:849).

- (93) ?He study Chinese in Beijing *one*, but also study in Shanghai *one*.
 ?‘It’s the case that he studied Chinese in Beijing, but also the case that he studied in Shanghai.’ (Teo 2014:849)

The sentence in (93) is problematic because the clause before *one* – *he studied Chinese in Beijing* – is under contrastive focus, which excludes any other alternative proposition. In other words, any other alternative, e.g. *he studied in Shanghai*, would be false. Therefore, (93) is contradictory: the first clause that it is only in Beijing that he studied Chinese, is in conflict with the second clause that it is only in Shanghai that he studied Chinese.

Another evidence provided in Teo (2014) for the argument of *one* as a contrastive focus marker is its compatibility with discourse markers in a single utterance. He asserts that *one* as a contrastive focus marker can co-occur with a discourse marker in an utterance, yet two discourse particles at the end of an utterance is regarded as ungrammatical (Teo 2014:849). Compare the pair of sentences in (94):

- (94) a. Can walk *one lah*. (Wong 2005:249)
 ‘It is within walking distance.’

- b. *Eat already *lah lor*. (adapted from Wee 2004a:125)
 ‘I have eaten.’

In (94)a, *one* functions as a contrastive focus marker, which puts *walk* under its focus scope, while *lah* serves as an assertive marker to persuade the interlocutors to walk with the speaker. Such expression is grammatical and idiomatic in CSE. In contrast, (94)b is ungrammatical, as one single utterance does not allow more than one discourse particle (Teo 2014:849).

However, the notion of exhaustibility (the fact that all alternatives are implied or presupposed to be false) is a cross-linguistically stable property of clefts, but not of a contrastive focus marker (see Rochemont 1986:127–60). Scholarly circles are increasingly converging on the point that clause-final 的 *de* – the Chinese counterpart of emphatic *one* in CSE – has the same distribution as 是...的 *shì ...de* clefts (Hole 2011; Hole and Zimmermann 2013; Paul and Whitman 2008). Such proposal considers that clause-final *de* is a special type of *shì ...de* clefts where *shì* is dropped, as exemplified in (95) (Hole 2011:1708):

- (95) a. 是 斯蒂芬 吃 的
 shì sīdīfēn chī de
 be Stephen eat EMP
 ‘It is Stephen who ate it.’
- b. 斯蒂芬 吃 的
 sīdīfēn chī de
 Stephen eat EMP
 ‘It is Stephen who ate it.’/ ‘The one(s) he ate.’

Example (95)a is a cleft and the focus is marked by copula *be* 是 *shì*, whereas the example without *shì* in (95)b also has the cleft interpretation. Yet (95)b is ambiguous as it could also be interpreted as a relative clause marker/nominalizer ‘the one(s) he ate’. Furthermore, contrastive foci can be marked by mere *shì* in Mandarin without *de*, as shown in (96). This shows that it is *shì*, not emphatic *de*, that serves as a contrastive focus marker.

- (96) 是 猫 偷 了 鱼
 shì mǎo tōu le yú
 be cat steal ASP fish
 ‘The CAT stole the fish.’/ ‘It is the cat that stole the fish.’

Therefore, this study follows Bao (2009) in classifying independent *one* as an emphatic marker instead of a contrastive focus marker.

5.4.4 Frequency of *one* according to its functions

From the theoretical discussions above, we can see that *one* as a relative clause pronoun and as an emphatic marker (represented by the structural frames XP-*one* and XP-ONE) are unique usages in CSE. The frequency results in Bao (2009) confirm this, as XP-*one* and XP-ONE only manifest themselves in ICE-SG, but not in ICE-GB. As shown in Table 5.6, ICE-GB does not exhibit such usages.

	frame	ICE-SG		ICE-GB	
		token	percent	token	percent
a.	A- <i>one</i>	156	28.1	123	38.6
b.	N- <i>one</i>	37	6.7	18	5.6
c.	P- <i>one</i>	3	0.5	5	1.6
d.	XP- <i>one</i>	6	1.1	0	0.0
e.	XP-ONE	74	13.3	0	0.0
f.	Others	280	50.4	173	54.2
	Total	556	100.0	319	100.0

Table 5.6: Comparison of counts of *one* in the Private Dialogue sub corpora between ICE-SG and ICE-GB (Bao 2009:344)

With regard to the pronominal usages (represented by the frames of A-*one*, N-*one* and P-*one*), there are only marginal differences. Parts of the pronominal frames (N-*one* and P-*one*) and the relative marker frame (XP-*one*) manifest low productivities in both ICE-GB and ICE-SG. However, these are well-established usages in Chinese (see Section 6.5). Bao (2009)

asserts that the low productivities of N-*one*, P-*one* and XP-*one* is due to the violation against the grammatical constraints of the lexifier language, which discourages a feature to be transferred. However, emphatic *one* also violates the morphosyntactic rules of English, but exhibits a much higher productivity (13.3 percent) than N-*one* (6.7 percent), P-*one* (0.5 percent), and XP-*one* (1.1 percent). It seems to suggest that emphatic *one* has been grammaticalized into a component of the pragmatic system. In this process, it becomes detached from its pronominal meaning and breaks free from the effect of the morphosyntactic constraints in Standard English.

5.4.5 Summary

We can conclude that Standard English *one* has a notably narrower scope of usage than CSE *one* since CSE *one* has extended from its pronominal functions to other functions such as the nominalizer function, the relative clause marker function, and the emphatic function. It is important to note that the nominalizer function overlaps with the relative clause marker function. The term nominalizer, which Chinese linguists often use, unites both the pronominal and relative clause marker functions and is adopted by previous CSE scholarly literature on *one*. It is argued that such an intersect is an important reason why Chinese *de* is regarded as the equivalent of *one* in CSE. The usages of the pronominal *one* in CSE do not differ substantially from Standard English, except that CSE *one* can follow nominal words (e.g. N-*one* silk *one* ‘a dress made of silk’), and possessive pronouns (e.g. my, your, his/her), which are ungrammatical in Standard English. The relative clause marker function and the emphatic function, however, are exclusive feature of CSE *one*.

5.5 Chapter summary and conclusion

This chapter discusses the use of *already*, *also*, *ever*, and *one* in CSE. We can observe a functional extension in all of these four linguistic variables and see that such a functional extension is due to cross-linguistic influences from Chinese. On top of functioning as a

phasal polarity expression, which associates two temporal reference points, CSE *already* has acquired aspectual functions, marking the completive, inchoative and prospective aspects. It is used in negative contexts, which is atypical in Standard English. *Also* has acquired subtle grammatical meanings when used with universal quantifiers (e.g. *all*, *everything*, *everyone*, etc.) and concessive *even*. Unlike Standard English *also*, CSE *also* is often used in negative contexts. *Ever* is used to express existential meaning in negative contexts in Standard English, whereas CSE *ever* is an experiential aspectual marker which means ‘to experience at least once’. And the pronominal *one* in Standard English, which is used to replace a general, vague substantive, or an already-mentioned noun to avoid repetition has extended its usages in CSE as a nominalizer and as an emphatic marker.

In addition to these functional extensions, CSE *already*, *also*, and *one* prefer different positions in a clause/sentence compared with their counterparts in Standard English. *Already*, *also*, and *ever* in Standard English are generally placed in mid-position between the subject and the main verb, or after the first auxiliary or a modal verb when there is more than one verb. In contrast, CSE *already* and *also* prefer phrase- or sentence-final position. Though CSE *ever* prefers its position directly in front of an actual verb, just as in Standard English, the actual verb is commonly used in its bare, non-inflectional form. Besides, CSE *ever* is used in affirmative responses to polar questions, which implies a backformation derived from *never*.

Interestingly, the syntactic positions of *already*, *also*, *ever*, and *one* do not completely mirror their counterparts in the Chinese substrates. For example, the equivalents of English *also* – 也 *yě* and 都 *dōu* – occur in preverbal position, exactly like English *also*. However, CSE *also* prefers sentence-final position. In addition, the preverbal position of CSE *ever* does not match that of the substrate 过 *guò* in Chinese. It seems that these transferred items were remodeled to fit the morphosyntactic structures of English. The results are in line with general linguistic constraints in contact situations, in which a wholesale of both semantic and syntactic features is not expected. As put by Siemund (2008:7), “probably [a] more realistic view held in language contact research is that whatever kind of material is transferred in a situation of language contact, this material necessarily experiences some sort

of modification through contact”. Bao (2015:4) also pointed out that “[e]ven though transfer targets the grammatical system, the contact language is not point-by-point identical with the substrate language”.

The next chapter explores the usages of 了 *le*, 过 *guò*, 也 *yě* / 都 *dōu*, and 的 *de* in Chinese. The purpose is to have a more solid understanding of the substrate sources in Chinese, which provide the functional extensions of CSE *already*, *also*, *ever*, and *one*. Issues concerning the meanings, functions, and morphosyntactic features of the substrates will be addressed, so that we can find out which functions along with their morphosyntactic features were transferred in CSE, and which were blocked. In addition, we will examine the grammaticalization processes of these expressions in Chinese. In this regard, we can reveal whether CSE speakers replicate the grammaticalization process which has taken place in Chinese (see Section 4.4 on replica grammaticalization).

6 The Chinese Substrates

This chapter focuses on developing a perspective from the Chinese counterparts of the four expressions *already*, *also*, *ever* and *one*. It highlights the existing link between CSE and the Chinese substrates and brings their differences under the same scope. Section 6.1 introduces the aspectual system of Chinese language, which sets the foundation for the later exploration of the Chinese aspectual markers: 了 *le* and 过 *guò*. Section 6.2 is concerned with the grammaticalization of 了 *le*, which demonstrates that verbal *le* (V-*le*) and sentence-final *le* (S-*le*) are two different categories that originate from different expressions. Section 6.3 discusses the aspectual marker 过 *guò*, which marks the experiential aspect. Section 6.4 describes the additive markers 也 *yě* and 都 *dōu*, which are possible sources of CSE *also*. Section 6.5 discusses the most widely used particle 的 *de* – the counterpart of nominalizer *one* and emphatic *one* – with its various functions (e.g. as a “relativizing” marker linking different types of phrases and clauses, and as an emphatic marker).

It is important to mention that there exist differences in opinion in modern literature on the meanings and functions as well as classifications of these Chinese markers, but the goal of this chapter is not to spread the debate over issues related to functional categorization, but to provide an accurate description of these markers and highlight the similarities and differences between them and the CSE markers.

6.1 Aspectual system of Chinese

Tense and aspect are two different linguistic properties to view events in relation to speech time. Tense is about temporal location, which is indicated by inflectional suffixes on verbs in native varieties of English and many other European languages (Smith 1997). Chinese, on the other hand, is often described as “tenseless”, because it lacks overt morphology to mark tense (Lin 2010; Smith and Erbaugh 2001, Platt and Weber 1980). Instead, temporal location is given by adverbials and modal auxiliaries (Smith 1997:263). Yet, there are clear aspectual distinctions, which deal with some intrinsic parts of clause meanings associated

with the internal temporal constituency of a situation (also “aktionsart” or “a state/an event”) (Comrie 1976:5). These aspectual distinctions relate to the way in which an event or a state is regarded. An event or a state is viewed as a process consisting of a series of phases, including the beginning, the continuation, and the completion. Each of these phases corresponds to an aspect.

According to Comrie (1976), aspect can be classified as either “perfective” or “imperfective”. The perfective aspect views an event or a state from outside, denoting a complete situation with beginning, middle, and end (Comrie 1976:3). It is important to note that the notion of “completion” does not necessarily mean that the state or the event has ended, as highlighting the end of a state or an event is only one possible meaning of the perfective aspect.⁴⁶ The perfective aspect includes all phases of the situation as a whole, not just the end of the situation. For example, the inchoative is included in the notion of the perfective. On the other hand, the imperfective aspect looks at the situation from within, with explicit reference to its internal temporal constituency of the situation. Therefore, the imperfective is incompatible with situations that lack internal temporal structure.

In Chinese, aspect is represented by aspectual markers. Taking Mandarin Chinese for example, the perfectivity and imperfectivity are expressed by means of two different types of aspectual markers, namely the perfective 了 *le* and 过 *guò* on the one hand, and the progressive 在 *zài* and the continuous 着 *zhe* on the other hand. The classification of the aspectual markers in Mandarin Chinese is illustrated in Figure 6.1, based on Comrie’s distinction between the perfective and imperfective (see Comrie 1976; Chao 1968; Li and Thompson 1981). Except for 在 *zài*, which is preverbal, the other above-mentioned aspectual markers in Chinese are often referred to as aspectual suffixes because they often appear in post-verbal position.⁴⁷ The habitual aspect is commonly unmarked. One less-mentioned aspect, which is referred to as the tentative aspect or delimitative aspect (Chao 1968:205), is expressed in the form of verbal reduplication (VV).

⁴⁶ Instead, the term *completive* is often used to imply the completion of a situation (see *already* as a completive marker in Section 5.1.1).

⁴⁷ Note that the term “suffixes” here is different from suffixes on verbs in English, i.e. Chinese suffixes do not change forms of verbs, they simply follow the predicate that they modify.

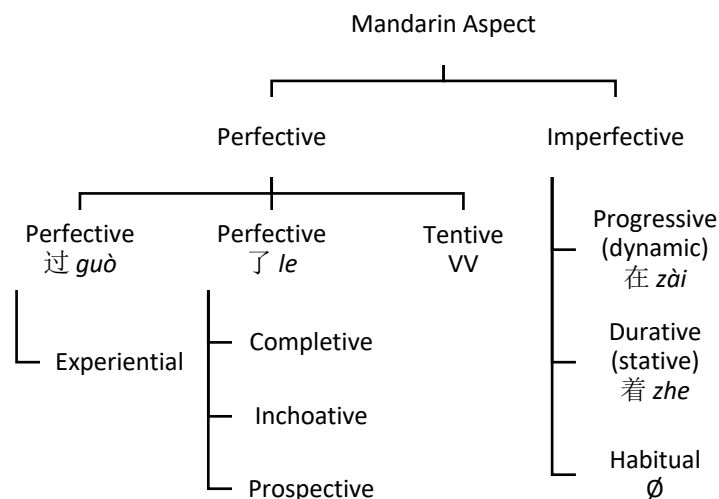


Figure 6.1: Schematic representation of Mandarin aspectual markers

Smith (1997) distinguishes two types of aspect: a universal situation aspect, and a language-specific viewpoint aspect. Situation aspect corresponds to what is traditionally called lexical aspect, or aktionsart, which denotes the relationship between the verb, its argument and the type of the event. It distinguishes situational types based on its internal temporal features, such as [\pm dynamic], [\pm durative] and [\pm telic], and subsequently classifies verb phrases into activity, achievement, accomplishment and state, plus a category of semelfactive (Smith 1997:17). Viewpoint aspect, which is realized grammatically, “enables the speaker to present the event talked about from a particular temporal perspective” (Smith 1988:230). Its subcategories correspond to Comrie’s (1976) classification of perfectivity and imperfectivity. According to Smith (1997:263), the major situation aspects are covert categories in Chinese, which are expressed by verbs and their arguments. On the other hand, there is a class of verbal complement suffixes that play an important role in conveying viewpoint aspect.

Following Smith (1997), Xiao and McEnery (2004) classify aspect into situation aspect and viewpoint aspect and argue that while the former is language-independent, the latter is language-specific (Xiao and McEnery 2004:20). They assert that previous research on Chinese aspectual system is not accurate enough (e.g. Chao 1968; Li and Thompson 1981). Therefore, their model further distinguishes two subcategories of viewpoint aspect, namely simplex and complex aspect. The subdivision of the simplex aspect follows the

classic classification of perfective and imperfective, whereas the complex aspect includes hierarchical combinations of two aspects in the simplex category. For example, a hierarchical combination of the actual aspect (represented by the completive *le*) and the delimitative aspect (represented by VV) in Chinese conveys a transitory event which has been actualized or completed (Xiao and MacEnery 2004:151). The complex viewpoints also apply to English, with the perfect progressive (*have+been+V-ing*) being a case in point. Their comprehensive overview of Chinese viewpoint aspect is illustrated in Figure 6.2. This section will not extend to describe each subcategories of the complex aspect, as the current research focuses on the aspectual markers *le* and *guò*, both falling in the category of perfective aspect.

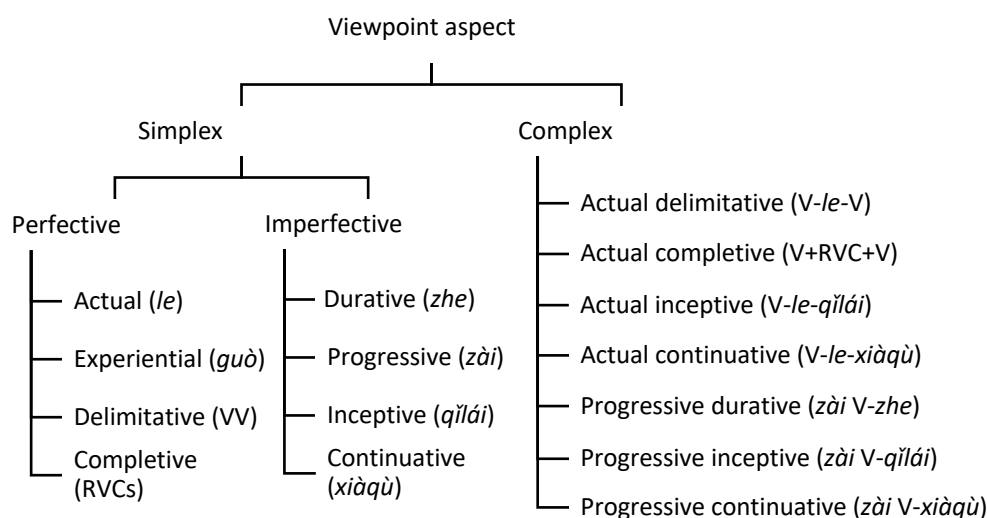


Figure 6.2: Chinese viewpoint aspect according to Xiao and McEnery (2004)

In Chapter 1, I introduced Bao's (2005) categorization of the Chinese aspectual system in comparison with that of English (see Section 1.2 and Section 4.7.4). He proposes that the aspectual meanings of *le* and *guò* have been transferred to CSE and relexified by English *already* and *ever* (see Section 5.1.1 and Section 5.3.2). Bao's (2005) classification of the Chinese aspectual system differs from Xiao and McEnery's (2004) mainly in the following aspects: (i) While Bao (2005) considers the inchoative/inceptive and the tentative are located at the same hierarchical level as the perfective and imperfective, Xiao and McEnery (2004) treat the inchoative/inceptive as belonging to the imperfect aspect (see Figure 6.2); (ii) The

inchoative/inceptive aspect is signaled by 起来 *qǐlái*, a “resultative verb complement” (RVC), lit. ‘up come’ in Xiao and McEnery (2004:110) while Bao (2004) attributes *le* to the inchoative/inceptive aspect; (iii) The “actual” aspect, which presents an actualization of a completed situation in Xiao and McEnery (2004), is termed the completive in Bao (2005), which is marked by the verbal *le*. The reason of giving the completive aspect such an unusual term as “actual” aspect is because *le* marks the actualization of an event or a state, not the completion of it (see Section 6.2.2).

6.2 了 *le*

As briefly discussed in Section 5.1.1, it is important to distinguish verbal *le* (V-*le*) from sentence-final *le* (S-*le*). Though both belong to perfective aspect markers (see Figure 6.1), they serve different functions. While V-*le* functions as a completive marker marking the completion/actualization of an event, S-*le* functions as an inchoative, which signals a “change of state” or the beginning of a new situation (Bao 2005). Li and Thompson (1981) consider that S-*le* indicates “current relevant state” (CRS), and refer to it as a “marker of CRS”. In the following sections, we will look at the usages of both V-*le* and S-*le* before examining their grammaticalization processes.

6.2.1 Two different *le*-s

The theoretical literature is increasingly converging on the point that there are two different *le*-s in Chinese which serve different grammatical functions (see Chao 1968; Li, Thompson 1981; Bao 2005), although they appear in the same form represented by one single character 了. The positions of the two *le*-s, however, are different. V-*le* (or suffix -*le*) is placed directly after a verb or adjective while S-*le* occurs in sentence-final position. Chao (1968) argues that V-*le* receives the preterit (past tense) interpretation in English while S-*le* is interpreted in terms of a perfect. This is exemplified in (97). Here in (97)a, the V-*le* marks the termination of the event, the same applies to the first *le* in (97)b. The second *le* in (97)b links it to the

time of utterance to emphasize that three letters exist at the moment of speaking. Besides, it seems to emphasize that three letters are quantitatively significant, and the speaker has finished them earlier than expected.

- (97) a. 我 今 早 写 了 三 封 信
 wǒ jīn zǎo xiě le sān fēng xìn
 I today morning write ASP three CL letter
 ‘I wrote three letters this morning.’ (talking in the afternoon or evening)
- b. 我 今 早 写 了 三 封 信 了
 wǒ jīn zǎo xiě le sān fēng xìn le
 I today morning write ASP three CL letter ASP
 ‘I have already written three letters this morning.’

Adopting Comrie’s (1976:81) view on aspect, *V-le* marks the completive aspect and the sentence-final *le* the “perfect aspect”. Li and Thompson (1981), on the other hand, suggest *S-le* signals a “currently relevant state” (CRS). Signaling a CRS is a communicative function from the perspective of functional reference grammar. Bao (2005), however, considers *V-le* responsible for the completive aspect while *S-le* gives rise to the inchoative interpretation.

However, it is difficult to distinguish *S-le* from *V-le* in terms of sentence position, especially when *le* is directly attached to adjective predicates.⁴⁸ When adjective predicates occur with *le*, they give rise to a dynamic, shifted interpretation (Smith 1997:265). In such cases, *le* triggers the inchoative interpretation, which presents the coming about of a state, as exemplified in (98).

- (98) 我 病 了
 wǒ bìng le
 I sick ASP
 ‘I got sick.’

⁴⁸ Therefore, some literature categorizes *le* directly following adjective predicates as *V-le* (Smith 1997) while some other studies consider it as *S-le* (Bao 2005; Teo 2019).

6.2.2 Aspectual meanings of V-*le*

As discussed in the previous chapter, CSE *already* is used to express various aspectual meanings, i.e. the completive, inchoative and prospective. All these usages find parallels in the Chinese perfective marker *le* (both V-*le* and S-*le*). This section first examines the aspectual meanings of V-*le*.

The completive use of *already* resembles that of *le* in marking “the termination of action” (Smith 1997:264). In (99), V-*le* is used to mark the completive aspect/actual aspect.

- (99) 我 洗 了 衣服
 wǒ xǐ le yīfu
 I wash ASP clothes
 ‘I washed the clothes.’

Citing Chu (1976), Xiao and MacEnery (2004:151) claim that the aspectual meaning of *le* in example (99) is ambiguous: the action of washing the clothes may be either terminated or completed. Unlike English, termination and completion are distinct concepts in Mandarin Chinese (Smith 1997:265). Completion is expressed unequivocally by resultative verb complements (RVCs). V-*le* and RVCs can occur together. As in (100), sentence a is acceptable in Mandarin, while sentence b is contradictory (see Xiao and MacEnery 2004:151). However, the type of closure depends on the type of situation. Example (100)a is only acceptable because the object is a bare (indefinite) noun. In this case, it depicts an atelic situation, in which there is no inherent final spatial endpoint. However, when the bare noun is replaced by a quantified direct object, the completive reading is unequivocal, as exemplified in (100)c:

- (100) a. 我 写 了 信, 可是 没 写 完
 wǒ xiě le xìn kěshì méi xiě wán
 I write ASP letter but NEG write finish
 ‘I wrote some letters, but I didn’t finish it.’

- b. *我 写 完 了 信, 可是 没 写 完
 wǒ xiě wán le xìn kěshì méi xiě wán
 I write RVC ASP letter but NEG write finish
 (Intended) 'I finished writing some letters, but I didn't finish it.'
- c. 我 写 了 一 封 信
 wǒ xiě le yì fēng xìn
 I write ASP one CL letter
 'I wrote one letter.'
- (Examples adapted from Xiao and MacEnery 2004:151)

When *le* is used in a telic situation, in which an inherent final spatial endpoint is naturally included, the completive reading (or closure) is unvarying for events such as activity, semelfactive, and achievement, as illustrated in (101), adapted from Smith (1997:264):

- (101) a. 他们 昨天 在 公园 吵 了 一 架
 tāmen zuótiān zài gōngyuán chǎo le yì jià
 they yesterday in park quarrel ASP one fight
 'They quarreled yesterday in the park.' (activity)
- b. 丽丽 忽然 打 了 喷嚏
 Lili hūrán dǎ le pēntì
 Lily suddenly hit ASP sneeze
 'Lily sneezed suddenly.' (semelfactive)
- c. 丽丽 在 中午 到 了 山顶
 Lili zài zhōngwǔ dào le shāndǐng
 Lily at noon arrive APS mountaintop
 'Lily reached the mountaintop at noon.' (achievement)

As mentioned earlier in (98), when *le* co-occurs with adjective predicates, it gives rise to the inchoative reading. Yet, time adverbials which indicate the final endpoint of a state trigger a shift to the completive reading, although the predicates belong to stative verbs. Example (102) exhibits a change out of a state, which induces the reading of a dynamic situation:

- (102) 我 在 那里 住 了 三 个 月
 wǒ zài nàlǐ zhù le sān gè yuè
 I at there live ASP three CL month
 'I lived there for three months.'

To summarize, *V-le* presents a termination and/or completion of an event. In an atelic situation, *V-le* presents a closed event, while in a telic situation *V-le* expresses completion.

6.2.3 Aspectual meanings of *S-le*

S-le, like CSE *already*, can express the completive, inchoative, and prospective when it co-occurs with situations of accomplishment. Consider the following examples:

- (103) 我 洗 衣服 了
 wǒ xǐ yīfu le
 I wash clothes ASP
 ‘I finished washing the clothes.’ (completive)
 ‘I have started washing the clothes.’ (inchoative)
 ‘I am about to wash the clothes.’ (prospective)

In contrast to *V-le* in (99), *S-le* in example (103) is compatible with different aspectual readings. The exact aspectual reading depends on the contexts of the utterance. The completive reading would mean that the speaker has finished washing the clothes; the inchoative would mean that the speaker has just started washing the clothes; and the prospective reading indicates that the speaker is about to wash the clothes.

When *le* is used to mark the inchoative/prospective aspect, it indicates a situation which is “new” or “new to the speaker” (Chao 1968:798).

- (104) a. (快) 下 雨 了
 (kuài) xià yǔ le (a new situation)
 (almost) fall rain ASP
 ‘It is raining (now).’ or ‘It is about to rain.’
- b. 是 啊, 十 一 点 半 了
 shì a shíyī diǎn bàn le (the quality or degree attained)
 yes PP eleven clock half ASP
 ‘Yes, it’s (as late as) half past eleven.’
- c. 汤 咸 了
 tāng xián le (an excessive degree)
 soup salty ASP
 ‘The soup is too salty.’

Example (104) shows that when *S-le* is an inchoative/prospective marker, it indicates that the situation becomes a reality. In addition, it signals a change of situation and its relevance to the situation at the time of reference. A “new situation” is observable as a fact, as in (104)a, a new realization of time, as in (104)b, or a new subjective judgment as in (104)c. Each of the examples above involves a comparison with the situation before the time of reference. Therefore, *S-le* is used to present new information to the addressee.

As *S-le* is compatible with various aspectual readings, narrowing it down to a particular reading depends on the context and the situation (i.e. aktionsart). Like CSE *already*, when *S-le* is used with stative verbs, it expresses a change of state; when it is used with types of aktionsart such as accomplishment, achievement, and activity, the completive, inchoative and prospective readings are possible (see Section 5.1.3). For example, if *S-le* occurs in an achievement situation, where it happens instantaneously, *S-le* expresses the completive aspect (Chao 2020:96).

- (105) 她 交 论文 了
 tā jiāo lùn wén le
 she submit dissertation ASP
 ‘She has submitted the dissertation.’

Perhaps due to the versatility in its aspectual readings, Chinese linguists have been trying to unify these different aspectual interpretations of *S-le* (Li and Thompson 1981; Li, Thompson, and McMillan Thompson 1982; Soh 2009). Li et al. (1982:22) propose that the basic function of *S-le* is to signal a “currently relevant state” (CRS): *S-le* has special current relevance with respect to some particular reference time. Soh (2009:624), on the other hand, makes use of the notion of “presupposition” and “common ground” (Stalnaker 1998), and proposes a “change of state” interpretation, which is associated with a change expressed by propositions within a common ground, while a “contrary to expectation” interpretation is associated with a change expressed by propositions across common ground.

However, the notions of “currently relevant state”, “change of state” and “contrary to expectation” are rather conceptual, which are difficult to apply when it comes to analyzing

the actual aspectual meanings of *S-le*. Besides, the term “current time” which is used to interpret currently relevant state is confusing because in some situations, current time means ‘here and now’, while in some other cases it refers to a particular “reference time” (Li et al. 1982:22–23). Apart from that, the interpretation of *S-le* overlaps with the “earlier than expected” interpretation of Standard English *already*.

In addition to the aspectual meanings, *S-le* seems to perform some pragmatic functions. For example, Chao (1968) notes that in certain situations, the use of *S-le* demands the presence of an audience and it appeals to the hearer to act in a certain way. Consider (106), in which *S-le* is used in imperative sentences to catch the attention of the audience (Chao 1968:798). It is also used to underline that the assertion is obviously true, as exemplified in (107). Li et al. (1982:37) add that *S-le* can function as a “finality marker” to emphasize that the sentence before *S-le* is all the speaker has to contribute at the moment, as presented in (108):

(106) In imperatives:

a. 吃 饭 了
chī fàn le
eat meal ASP
‘Let’s eat now!’

b. 咱们 坐 了
zánmen zuò le
we sit ASP
‘Let’s sit now!’

(107) To express obviousness:

a. 这 个 你 当然 懂 了
zhè ge nǐ dāngrán dǒng le
this CL you of course understand ASP
‘Of course you understand this!’

b. 再 好 没有 了
zài hǎo méiyǒu le
COMP good not have ASP
‘There is no better (option) (, which is very obvious).’

- (108) As a finality marker (Li et al. 1982:37):
 (to a friend)
 他 打 了 我 一 拳 了
 tā dǎ le wǒ yì quán le
 he hit ASP I one punch ASP
 ‘(What I want to say is that) he punched me! (And that’s it.)’

6.2.4 V-*le* and S-*le* in negative sentences

Unlike English *already*, S-*le* is compatible with negative sentences. It should be noted that there are two different negators in Chinese, one is 不 *bù*, the other is 没有/没 *méiyǒu/méi*. Li and Thompson (1981) argue that while *bù* is a neutral negative marker that denies the existence of a state, *méiyǒu/méi* denies the completion of an event or action. Other proposals (e.g. Lü 1980; Bai 2000) suggest that the distinction does not lie in temporal references but between objective narration and subjective evaluation: while *bù* negates subjective evaluation, *méiyǒu/méi* is used in objective narration. Returning to V-*le* and S-*le*, while V-*le* denoting the completive/actual aspect is incompatible with both negators, S-*le* can co-occur with both negators, which gives rise to the inchoative reading (Bao 2005:246, Li 2014:67). Consider the following examples drawn from the CCL corpus:

- (109) a. 她 [...] 几乎 上 课 都 没有 心情 了
 tā jīhū shàng kè dōu méiyǒu xīnqíng le
 she almost attend class ADD NEG mood ASP
 ‘She even has almost no mood to attend the classes anymore.’
 (CCL Contemporary-Literature-Mainland-Su.X.L.)
- b. 我 也 不 年轻 了
 wǒ yě bù niánqīng le
 I ADD not young ASP
 ‘I am not young anymore either.’
 (CCL-Contemporary-Magazine-Reader)

In both examples in (109), S-*le* denotes a change of state. By using S-*le*, the sentence in (109)a presupposes that the student had a good mood previously and at the time of reference, she does not have a good mood anymore. It represents a change of state from “having a good

mood” to “having no good mood”. Similarly, example (109)b suggests that the speaker was young previously and now he is no longer young.

Besides carrying different grammatical functions, diachronic research suggests that *V-le* and *S-le* originate from different etymons. While *V-le* originates from the verb 了 *liǎo* ‘to finish’, *S-le* may originate from multiple sources (Sun 1996). The following subsections on *le* will examine how *V-le* and *S-le* have developed into today’s usages in Chinese.

6.2.5 Grammaticalization of *V-le*

V-le and *S-le* are described to have undergone different grammaticalization processes. *V-le* originates from the lexical verb 了 *liǎo* ‘to finish’ in Early Middle Chinese (ca. AD 600 – mid-12th century) and later developed into an aspectual marker suffixed to a verb. On the other hand, one of the precursors of *S-le* is hypothesized to come from the expression 来 *lái* meaning ‘to come’. It was phonologically reduced and became a homonymy of *V-le*. We will first look at the grammaticalization process of *V-le* analyzed in Sun (1996).⁴⁹

Figure 6.3 shows that the use of 了 *liǎo* as an actual verb meaning ‘to close/finish/end’ made up more than 30 percent of all tokens of 了 in the 10th century. This use, however, has significantly decreased in number in the 13th century and reduced to only 0.1 percent in the 15th century.

⁴⁹ The data are obtained as quoted in Sun (1996) from the Index of *Zutangji* by Yanagida Seizan (1980) (for the distribution in the 10th century), the studies of *Zhuzi yulei* by Cui (1980) and Mu (1986) (for the 13th century), and the text of *Laoqida* in the 15th century.

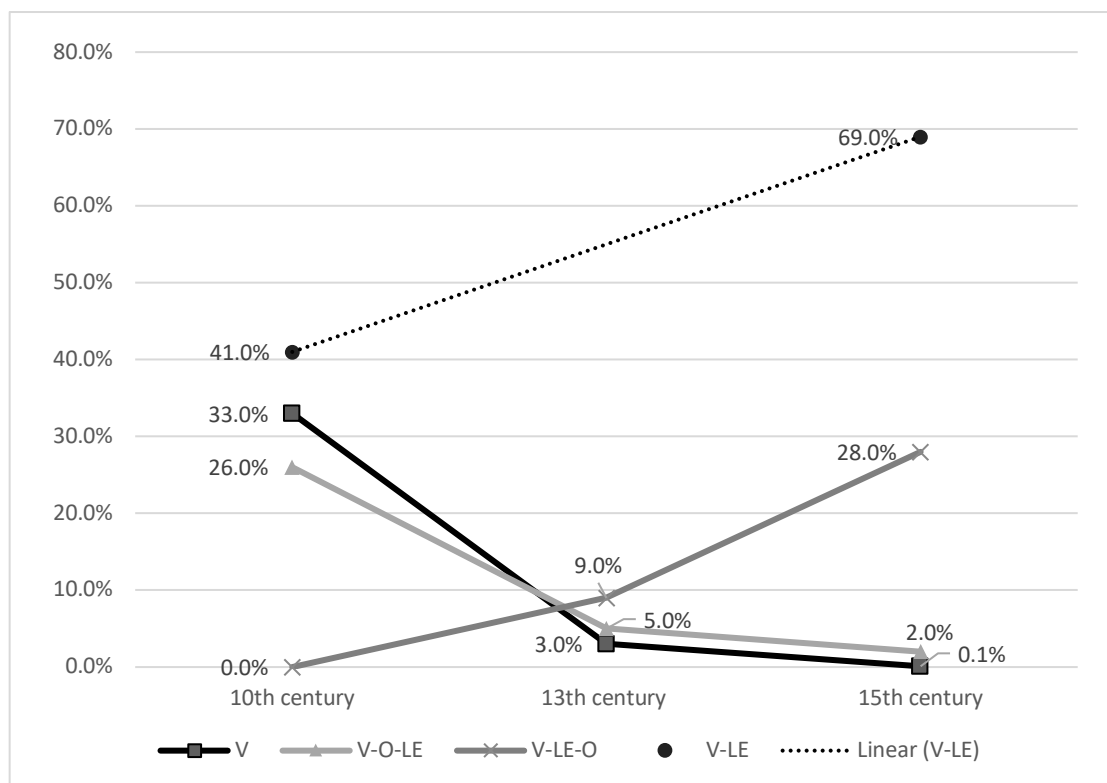


Figure 6.3: Grammaticalization of *V-le* in historical texts (adapted from Sun 1996:106)

A similar trend in the same period applies to the distribution of the *V-O-le* sequence,⁵⁰ which has decreased from more than 25 percent to 2 percent. The *V-O-le* sequence was used to describe the completion of the preceding event (Sun 1996:88), as exemplified in (110):

- (110) a. 作 此 语 了
 zuò cǐ yǔ liǎo
 make DEM words complete
 ‘(One) finished making this statement.’
 (from *Dunhuang bianwen*, cited in Sun 1996:88)
- b. 过 江 了 向 行 者 云
 guò jiāng liǎo xiàng xíngzhě yún
 cross river complete towards walking person say
 ‘Upon having completed the crossing of the river, (one) talked to the person who was walking.’
 (from *Zutangji*, cited in Sun 1996:88)

⁵⁰ Different from *S-le*, Here 了 in the sequence *V-O-le* is pronounced as *liǎo*, a triphthong on the third tone (Sun 1996:104).

On the other hand, the *V-le*-O sequence rose from practically zero in the 10th century to approximately 30 percent in the 15th century. The same trend applies to *V-le*, which increased from 41 percent to 69 percent during the same period. The trendline of *V-le* in Figure 6.3 is represented by a dotted line as the data for the 13th century is missing.⁵¹ In both sequences, *le* became a morphologically bound affix. Furthermore, *V-le* was increasingly favored over *V-què*, which is said to be the first affix-like perfective marker in Chinese (Mei 1981 and Cao 1986), as demonstrated in (111):

- (111) a. 过 却 多少 林木
 guò què duōshǎo lín mù
 pass ASP much woods
 ‘One passed so many woods.’
 (*Zutangji*, cited in Sun 1996:87)
- b. 军官 食 了 便 即 渡 江
 jūnguān shí le biàn jí dù jiāng
 officer eat ASP then soon cross river
 ‘Immediately after finishing eating, the officers cross the river.’
 (*Dunhuang bianwen*, cited in Sun 1996:89)

The counterpart of V-le in Cantonese

A further piece of evidence in support of the separation of *V-le* from *S-le* is the Cantonese completive/actual aspect marker 左 *jo2*. In Cantonese, 左 *jo2*, meaning ‘to make it become the past’ or ‘to finish’ is the equivalent of *V-le* in Mandarin. Like *V-le*, it is used as a completive/actual aspect marker, which occurs immediately after the verb (Matthews and Yip 1994:204–5), as illustrated in (112):

- (112) 我 食 左 饭 (Cantonese)
 ngo5 sik6 jo2 faan6
 I eat ASP meal
 ‘I just ate/I have eaten.’

⁵¹ Besides the lack of data for the 13th century, the distinction between Verbal-*le* and Sentence-*le* in a *V-le* sequence heavily depends on the context. Therefore, Sun (1996:89) notes that the findings on the *V-le* sequence may be rather inaccurate.

It is worth noting that *jo2* can only appear post-verbally, but it cannot occur sentence-finally, as exemplified in (113). This confirms that this counterpart of *V-le* is different from *S-le*. Hence, we can deduct that *S-le* is different from *V-le*, which must have come from different sources.

- (113) a. 我 食 左 饭 啦
 ngo5 sik6 jo2 faan6 la1
 I eat ASP meal ASP
 ‘I just ate/I have eaten.’
- b. *我 食 左 饭 左
 ngo5 sik6 jo2 faan6 jo2
 I eat ASP meal ASP
 ‘I just ate/I have eaten.’

6.2.6 Grammaticalization of *S-le*

The etymology of *S-le* is more complicated than *V-le*. Sun (1996:93) proposes that the grammaticalization of *S-le* is related to the interaction among a set of sentence-final particles including 也 *yě*, 矣 *yǐ*, 来 *lái* and also *V-le*. According to Sun (1996:106), the above-mentioned sentence-final particles could all function as perfect markers in Early Mandarin. Yet, *S-le* remains the only survivor due to grammaticalization, which tends to promote a small, homogeneous paradigm. We will examine one of the most possible chains of the grammaticalization of *S-le*, namely 以来 *yǐlái* > 来 *lái* > *S-le*.

The perfect marker 来 *lái* is considered to be one of the most likely precursors of *S-le* (Sun 1996:100). The example in (114) shows that the compound *yǐlái* meaning ‘until’ was used to relate two time points as a conjunction in Old Chinese. The compound was reduced to the monosyllabic *lái* later in Middle Chinese (Mei 1978), as demonstrated in (115). By the 15th century *lái* had functioned as a perfect marker, as shown in (116). The diphthong *lái* went out of use around the 18th century and was replaced by its phonological reduced form *le*. Since then *le* has been used as a perfect marker in sentence-final position.

- (114) 自 生 民 以 来 未 有 孔 子 也
 zì shēng mǐn yǐlái wèi yǒu kǒngzǐ yě
 from birth people till NEG have Confucious PTC
 ‘From the birth of mankind till (now), there has never been another like Confucius.’ (*Mengzi GongsunChou shang*, cited in Sun 1996:97)
- (115) 顾长康 画 有 苍 生 来 所 无
 gùchángkāng huà yǒu cāngshēng lái suǒ wú
 NAME painting have people till DEM NEG.have
 ‘From the beginning of mankind till (now), there has been nothing (like) Gu Changkang’s paintings.’ (*Shishuo xinyu*, cited in Sun 1996:97)
- (116) 大 雄 山 下 采 菌 子 来
 dàxióng shān xià cǎi jūnzi lái
 NAME mountain below pick fungi CRS
 ‘I have been to the foot of the Daxiong mountain to pick mushrooms.’
 (*Jingde chuandenglu*, cited in Sun 1996:98)

It is worth noting that *S-le* can be interpreted either as a perfective marker or as a marker of currently relevant state (see Section 6.2.3). Therefore, even for native speakers, it is difficult to distinguish the phonological reduced *lái* from *V-le*. For that reason, the uses of *V-le* may also converge with *S-le* (Sun 1996:102).

6.2.7 Summary

After comparing the studies on *already* in CSE with the studies on the *V-le* and *S-le* in Chinese, we can see that the Chinese *le*-s have a wider range of aspectual functions than that of *already* in CSE. Previous studies on *already* in Singapore have identified three major aspectual functions of CSE *already*, namely completive, inchoative/inceptive and prospective. Apparently, all these special aspectual uses find their counterparts in Chinese. The completive reading of CSE *already* corresponds to the completive/actual aspect of *V-le* and *S-le*. CSE inchoative/inceptive *already* is parallel to the *S-le* which signals a change of state. Like CSE *already*, *S-le* can express the completive, inchoative, and prospective when it co-occurs with situations of accomplishment. When *S-le* is used in negative sentences, it gives rise to the inchoative/change-of-state reading.

However, some pragmatic functions of S-*le* have not been discovered in CSE *already*, such as its emphatic function in imperative sentences and its function as a finality marker. We will further explore whether these functions have been appropriated into CSE *already* through OHI.

In addition, diachronic research suggests that V-*le* and S-*le* originate from different sources, which further supports that V-*le* is different from S-*le*. In light of the discussions of “contact-induced grammaticalization” (Heine and Kuteva 2003, 2005), according to which language contact may also lead to replication of the process of grammaticalization, it will be fascinating to explore whether CSE speakers have replicated the grammaticalization processes of V-*le* and S-*le* into CSE *already* (see Chapter 9.2.2 for the discussion of grammaticalization of *already* in relation to its substrate *le*).

6.3 过 *guò*

6.3.1 Aspectual meanings of *guò*

Like V-*le*, the aspectual marker 过 *guò* follows the actual verb and expresses the perfective viewpoint. According to Li and Thompson (1981:227), the interpretation of *guò* is close to ‘once’, ‘ever’ or ‘before’. As an aspectual marker, it presents a prior closed situation of any type, and emphasizes that its final state no longer obtains (Smith 1997:266). Therefore, *guò* is often referred to as the “experiential” marker as its perfective reading usually carries experiential force.

Unlike inchoative *le*, which conveys the beginning of a state or an event, experiential *guò* expresses a discontinuity with the present or other reference time, i.e. the final state of the earlier situation no longer upholds (Smith 1997:266). Yet, the closure of the previous event continues to hold implication at the utterance time or other reference time, which is similar to the English perfect (Ziegeler 2015:215). Consider the contrast between (117)a and (117)b, they are identical except for the aspectual markers:

- (117) a. 他 一 个 月 以 前 去 过 新 加 坡
 tā yí gè yuè yǐqián qù guò xīnjiāpō
 he one CL month ago go ASP Singapore
 ‘He went to Singapore one month ago.’ (He is no longer there)
- b. 他 一 个 月 以 前 去 了 新 加 坡
 tā yí gè yuè yǐqián qù le xīnjiāpō
 he one CL month ago go ASP Singapore
 ‘He went to Singapore one month ago.’ (He may still be there)

Guò in (117)a expresses a discontinuity of the event, while *V-le* implies that the speaker may still be in Singapore.

Besides the subtle aspectual differences in aspectual meaning, *V-le* and *guò* prefer different temporal adverbs: while *V-le* usually co-occurs with temporal adverbs referring to the recent past such as 刚刚 *gānggāng* ‘just now’, *guò* prefers indefinite distant past adverbs such as 以前 *yǐqián* ‘before’ (Matthew and Yip 1994:206; Smith 1996:267):

- (118) a. 我 刚 刚 看 了 这 本 书
 wǒ gānggāng kàn le zhè běn shū
 I just now read ASP this CL book
 ‘I’ve just read this book.’
- b. 我 以 前 看 过 这 本 书
 wǒ yǐqián kàn guò zhè běn shū
 I before read ASP this CL book
 ‘I’ve read this book before.’

Experiential *guò* requires a situation which is regarded as unusual and uncommon for the speaker, yet such a situation is often repeatable (Iljic 1987:71). The situation in question has to be of a type that can occur more than once. One can assume that this situation has taken place at least once. Consider the examples in (119):

- (119) a. 你 吃 过 榴 莲 没 有 ?
 nǐ chī guò liúlián méiyǒu
 you eat ASP durian no
 ‘Have you ever eaten durian?’

- b. 我 吃 过 臭 豆腐
 wǒ chī guò chòu dòufu
 I eat ASP stinky tofu
 ‘I have tried stinky tofu once.’

In this sense, its semantic interpretation is close to English *ever*. Yet, while English *ever* is a negative polarity item, Chinese *guò* is not, i.e. it is compatible with both the negation markers 不 *bù*, and 没有/没 *méiyǒu/méi* (see Xiao and McEnery 2004). This explains why CSE *ever* occurs in both affirmative and negative sentences, which is due to substratum interference. This results in a semantic extension of CSE *ever* from a negative polarity items (NPI) to non-negative-polarity contexts. Consequentially, it is grammaticalized to a marker of experiential aspect.

6.3.2 Grammaticalization of *guò*

The chain of grammaticalization of Chinese *guò* is different from that of CSE *ever*. Originally, *guò* was used as an actual verb meaning ‘to pass a spatial point’ or ‘to pass a temporal point’ before it was grammaticalized into an aspectual marker. When it is used as a transitive verb, its direct object is a locative or temporal noun (Shi 2002:139). Consider the following examples of the first two uses of *guò*:

- (120) a. 子墨子 归 过 宋
 zǐmòzǐ guī guò sòng
 Micius return pass State of Song
 ‘Micius returned. He passed the State of Song.’
 [CCL-Classical Chinese-02*Chunqiu-Mozi*, AD 760]
- b. 过 期 三 日
 guò qī sān rì
 pass due time three day
 ‘It expired for three days.’
 [CCL-Classical Chinese-02*Chunqiu-Zuozhuan*, AD 760]

In the following examples, *guò* appears in the structure of V [*guò* O]. Here, [*guò* O] could be analyzed as a resultative structure, in which *guò* functions as an adverbial modifier meaning ‘across/past/through’.

- (121) a. 虾蟆 跳 过 雀儿 浴
 há má tiào guò què er yù
 frog jump pass bird bathing pond
 ‘The frog jumped across the bird-bathing pond.’
 (Poem *Zeng Houxi*, by Han Yu, AD 800)
- b. 杜鹃 你 休得 叫 过 通宵
 dùjuān nǐ xiū dé jiào guò tōngxiāo
 cuckoo you not cry pass all night
 ‘Cuckoo, don’t you cry through all night!’
 (*Zhang Xie Zhuangyuan*, AD 1200)

It is not until the 8th century that *guò* started to acquire the meaning of “activity-experiencing” (Shi 2002:139), which is akin to its aspectual function. Instead of forming a verb-object relation with the locative or temporal object as in (121), it builds a stronger tie with the preceding verb, as shown in (122):

- (122) 本 司 检 过
 běn sī jiǎn guò
 our office check pass
 ‘Our office checked it.’ (Ru Tang qiuфа Xunli ji, AD 800)

The activity-experiencing *guò* was limited to the structure of V *guò*, which does not allow an object. It is not until the 13th century that the activity-experiencing *guò* as an RVC started to occur in the form V- *guò* O (Shi 2002:140), which indicates that *guò* has grammaticalized into an experiential aspect marker. In this process, *guò* and its preceding verb became fused. Therefore, *guò* lost its lexical status, as shown in (123):

- (123) 饮 过 酒 今 番 不 忘
yǐn guò jiǔ jīn fān bú wàng
drink ASP wine this time NEG forget
'Once we drink wine, we don't forget this moment'
(Zhuge Liang bowang shao tun, AD 1300)

In (123), the object 酒 *jiǔ* 'wine' is not the patient of *guò*, but the patient of the verb 饮 *yǐn* 'to drink'. In other words, *guò* neither denotes 'to pass a locative or temporal point' nor does it function as an RVC meaning "activity experiencing". It became the verbal suffix, and is used to mark the perfective aspect, meaning that the activity or event has been experienced once at a certain temporal reference point in the past.

6.3.3 Summary

In sum, the aspectual *guò* expresses the perfective viewpoint, which refers to situations that happened at least once in the past. Different from *V-le*, it expresses a discontinuity of an event or a state. When there is no specific time adverbial, *guò* typically requires a situation that is unusual and uncommon, yet repeatable. Moreover, *guò* is compatible with both negative and positive contexts.

The aspectual meaning of *guò* corresponds to the usage of *ever* in CSE. In addition, the substratum influence from *guò* is the major reason for the grammaticalization of CSE *ever* as an experiential marker. However, diachronic research in *guò* reveals the perfective *guò* comes from its verbal source meaning 'to pass'. It suggests that CSE *ever* did not replicate the grammaticalization process of its substrate *guò* (see further detailed discussion of contact-induced grammaticalization in Chapter 9).

6.4 也/都 yě/dōu

6.4.1 The additive marker yě/dōu

In Mandarin Chinese, 也 yě usually functions as an additive marker. It is used in preverbal position, meaning ‘also’ or ‘too’, as illustrated below.⁵²

- (124) A: 你 保重 啊
 nǐ bǎozhòng a
 you take care PTC
 ‘Take care!’
- B: 你 也 是
 nǐ yě shì
 you also COP
 ‘You too.’

The additive marker yě can be used in negative sentences as shown in (125). In both examples, yě connects with a previous sentence. It is used to express a similar situation or add complementary information. In this sense, the additive marker yě is not very different from English *also* and *either*, which also connect with a previous sentence. However, using *also* or *either* depends on whether the additive marker occurs in an affirmative or a negative sentence (see Chapter 5.2.1).

- (125) A: 我 不 认识 他
 wǒ bú rènshi tā
 I not know him
 ‘I don’t know him.’
- B: 我 也 不 认识 他
 wǒ yě bú rènshi tā
 I also not know him
 ‘I don’t know him, either.’

⁵² The Mandarin examples in (124) are inspired by the Cantonese examples from Matthews and Yip (1994:188).

In (125), 也 *yě* cannot be replaced by 都 *dōu*, as using *dōu* would express ‘I don’t even know him’ in Mandarin Chinese (Lou February 2020 p.c.). In Cantonese, however, 都 *doul* can be used in all the above contexts, while Cantonese 也 *ya5* is restricted to written contexts.

6.4.2 With universal quantification

Besides expressing the meaning of ‘in addition to’, which corresponds to English *also*, Mandarin 都 *dōu* is commonly used in nearly all forms of universal quantification, in expressing the meanings ‘all’, ‘every’ and ‘each’. Xiang (2008) proposes that *dōu* is a maximality operator, which operates on any plural set and yields maximal plural individuals as exemplified in (126). The same applies to Cantonese 都 *doul* (Matthews and Yip 1994:266). It is worth noting that the additive 也 *yě* does not co-occur with universal quantifiers to express this particular meaning.

- (126) a. 他 个个 女孩 都 追 的
 tā gègè nǚhái dōu zhuī de
 he CL-CL girl all pursue EMP
 ‘He chases all the girls.’

(adapted from Matthews and Yip 1994:262)

- b. 人 人 都 应该 保护 环境
 rén rén dōu yīnggāi bǎohù huánjìng
 person person all should protect environment
 ‘Everyone should protect the environment.’

(adapted from Matthews and Yip 1994:266)

Another influential analysis is to treat *dōu* as a distributivity operator, which serves as a predicate over a plural noun phrase preceding it (Yang and Wu 2019:24). While English distinguishes dual and plural (more than dual) with *both* and *all*, Chinese *dōu* can be used in all plural situations, as demonstrated in (127):

- (127) a. 两 个 都 好
 liǎng gè dōu hǎo
 two CL all/both nice
 ‘They are both nice.’

- b. 他们 三 个 都 去
 tāmen sān gè *dōu* qù
 they three CL all go
 ‘All three of them are going.’

As observed in the above examples, the additive marker *dōu* occurs in preverbal position. In addition, *dōu* can be used in negative sentences. The negator immediately precedes the universal quantifier:

- (128) a. 不是 个个 都 那么 空闲
 búshì gègè *dōu* nàme kòngxián
 NEG.COP CL-CL all so free
 ‘Not everyone has so much free time.’
- b. 大学 不是 谁 都 录取 的
 dàxué búshì shéi *dōu* lùqǔ de
 university NEG.COP who all accept EMP
 ‘The university does not just accept anyone.’

6.4.3 Concessive *lián* with *dōu/yě*

The adverbs *yě* and *dōu* can co-occur with 连 *lián*, which corresponds to the CSE *also* with concessive *even*. Chinese scholarly literature generally treats *lián* as a focus marker, which is used to emphasize the elements following it in the communicative context (see Chao 1968; Lü 1980; Matthews and Yip 1994). The collocation of *lián* and *yě/dōu* gives rise to a concessive *even* reading, as *lián* invokes contextually quantified individuals, and *yě/dōu* distributes the properties of the main predicate over the individuals (Yang and Wu 2019:26). This analysis of *lián* and *yě/dōu* corresponds to the scalar additive analysis (see Chapter 5.2.1). Like English *even*, *lián* can be analyzed as an additive marker with an additional scalar component except that such an additional scalar component has to be licensed by *dōu*. The associate of the additive is the least likely candidate among the set of alternatives for which the proposition holds, as illustrated in (129):

- (129) (连) 你 都/也 不 给 我 面子
(*lián*) nǐ *dōu/yě* bù gěi wǒ miànzi
(even) you also not give me face
'Not even you show respect for my feelings.'

In the above example, the person referred to by the pronoun 你 *nǐ* 'you' is the least expected person that does not respect the speaker's feelings. Interestingly, the concessive *lián* can be omitted here without making the sentence incomplete. The omission of *lián* results in an additional scalar reading of *yě/dōu*, on top of its additive meaning (see Chapter 9.2.5 on the grammaticalization of *also* in relation to Chinese *yě/dōu*).

6.4.4 Summary

In sum, unlike English *also*, which is an additive marker restricted to positively oriented contexts, Chinese *dōu* and *yě* can be used in a wider range of contexts. In other words, Chinese *dōu* and *yě* are not limited to positively oriented contexts, as they can be used in negatively oriented contexts as an equivalent of English *either*. In addition, *dōu* is commonly used in nearly all forms of universal quantification, in expressing the meanings 'all', 'every' and 'each'. Besides, *dōu* can be used in both dual and plural situations, while English typically distinguishes these two situations with *both* and *all*. Finally, the concessive reading of *even* is achieved by the *lián*-NPs in collocation with *yě/dōu*. Again, these additional functions and contexts of *yě* and *dōu* beyond the English additive *also* are accountable for the semantic extension of CSE *also*.

6.5 的 *de*

6.5.1 The “relativizer” and “nominalizer”

De is the most widely used morphosyntactic particle in Modern Chinese, which is commonly used to connect a modifier to its head in five types of constructions – genitive, adjectival, associate and adverbial phrases and relative clauses. Shi and Li (2002:1) generalize these

various functions under the umbrella term “relativizing” and termed *de* as a “relativizer”. Similar to the English possessive apostrophe ’s, *de* is used as a connector in genitive or associate phrases to express that something belongs to something else or to express a type of relationship between things. For example, *de* is used in a Noun-Noun cluster to signal that the noun on the left modifies the one on the right, as exemplified in (130)a. In addition, *de* is used to connect an adjective with a noun (Huang et al. 2009:14). As shown in (130)b, *de* is used to attribute the color “red” to the “bicycle”.

- (130) a. 教室 的 桌子
 jiàoshì de zhuōzi
 classroom POSS table(s)
 ‘the classroom’s tables’ or ‘the tables of the classroom’
- b. 红色 的 自行车
 hóngsè de zìxíngchē
 red ATTR bicycle
 ‘a/the red bicycle’

The examples above exhibit the feature of having a modifier marked by *de*, which precedes the head constituent. Such use of *de* does not resemble that of English *one*. Thus, one may ask why CSE speakers regard Chinese *de* as equivalent to English *one*. Yet, if the head noun is already established in the discourse context, the form “modifier + *de*” may occur without the head (Paris 1979), as illustrated in (131). In this sense, *de* is very similar to the anaphoric *one* and the independent *one* (see Section 5.4.1).

- (131) a. 最 贵 的 (车)
 zuì guì de (chē)
 most expensive one (car)
 ‘the most expensive one/car’
- b. 我 昨天 看见 的 (朋友)
 wǒ zuótiān kànjiàn de (péngyǒu)
 I yesterday see POSS (friend)
 ‘the friend that I saw yesterday’

However, anaphoric *one* and independent *one* in English is restricted to short phrases such as “the big *one*” or “the expensive *one*”. Chinese *de* carries more functions than English *one*. As shown in (131)b, *de* can occur between a relative clause and a nominal phrase. When the head is omitted and *de* occurs in the frame “modifier + *de*”, *de* is referred to as a nominalizer (Li and Thompson 1981:579; Liu 2003; Wu 2004). As put by Li and Thompson, “[a] nominalization can be called a relative clause if the head noun it modifies refers to some unspecified participant in the situation named by the nominalization” (1981:579). It is worth noting that unlike English, Chinese relative clauses do not need a relative pronoun like the English *that* or *who*. Besides, quantifiers and demonstratives are optional in Chinese, as shown in (132):

- (132) a 我 昨天 买 的 书
 wǒ zuótiān mǎi de shū
 I yesterday buy REL book
 ‘the book(s) that I bought yesterday’
- b. 她 来 德国 的 事实
 tā lái déguó de shìshí
 she come Germany REL fact
 ‘the fact that she came to Germany’

In essence, *de* as a relativizer is semantically not different from *de* as a nominalizer. Following Li and Thompson (1981:579), when an NP is absent after the “relativizer” *de*, *de* is structurally attached to a verb phrase or an adjectival phrase, and consequently becomes a nominalizer. In the following examples, the nominalization function of *de*, i.e. converting phrases which are not nominal in nature into the nominal class (Liu 2003:74), can be simply viewed as a relativizer with its nominal phrase omitted.

- (133) a. 冰箱 里 有 什么 吃 的 吗
 bīngxiāng lǐ yǒu shénme chī de ma
 refrigerator inside exist what eat NOM PP
 ‘Is there anything that can be eaten in the refrigerator?’
 or ‘Is there any food in the refrigerator?’ (adapted from Liu 2003:74)

- b. 他 是 教 书 的
 tā shì jiāo shū de
 he is teach book NOM
 ‘He is the one who teaches.’
 or ‘He is a teacher.’

The examples above can be viewed as “head-less” relative clauses, where the NP following *de* is not realized. In (133)a, the phrase 吃的 *chī de* can be either interpreted as ‘something that can be eaten’ or ‘something to eat/food’. Syntactically, the particle *de* converts the verb 吃 *chī* ‘to eat’ into a noun meaning ‘something to eat’. Also, *de* can be attached to a verb of activity to describe a person’s occupation, as shown in (133)b, where *de* nominalizes the verb phrase 教书 *jiāoshū*, lit. ‘to teach book’ into a noun phrase meaning ‘someone who teaches’.

6.5.2 Sentence-final *de*

The sentence-final particle *de*, which often occurs with the copula 是 *shì* ‘be’ in forming the *shì...de* construction, is increasingly recognized as a cleft which can be compared to the English cleft structure *It is ... that* in theoretical literature on Chinese syntax (Paul 2008; Hole 2011). The focused element commonly follows the copula *shì* and precedes sentence-final *de*. The schema of the type follows the pattern of “topic + COP + cleft”, as shown in the following example:

- (134) 他 (是) 昨天 来 的
 tā (*shì*) zuótiān lái de
 he (COP) yesterday come EMP
 ‘The situation is that he came yesterday.’ or ‘It is yesterday that he came.’
 (adapted from Li and Thompson 1981:59)

In *shì...de* construction, the constituent bounded by *shì* and *de* is emphasized, which can be interpreted as focused new information (Chao 1968; Li and Thompson 1981). Chao (1968:296) refers to sentence-final *de* as “situational *de*”, which is used to refer to the whole situation of the utterance, meaning ‘such is the case’, or ‘this is the kind of situation’. Li and

Thompson (1981:589) point out that instead of purely reporting an event, “the *shì ...de* construction serves to characterize or explain a situation by affirming or denying some supposition”. As illustrated in (134), by using the cleft structure, the speaker affirms that *he came yesterday* and denies that *he came on any other day rather than yesterday*. Such notion is referred to as “exhaustibility” (the fact that all alternatives are implied or presupposed to be false), which is a cross-linguistically stable property of clefts (Rochemont 1986:127).

Another piece of evidence to support the cleft reading is provided by the replacement of the *shì...de* construction with the perfective *le*. In (134), the situation that *it is yesterday that he came* is emphasized. The implication is that *it is not today, nor is it any other day that he came, but yesterday*. When the *shì...de* construction is replaced by the perfective *le* as shown in (135), such emphasis on *yesterday* is not present. The focus of (135) with the use of the perfective marker *le* is on the closure of the activity.

- (135) 他 昨天 来 了
 tā zuótiān lái le
 he yesterday come PFT
 ‘He came yesterday.’

Ross (1983), on the other hand, argues that cleft *de* and relativizer/nominalizer *de* can be unified into one single category, which follows the string [NP MOD *de* NP]. He termed this unified function as the “NP modification marker”. Following the string, the various types of *de* phrases and clauses are illustrated in (136), adapted from (Ross 1983:216–18):

- (136) a. Possessive *de* [NP [MOD Zhāng sān] de [NP shū]]
 ‘Zhangsan’s book’
 b. Relativizer/nominalizer *de* [NP [MOD wǒ zuótiān mǎi] de [NP shū]]
 ‘the book I bought yesterday’
 c. Cleft *de* wǒ shì [NP [MOD zuótiān mǎi shū] de [NP e]]⁵³
 ‘It is yesterday that I bought the books.’

⁵³ ‘e’ denotes an empty node.

One argument for the unification of all *de* constructions is that the clauses with emphatic *de* can simply be viewed as headless relative clauses (Liu 2003:77; Lee 2005). Furthermore, the subject of the sentence must be the same as the missing subject participant in the nominalization (Li and Thompson 1981:588). In addition, when a direct object occurs in the nominalization of the *shì...de* construction, a reversal can take place, as shown in (137):

- (137) a. 他们 (是) 八 点钟 开 门 的
 tāmen (shì) bā diǎn zhōng kāi mén de
 they (COP) eight o'clock open door EMP
 'It is at 8 o'clock that they open the door.'
- b. 他们 (是) 八 点钟 开 的 门
 tāmen (shì) bā diǎn zhōng kāi de mén
 they (COP) eight o'clock open REL door
 'It is at 8 o'clock that they opened the door.'

The NP modification marker approach is very convenient as it unifies *de* phrases and clauses into one single structure. However, there are two major shortcomings with this unifying approach. First of all, neither possessive *de* nor relativizer/nominalizer *de* has emphatic interpretations while the cleft interpretation emphasizes the clause bounded by copula *shì* and sentence-final *de*. Secondly, while emphatic *de* can be omitted without making the utterance incomplete, possessive *de* and relativizer/nominalizer *de* have to be present (Lee 2005). Furthermore, the reversed structure [Verb-*de*-Object] of the *shì...de* gives rise to a past tense interpretation. Simpson and Wu (2002) argue that *de* in the cleft structure undergoes a process of lateral grammaticalization syntactically, in which *de* changed from category determiner to category tense and becomes a past tense morpheme in [Verb-*de*-Object] form. The grammaticalization results from the reinforcement of a past time conversational implicature commonly associated with the *shì...de* construction (Simpson and Wu 2002:170). This proposal seems plausible, as [Verb-*de*-Object] is not compatible with a future tense, as illustrated in (138).

- (138) *他是明天才会来的学校
 tā shì míngtiān cái huì lái de xuéxiào
 he COP tomorrow first will come REL school
 (Intended meaning) ‘It is not until tomorrow that he will come to school.’

Evidently, the construction [Verb-*de*-Object] must involve an event that has already happened. Its incompatibility with a future tense suggests that past tense is a necessary implication of a *de* clause.

6.5.3 Grammaticalization of *de*

Yap, Deng and Caboara (2017) suggest that *de* originates from the locative/spatial noun 底 *dǐ* ‘bottom’ (initially 氏 *dǐ* ‘foundation’ or ‘base’) in Old to Modern Chinese and developed into a relativizer/nominalizer in Middle to Modern Chinese, as exemplified in (139). Relativizer/nominalizer 底 *dǐ* underwent phonological reduction and became the phonological variant 的 *de* around the 8th to 10th century, which further developed into an emphatic marker (also “attitudinal marker” or “stance marker”, see Yap et al. 2017). From the 13th century on, 底 *dǐ* went out of use as a nominalizer/relativizer and an emphatic marker, and was completely replaced by 的 *de*.

- (139) 如何是一丸疗万病底药
 rúhé shì yì wán liáo wàn bìng dǐ yào
 how COP one CL cure ten thousand illness REL medicine
 ‘What is a medicine that cures all illnesses?’
 (*Jingde chuandenglu*, Southern Song, AD 1127–1279; cited in Yap et al. 2017:3)

According to Yap et al. (2017), “bridging contexts” (Heine 2002) provided important forces for the grammaticalization of 底 *dǐ* into an emphatic marker. As illustrated in (140), both the nominalizer and the emphatic functions are possible readings of 底 *dǐ*. Such ambiguity triggers the reanalysis of 底 *dǐ*.

- (140) 上 下 两 轮 月 若 个 是 真 底
 shàng xià liǎng lún yuè ruò gè shì zhēn dǐ
 up down two CL moon which CL COP real NOM/EMP
 ‘The moon up (in the sky) and the moon down (in the water), which one is the real one?’ or ‘which one is real?’
 (*Xia Ye Wan Yue*, Yang Wanli, Southern Song, AD 1127–1279; cited in Yap et al. 2017:3)

In Example (140), 若个是真底 *ruò gè shì zhēn dǐ* can be interpreted as [若个是[真底]] ‘which moon is the real one’, which focuses on a nominal predicate construction with a referential reading [which one is [the real (one)]]. The referential reading implicates contrast between a “real moon” and the “reflection of the moon”. However, due to the absence of the head of the NP, it can also be reinterpreted as [[若个是真]底] [[which one is real] EMP], which shifts the focus from the referential reading to the speaker’s personal evaluation of the proposition of the utterance (Yap et al. 2017:4). Such development constitutes a case of semantic expansion. As 底 *dǐ* falls in the final prosody of the utterance, it is an ideal candidate for hosting the prosodic cues to reflect the speaker’s stance/attitude towards the proposition.

In sum, the chain of the grammaticalization of *de* can be formulated as: [底 *dǐ* ‘foundation’: lexical noun] > [底 *dǐ* ‘bottom’: spatial noun] > [底 / 的 *dǐ/de*: relativizer/nominalizer] > [的 *de*: emphatic marker].

6.5.4 Summary

Again, in the case of Chinese *de*, we observe a wider range of semantic scope than English *one*. Not only does *de* function as a possessive marker, but it also serves as a relativizer/nominalizer which connects a relative clause with a nominal phrase. Furthermore, it co-occurs with the copula 是 *shì* ‘be’ in forming the cleft *shì...de* construction to emphasize the constituent that occurs between them as exclusive, i.e. that any other alternative is implied or presupposed to be false.

The only parallel between Chinese *de* and English *one* is their usage as an anaphoric pronoun, e.g. 最大的 *zuidàde* ‘the biggest one’. Apart from that, English *one* does not manifest the functions of linking relative clauses and monomializing phrases that are of non-nominal nature. It is perhaps this particular connection of both being able to serve as an anaphoric pronoun that triggers CSE speakers to consider them as equivalents and therefore transfer the other semantic functions of *de* into CSE *one*.

The historical development of Chinese *de* suggests that it developed from a locative/spatial noun into a nominalizer/relativizer, which further developed into an emphatic marker. Chinese *de* as a nominalizer is head-final, and when it occurs in sentence-final position, it can naturally be reanalyzed as a carrier of sentence-final prosody.

6.6 Chapter summary and conclusion

The previous sections discussed the meanings of the Chinese expressions 了 *le*, 过 *guò*, 也 *yě* / 都 *dōu*, and 的 *de*, which are considered as counterparts of the four CSE expressions *already*, *ever*, *also*, and *one*. The parallels between the various functions of the four CSE markers and their Chinese counterparts confirm the assumption that the CSE variants and their semantic extension and/or grammaticalization are due to cross-linguistic influence from the Chinese substrates.

Many grammatical functions (e.g. aspectual functions of *le* and *guò*, universal and concessive meaning of *dōu*, and emphatic function of *de*) are only possessed by the Chinese substrates, but are not shared by the superstrate language, which is Standard English. This raises the question as to why CSE speakers consider these expressions in Chinese as equivalent to *already*, *also*, *ever* and *one* in English. The previous analysis reveals that there is at least one parallel construction that is shared by each of these Chinese expressions and their corresponding expressions in Standard English. For example, English *already*, like Chinese *le*, can also mark the inchoative aspect, though it is limited to occurrences with stative predicates. *Ever* in Standard English also possesses the existential meaning, which mirrors the experiential meaning of *guò*. However, while the existential meaning in English

is restricted to negative contexts, Chinese *guò* can be used in both positive and negative oriented contexts. Furthermore, the existing link between English *also* and Chinese *yě/dōu* is their additive meaning. Finally, the parallel between Chinese *de* and English *one* is their usage as an anaphoric pronoun.

This chapter has also examined the grammaticalization processes of 了 *le*, 过 *guò*, and 的 *de*. Compared to the CSE markers in the focus of the present study, their Chinese substrate equivalents have gone through different grammaticalization paths. According to Heine and Kuteva (2003, 2005), change brought about in the replica language may lead to grammaticalization. The changes of the CSE markers suggest an ordinary contact-induced grammaticalization, which proceeds in a way unrelated to that of the substrate language.

Having established the descriptive generalizations concerning the differences between *already*, *also*, *ever* and *one* in standard varieties of English and in CSE, as well as their similarities with their Chinese counterparts, we are now in a position to outline the database to be used in the analysis of the grammaticalization of *already*, *also*, *ever*, and *one* that follows in Chapter 7.

7 Methodology

This chapter comprises four parts: The first part introduces the corpora used in this study and discusses the reasons for choosing them. The second part illustrates the procedure of extracting the occurrences of *already*, *also*, *ever*, and *one* as well as the methods of calculating the relative frequencies of these four expressions. The third part of this chapter focuses on categorizing and annotating these expressions based on their different functions and sentence positions. Lastly, the final part of this chapter gives a summary of this chapter.

7.1 Database

The analysis of *already*, *also*, *ever*, and *one* in CSE in this study is based on data drawn from the Singaporean component of the International Corpus of English (ICE-SG) and the Oral History Interviews (OHI). The ICE project was initiated in the early 1990s, in which informants were educated speakers of at least eighteen years of age at the time of data collection (see Nelson 2002). The Singaporean component of the ICE project, on the other hand, is a one-million-word corpus containing spoken and written material in roughly equal proportions. The analysis in this study is restricted to the spoken section of the Singaporean component of ICE to compare it with OHI, as the latter data source only includes spoken data. OHI represent speech by informants who were born between 1899 and 1983, and the recordings include detailed metadata. As OHI represent a type of CSE that was spoken at least five decades before CSE was sampled in ICE-SG, a comparison between the two datasets can provide us with useful information on the development of CSE from past to present.

Furthermore, Mandarin Chinese examples will be drawn from the corpus for Chinese linguistic studies compiled by the Center for Chinese Linguistics (CCL) for comparison. The following section provides a finer-grained description of the three corpora.

7.1.1 The Oral History Interviews

The project of OHI was initiated by the Oral History Center (OHC) at the National Archives of Singapore (NAS). The collection represents an important project to document the history of Singapore and can be seen as part of a wider nation-building initiative. The interviews have been recorded since 1979. Most of the interviews contain rich metadata: biographical information – such as age, gender, ethnicity, heritage language, educational background, and occupation of the speakers – precedes the text data. In total, 4,134 interviews (in different languages) have been in the OHI project by the end of November 2020. Among these, more than 1,000 interviews are in English.⁵⁴

One of the most prominent features of OHI is its authenticity and naturality. The interviews are not grammar-oriented. The main genre in OHI is narrative, as the speakers were recounting past events in their personal recollections. They cover various topics including the Japanese occupation, bilingual education, political change, and communities of Singapore (OHI-NAS 2020). All speeches are naturally and orally produced.

Secondly, the ethnic composition of the interviewees – including Chinese, Malaysian, British, Indonesian, and Indian – makes OHI a suitable dataset to investigate recent grammatical developments in the contact varieties of English in Singapore from a typological and contact-linguistic perspective. In addition, the socio-economic status of the informants varies as the informants come from all walks of life in Singapore, including politicians, small business owners, midwives, prisoners of war, engineers, teachers, etc. Interviewees are of different ages, which allows the comparison of speakers of different generations. These rich metadata offer valuable insights into the sociolinguistic domain of CSE and enable us to determine what social factors are most related to the use of CSE markers.

Lastly, OHI well complement the ICE-SG corpus. The Singaporean component of the ICE corpora does not offer metadata of individual speakers, nor does it provide

⁵⁴ The OHI interface does not offer the sum of the English Interviews, but provides the total number of reels of the recordings: 20,575 (OHI-NAS 2020). The average number of the reels per interview is 13 based on 101 interviews.

timestamps of the recordings. As mentioned earlier, OHI provide metadata of the interviewees, and accurately documented the date and time of the interviews. Apart from that, the informants of ICE-SG mainly consist of university students whereas more elderly speakers contributed to OHI. As we attempt to explore the development of CSE across a wide timespan, the comparison between OHI with older speakers and ICE-SG with younger speakers could help bridge the diachronic gap in studying contact-induced grammaticalization of certain lexical items in CSE.

7.1.2 The selection of the informants of OHI

The collection of OHI is not a dedicated linguistic corpus. It is primarily designed for history documentation purposes, but not for linguistic research. Therefore, the data of OHI are not linguistically tagged. It is necessary to manually annotate and count the occurrences of the target tokens. Apart from that, all data collection has to be conducted via the online interface of the National Archives of Singapore. Given the large volume of the data, the present study chose 100 interviews from the recordings of OHI for the analysis, which amount to 16,147 pages of running texts.⁵⁵ These 100 interviews necessarily represent an informed convenience sample, which was selected based on the following criteria: (i) the transcriptions of the interview recordings and the metadata of the interviewees are available and accessible via the online interface of the Archives; (ii) the interviewees produce markers of CSE, such as sentence-final *already* and aspectual *already*, resembling the Chinese substrates (see Chapter 5.1.1); (iii) a higher number of older speakers are included in the samples; (iv) there is a mix of speakers coming from different ethnic backgrounds such as Chinese, Peranakans, Malays, Indians and Eurasians; (v) speakers with a British ethnic background are included as a quasi-control group; and (vi) speakers represent different levels of education. Table 7.1 shows a sample of the biographical information of 15 speakers (see Appendix I for a full list of the 100 speakers selected in the study).

⁵⁵ Before around 2015, it was possible to use the pdfs of the interviews to approximate the word tokens contained in them and thus their length. The National Archives of Singapore have meanwhile enforced a more restrictive access policy so that the length of the interviews now needs to be given in terms of their overall page number.

ACC ID	Initials	Birth year	Birthplace	Sex	Ethnicity	Language	Education	Interview year
000001	TMK	1921	Malaysia	m	Chinese	Hokkien	medium	1979
000013	SEA	1918	Singapore	m	Malay	Malay	medium	1980
000021	CCS	1932	Singapore	m	Chinese	Chinese	high	1980
000057	LYC	1906	China	m	Chinese	Teochew	medium	1981
000071	CKM	1922	Malaysia	m	Peranakan	Chinese	high	1981
000095	MB	1920	Malaya	m	Eurasian	English	high	1993
000123	NB	1933	Indian	m	Indian	Tamil	low	1981
000213	EQ	1899	Singapore	f	Chinese	Hokkien	low	1982
000237	HCY	1905	China	f	Chinese	Cantonese	medium	1982
000265	LTS	1925	Singapore	m	Chinese	Chinese	low	1983
000259	AJ	1915	UK	m	British	English	low	1983
000404	RW	1948	UK	m	British	English	medium	1984
001953	LAS	1940	Singapore	m	Chinese	Chinese	low	1997
002206	MH	1937	Singapore	f	Chinese	Hakka, Cantonese	low	1999
003409	LSC	1983	Singapore	m	Chinese	Mandarin	high	2009

Table 7.1: Metadata of selected Oral History Interviews (a sample of 15 speakers)

The oldest speaker in the selected database was born in 1899 whereas the youngest speaker was born in 1983. Thus, the selection covers approximately 100 years of the development of CSE. Yet, most speakers were born between the 1920s and 1940s. As shown in Table 7.2, the average year of birth of the selected interviewees is 1927. The selected interviews were conducted from 1979 to 2009. The length of the interviews varies, with the shortest one consisting of only 12 pages and the longest one amounting to 651 pages.

	Born in	Interviewed in	Length (pages)
min	1899	1979	12
max	1983	2009	651
mean	1927	1992	161

Table 7.2: Overview of year of birth and year of interviews conducted

The interviews can be divided into seven groups according to the ethnic background of the interviewees, consisting of Chinese, Indian, Malay, British, Peranakan, Eurasian, and Other. Table 7.3 provides an overview of these groups and their respective sample sizes. Much as in the wider population of Singapore (see Chapter 3.2), the Chinese subsample contains the

highest number of speakers (n=57), whose heritage languages are Chinese, Hokkien, Hakka, Teochew, and Cantonese. The second group of 12 interviewees are of Indian ethnic background. The third group consists of 8 interviews with speakers of Malay ethnic background. The Peranakan group consists of 8 speakers. The British group – Singaporean with a British ethnic background – serves as a control group of 8 interviews as well, followed by 5 speakers with a Eurasian ethnic background. The last group “Other” includes 2 informants, one from Iraq and the other having a mixed ethnicity of Chinese-Indian (Chindian).

Ethnicity	Interviews	Length (pages)
Chinese	57	8688
Indian	12	2699
Malay	8	1129
Peranakan	8	1425
British	7	951
Eurasian	6	587
Other	2	668
Total	100	16147

Table 7.3: Seven groups of interviewees according to ethnic background

The summary of the metadata of the selected interviews is listed in Table 7.4. The gender ratio of the interviewees is around 4 male speakers per 1 female speaker (79 male, 21 female). The majority of the participants were born in Singapore (n=65), followed by Malaysia (n=14), India (n=6), the UK (n=5), and China (n=4). In view of the education level, more than half (52%) of the participants received secondary (high school) education. Over 36% of the interviewees finished higher education, earning college and university degrees, whereas 12% of them only finished primary school. Language profile is not listed in the metadata pages but can be elicited from the content of the interviews. The informants speak various types of English that fall into a lectal continuum. The majority (66%) of the interviewees can speak an unspecified vernacular of Chinese, among which 26% were noted as Hokkien speakers, 16% as Cantonese speakers, and 3% as Teochew speakers. One

speaker reported himself as a Hakka speaker. The second-largest cohort is the Malay speakers (n=20), followed by the Tamil speakers (n=13). Other languages spoken by the informants include languages such as Scottish, Siamese, Burmese, German, Baba-Malay, Hindi, Urdu, and Pashto.

Social category		Percentage
Gender	male	79%
	female	21%
Ethnicity	Chinese	57%
	Indian	12%
	Malay	8%
	Peranakan	8%
	British	7%
	Eurasian	6%
	Other	2%
Birthplace	Singapore	66%
	Malaya (now Malaysia)	14%
	China	4%
	UK	5%
	India	5%
Education	low	12%
	medium	52%
	high	36%
Language spoken	Chinese	67%
	Chinese (Hokkien)	27%
	Chinese (Cantonese)	17%
	Chinese (Teochew)	3%
	Chinese (Hakka)	1%
	Chinese (Hainanese)	1%
	Malay	20%
	Tamil	12%
	English	100%
	Other	7%

Table 7.4: Summary of metadata of selected interviews in OHI

The percentages of the category “language spoken” add up to more than 100% because many informants of OHI speak more than one language. As mentioned earlier, multilingualism is the norm in Singapore, and this is reflected in the linguistic profiles of the speakers who

contributed to OHI. For example, most of the Chinese interviewees included in OHI speak Mandarin Chinese as well as at least one of the Chinese vernaculars such as Hokkien and Cantonese in addition to English. If they happened to live in a Malay neighborhood, they would speak some Malay as well (see OHI-002108-TSG).

7.1.3 ICE-SG

ICE-SG (International Corpus of English – the Singaporean component) is a one-million-word corpus containing spoken and written material in roughly equal proportions. The project was initiated in 1997 by members of the Department of English Language and Literature at the National University of Singapore. Three criteria have to be met for speakers and writers to be included in the ICE corpora: (i) they have to be native speakers of the variety; (ii) they have received English-medium education in that country; and (iii) they have to be at least 18 years old (Greenbaum and Nelson 1996:5). As the ICE corpora were designed for cross-regional comparison of different varieties of English, but not dedicated for sociolinguistic studies, not every component of the ICE corpora contains metadata of the speakers. Besides, the sociolinguistic structure of a sub-corpus often only represents a particular group of the whole population. ICE-SG, for example, does not offer detailed metadata of the informants and it reflects language use of educated speakers who were either born in Singapore or moved there at an early age and received education through English in Singapore (Nelson 2002:3).

This study only includes the spoken section of the ICE-SG to compare with the spoken data drawn from OHI. The spoken section (approximately 661,000 words) of the ICE-SG emulates the day-to-day conversations of the colloquial register of Singaporeans. It is further divided into three subcategories: (i) the private dialogue section (~216,000 words), which includes face-to-face conversations and phone calls (corresponding filenames S1A-001 to S1A-100); (ii) the public dialogue section (~181,000 words), which includes e.g. classroom lessons and business transactions (corresponding filenames S1B-001 to S1B-80); and (iii) the monologue section, which includes unscripted speeches (~147,000 words, e.g.

spontaneous commentaries and legal presentations) corresponding to filenames S2A-001 to S2A-070, as well as scripted talks (~116,000 words, e.g. broadcast news and broadcast talks) corresponding to filenames S2B-001 to S2B-050).⁵⁶

7.1.4 CCL Corpus

The CCL Corpus is the largest corpus for Chinese linguistic studies compiled by the Center for Chinese Linguistics (CCL) at the University of Peking, China. This study uses the Modern Chinese section of the CCL Corpus, which consists of 477 million Chinese characters. It has been collecting samples of Modern Chinese since 1919, both written and spoken. It covers genres such as literature, TV, movie, newspaper articles, and historical biography.

To keep the results comparable with OHI and ICE-SG, I focused on the spoken section of Modern Chinese in the CCL Corpus. The study of the CCL Corpus will target the Chinese counterparts of the investigated CSE expressions. The CCL Corpus is equipped with an online interface to allow users to extract linguistic generalizations. The online interface can show a full output sentence up to 200 Chinese characters in length. In terms of the display of extracted data, CCL limits data to 500 lines per page but allows downloads of output data. However, it does not have part of speech (POS) tags, which makes it difficult to retrieve the frequencies of the substrates. In addition, the Chinese counterparts are often either polyphonic or polysemous characters, which increases the difficulty of finding the desired targets.

The problem is solved by manually filtering out compounds of the targeted expressions, as well as characters in other pronunciations and meanings to extract possible patterns of the Chinese substrates.

⁵⁶ The word count was retrieved by Antconc Wordlist (“word tokens”). The exact word counts are: 661,004 for ICE-SG (216,305 for private dialogues, 181,279 for public dialogues, 147,192 for unscripted monologues, and 116,228 for scripted monologues).

7.2 Procedure

7.2.1 Searching the linguistic variables

The study adopts a “corpus-based approach” (see Anderson 2016) where the four CSE expressions (*already*, *also*, *ever*, and *one*) are searched for in OHI and ICE-SG. For the occurrences in OHI, I used the online search engine provided by the Archives, which allows users to enter keywords to search for targeted expressions. After typing in the search strings, the online interface displays the targeted keywords highlighted in yellow in their contexts. I noted down in Excel the sentence in which the targeted expressions occur.

As for ICE-SG, it is freely downloadable and can be used for non-profit linguistic research purposes. As the text files of ICE-SG are machine-readable, I used the concordance program of AntConc 3.5.8 (Anthony 2019) for the extraction of the tokens of *already*, *also*, *ever*, and *one* found in ICE-SG.

In sum, 18,198 tokens were elicited in the selected 100 interviews, with *already* (n=3,676) and *also* (n=11,293) being the most frequently used ones among the four CSE expressions. For the tokens of *one*, I only focus on the clause-final *one* (CF-*one*) due to the large number of occurrences of *one* in the data, with most of them serving a numeral function, which is not the major interest of this study (see Chapter 5.4 and Chapter 6.5). In ICE-SG, on the other hand, I found 489 occurrences of *already*, 1364 occurrences of *also*, 85 occurrences of *ever*, and 522 occurrences of phrase-final *one*. Table 7.5 provides an overview of the absolute frequencies of these four expressions in OHI and ICE-SG.

	OHI	ICE-SG	Total
<i>already</i>	3,676	489	4,165
<i>also</i>	11,293	1,364	12,657
<i>ever</i>	760	85	845
CF- <i>one</i>	2,469	522	2,991
Total	18,198	2,460	20,658

Table 7.5: Overview of absolute frequencies of *already*, *also*, *ever*, and *one* in OHI and ICE-SG

The absolute frequencies of these expressions are substantially higher in OHI than ICE-SG. However, we need to bear in mind that OHI is a much larger corpus than ICE-SG. A rough estimation of the total word count of the 100 selected interviews from OHI is about 4.69 million words (see Section 7.2.2). On the other hand, the total word count of the spoken section of ICE-SG equals 661,004 (see Chapter 7.1.3). This means that the selection of OHI represents a corpus more than 7 times larger than ICE-SG.

In the next step, I calculated the relative frequencies of these four linguistic variables, which allows us to identify the changes of their distributions in different functions and discover possible associations between the relative frequencies of these expressions and additional background variables, such as ethnicity, education, gender, and age groups.

7.2.2 Calculating the relative frequencies of the linguistic variables

The average word count per page of the selection of OHI is 290.74. This is achieved by manually counting the number of words of the first page, the middle page, and the last full page of each interview. As mentioned earlier, the total pages of the selected interviews add up to 16,147. Therefore, the total word count of the 100 selected interviews amounts to approximately 4,694,579 words. However, the study adheres to the number of occurrences per page (pp) to measure the relative frequencies of the variables in OHI, as there are several inconsistent factors influencing the extraction of the total word counts of the selected interviews of OHI, e.g. inconsistent font, page size, and space among different interviews. On the other hand, the relative frequencies of the four CSE expressions in ICE-SG are calculated based on the measure of per thousand words (ptw). Table 7.6 lists the relative frequencies of the four expressions across the two corpora:

	OHI (16,147 pages)		ICE-SG (661,004 words)	
	N	pp	N	ptw
<i>already</i>	3,676	0.23	489	0.74
<i>also</i>	11,293	0.70	1,364	2.06
<i>ever</i>	760	0.05	85	0.13
CF- <i>one</i>	2,469	0.15	522	0.79

Table 7.6: Relative frequencies of *already*, *also*, *ever*, and *one* in OHI and ICE-SG

The next step was to code the relevant tokens according to their position and semantic meanings. The extracted tokens were manually coded in their given context, since the context is vital in determining the semantic meanings of a given token. The next section gives a detailed description of the categorization of the four CSE expressions.

7.3 Coding the linguistic variables

This section elaborates on the four CSE expressions *already*, *also*, *ever*, and *one* as well as their categorization. There is a certain amount of overlap with Chapter 5, but this section focuses more on the methodological perspectives of the above expressions, i.e. how they were categorized based on their sentence positions and meanings.

7.3.1 Already

I looked for the occurrences of *already* by using the search engine provided by the online interface of OHI. The sentences in which *already* occurs were noted down in Excel. The CSE variants of *already* are mostly located in sentence-final position while Standard English *already* usually occurs in pre-predicate position (see Chapter 5.1). Therefore, an important parameter for the annotation is the position of *already*. I annotated the position – i.e. initial, medial, or final – of *already*. Some typical examples are demonstrated in (141):

- (141) Classification of *already* according to sentence position
- a. Initial: *Already* there were controversies... [OHI-002307-VTA]
 - b. Medial: There was *already* a Chinese newspaper called... [OHI-000064-SCY]
 - c. Final: They can get married *already*. [OHI-002686-SMY]

Following that, I classified the meanings of *already* in each occurrence according to whether *already* functions as an aspectual marker, which gives rise to the readings of completive/actual aspect, inchoative aspect, and prospective aspect as shown in (142) (also see the theoretical discussions on *already* in Chapter 5.1). Special focus was also given to those occurrences where *already* occurs in negative sentences though this category overlaps with the semantics of the inchoative, as exemplified by (143).

- (142) Aspectual markers (inchoative, completive, and prospective):
- a. Ah, until I finish *already*, then went off. [OHI-003206-MA] (+completive)
‘Ah, after I finished the task, I went away.’
 - b. They, from young they learn *already*. [OHI-001109-MN] (+inchoative)
‘They started to learn this skill at an early age.’
 - c. They can get married *already*. [OHI-002686-SMY] (+inchoative)
‘They can get married now (which was not the case before).’
- (143) *already* in negative sentences:
The Japanese money was of no use *already* after the surrender. (+negation, +inchoative)
‘The Japanese currency was no longer in use after the surrender of the Japanese (it was in use before the surrender).’

During the annotation, *already* was found in pre-adjective position which modifies the adjective that it precedes, e.g. *already* established, *already* existing. Such usages were well-attested in mainstream varieties of English, for instance in the Corpus of Contemporary American English (COCA) (Davies 2008-), as shown in (144). Therefore, pre-adjective usages were considered as standard usages.

- (144) a. [...] rule based on the risk – or perhaps the perception – that police were violating *already-established* law.
[COCA-2018-ACAD-American Criminal Law Review]

b. I have discussed the decision to frame research around *already-existing* data at length elsewhere [...] [COCA-2019-ACAD-Geographical Review]

Apart from the aspectual categories, I created a discourse-pragmatic category to explore the possibility of CSE *already* functioning as a discourse marker. A category named “Other” was also created to capture usages that were not yet described in previous scholarly literature. In addition, instances of *already* without further context were placed in the category of “Unclear”, as shown in (145):

(145) [...] in Sarawak, no money too. So for two, three months we were *already*...
Then of course after the third month probably you went out to [...]
[OHI-00485-LLH]

The following table lists all the categories of the semantic interpretations of *already*, some examples are listed in the right column.

	Category	Example
(a) Types of predicates	(i) stative (ii) non-stative (iii) modal verbs	was <i>already</i> , know <i>already</i> finish <i>already</i> , learn <i>already</i> can <i>already</i>
(b) Aspectual marker	(i) completive/actual aspect (ii) inchoative aspect without connotations of anteriority (iii) prospective aspect	
(c) Negative sentence		No use <i>already</i>
(d) Discourse functions		That is all <i>already</i>
(e) Other	(i) lack of tense/inflectional markers (ii) double use of <i>already</i> (iii) copula <i>be</i> omission/predicate ellipsis (iv) word order change (v) with <i>still</i>	It finish <i>already</i> <i>Already</i> ... <i>already</i> So big <i>already</i> I engage <i>already</i> two people. It was <i>already</i> ... still
(f) Unclear	single usage or utterance without further context	It was <i>already</i> ...

Table 7.7: Overview of categories of the semantic interpretations of CSE *already*

Category (a) first classifies the predicates in which *already* occurs. The classification of the predicates helps to determine the aspectual functions of *already* (see Chapter 5.1.3). Categories (b)–(e) are regarded as substrate-influenced features of *already*, while category (f) represents unclear and ambiguous cases of *already*.

The meanings and functions of *already*, especially the aspectual meanings, were determined in the textual context in which *already* occurs. I paid extra attention to the sentences directly before and after the occurrence of *already*, as well as the sentence where *already* appears in. For example, in the case of (146), *already* is an inchoative marker. The context was crucial for this interpretation. In this example, the speaker recounts that his brother went to the Royal College to study and he went there to visit his brother. The sentence in focus is “He *already* studied there.” Both the inchoative reading ‘he started to study there’ and the completive reading ‘he finished his studies there’ are possible. Yet, given the context, it would only make sense that the speaker thought that his brother was still studying at the Royal College as the speaker went to the Royal College to visit him. Therefore, *already* was categorized here as an inchoative marker instead of a completive marker.

- (146) He went to the Royal College or something like that. He *already* studied there. But after sometime I went there, I can’t find him. And in the end we found him in one of the hotels – as a kitchen boy. [OHI-000009-LGS]

In addition to marking the completive aspect, *already* seems to have developed a pragmatic function, resembling that of the Chinese substrates. Li et al. (1982) noted that Chinese *le* can be used to close a statement and named such function as “finality marker” (see Chapter 6.2.3). I observed similar usage of CSE *already*, in which it serves as a piece of information that the comment made by the speaker is all that he has to contribute at the moment, as exemplified in (147):

- (147) Interviewer: What about the workers, they just stand by?
 LAS: They don’t do anything. They stopped *already*.
 ‘They don’t do anything. They have stopped. (That is all I can contribute to the topic at the moment).’ [OHI-001953-LAS]

Occurrences of *already* which demonstrate morphosyntactic incompatibility with Standard English, yet do not belong to the aforementioned categories were categorized in the group “Other”. This includes, for example, cases in which the predicate that *already* modifies remains in its bare, unmarked form as shown in (148). Use of *already* was also found in cases where the copula *be* is absent as in (149). Although the default assumption is that there is only one occurrence of aspectual *already* per clause, actual usage shows that speakers may also produce *already* twice in the same clause. For more on the discourse functions and other functions of *already*, see Chapter 8.1 for the analysis.

- (148) Lack of tense/inflectional markers:
- a. But the water was not very. Quite high. It is up to the chest *already*.
[OHI-002206-MH]
 - b. So obviously there were smell? Ah, now, it’s cleared *already*. Since that time they have cleared a lot. [OHI-001953-LAS]
- (149) Copula *be* ellipsis:
- a. The Government’s requisition *already* one million over.
[OHI-000057-LYC]
 - b. Where got the energy? Even our mind, all blank *already*.
[OHI-001953-LAS]
- (150) Double use of *already*:
- a. They were *already* actually piling their arms *already*. [OHI-000001-TMK]
 - b. After all this news going on, after the screening, the people were *already* living in fear *already*. [OHI-000265-LTS]

7.3.2 Also

In sum, 11,293 tokens of *also* were extracted from the 100 interviews in OHI. It was necessary to limit the research scope for the qualitative analysis. Therefore, the study focuses on tokens of *also* in phrase- and sentence-final position, as it has been reported in the literature that substrate-derived *also* occurs more frequently in sentence-final position (see Chapter 5.2). These tokens of *also* were elicited by using search strings of *also* succeeded by the following punctuation marks: full stop (.), comma (,), semicolon (;), and question

mark (?).⁵⁷ These are exemplified in (151). In addition, I included tokens of phrase-final *also* occurring directly before a conjunction such as *and*, *or*, and *because*,⁵⁸ as well as an adverbial starting with *when*.⁵⁹ These are exemplified in (152).

- (151) *also* followed by
- a. Full stop: They brought their own boats *also*. [OHI-000263-S]
 - b. Comma: We had the report book *also*, yes. And must be signed by our parents. [OHI-000211-LKT]
 - c. Semicolon: Children, a lot *also*; Actually, it was mainly [...] [OHI-000081-K]
 - d. Question mark: Can you describe this *also*? [OHI-001631-OCN]
- (152) *also* followed by conjunctions and adverbials
- a. And: We always make sure that the neighbours get some *also* and all that. [OHI-002108-TSG]
 - b. Or: Even my place the Japanese will come at night *also* or during the day, anytime they like. [OHI-002017-FKS]
 - c. Because: Or my father must have known *also* because he was the old type. [OHI-000237-HCY]
 - d. When: And then of course we have field work *also* when we do village survey. [OHI-002107-CPK]

In a subsequent step, I annotated the tokens of sentence-final *also* according to their substrate-influenced features, including (a) *also* occurring in negative sentences, (b) with universal quantifiers including *all*, *always*, *every**, and *any**,⁶⁰ and (c) with the concessive adverb *even*, as exemplified in (153):

- (153) a. In negative sentences: You, one person go and eat? Cannot *also*, you know. [OHI-001953-LAS]
- b. With universal quantifiers: [...] like that. And this song you can hear *everywhere* in cinemas *also*, very patriotic song. [OHI-000483-TTC]
- c. With concessive *even*: Malays? I think, just...not only the Malays, *even*

⁵⁷ I also attempted to search tokens of *also* followed by the exclamation mark (!), but no result was found in OHI.

⁵⁸ The conjunction *but* was searched for, yet all the tokens of *but* occur after punctuation marks, which were already included in the previous round of research.

⁵⁹ The tokens of *also* followed by an adverbial starting with *with* were searched for, but no occurrence was found in OHI.

⁶⁰ The entry *every** includes tokens such as *every*, *everyone*, *everybody*, *everything*, *everywhere*, and *every time*. The same applies to the entry of *any**.

the Chinese *also*, even everybody during the Japanese, they suffered a bit.
[OHI-000242-II]

7.3.3 Ever

Following Israel (1998), I annotated three functions of *ever* as standard usages: existential, universal, and emphatic.⁶¹ As previously discussed in Chapter 5.3.1, *ever* with the existential meaning is a negatively-polarized item which is restricted to negative contexts meaning ‘even once’. Existential *ever* is realized in *have ever V-ed/did ever V* in an interrogative, in a negative sentence, or after *if*, as exemplified in (154). Apart from that, the universal usage – meaning ‘constantly’ or ‘at all times’ – is realized as *ever since/as ever/ever+adj*, as shown in (155). The emphatic *ever* is realized as *ever so*, as in (156). Other usages that belong to the emphatic function discussed in Israel (1998) were searched, e.g. in inversion exclamatives as in *Was I ever shocked!* and in rhetorical *wh*-questions, as in *What ever did you expect?* However, no examples in these emphatic usages were found in either OHI or ICE-SG.

(154) Existential *ever*:

- a. Questions: Did you *ever* ask her how she came to be in Singapore?
[OHI-000371-BNT]
- b. Negatives: Nobody *ever* apologized to us. [OHI-003223-CSS]
- c. After *if*: But if *ever* a person deserves recognition, he does.
[OHI-001309-MG]

(155) Universal *ever*:

- a. *Ever since* I moved in *ever since* I *ever since* I moved in I never paid [...]
[ICE-SG:S1A-040#88:1:B]
- b. Oh fine *as ever* I’m glad to say. [ICE-SG:S1A-026#103:1:A]
- c. The burden of the *ever increasing* expenditures ultimately falls back on the people. [ICE-SG:S2B-048#X124:2:B]

(156) Emphatic *ever*: Because I’ve heard it’s *ever so* easy.
[ICE-SG:S1A-019#100:1:A]

⁶¹ In this study, we only focus on *ever* as a stand-alone morpheme. The derivational usages as realized in *whenever*, *whatever*, *whoever*, and *wherever* were excluded from this study.

With regard to CSE *ever*, the realization is *ever-V*, which functions as an experiential marker (Bao 2005:238; Leimgruber 2013:81; Ziegeler 2015:120). The use *ever* in affirmative responses to a polar interrogative was searched (see Ho and Wong 2001; also Chapter 5.3.3), as in (158), though no examples of such usage were found in OHI or ICE-SG. Again, a category “Other” was created to discover innovative usages that have not yet been discovered in scholarly circles.

- (157) Aspectual *ever* (experiential): I mean I *ever* see that being mentioned in the media. [OHI-003223-CSS]
- (158) *ever* in affirmative responses to polar interrogatives:
 A: Your husband *ever* bring fish home to eat or not?
 B: *Ever*. (Ho and Wong 2001:81)

7.3.4 One

This study focuses on phrase-final *one*, as substrate-influenced *one* usually occurs in phrase-final position (see Chapter 5.4). As with *also*, phrase-final *one* was elicited through search strings followed by punctuation marks and conjunctions (see Section 7.3.2). As we discussed in Chapter 5.4, CSE *one* can follow nominal words (e.g. N-one: silk *one* ‘a dress made of silk’), and possessive pronouns (e.g. my, your, his/her), which are ungrammatical in Standard English, as exemplified in (159):

- (159) a. N-*one*: That one’s his own la silk *one*. [ICE-SG:s1a-002#187:1:B]
 b. Possessive pronoun-*one*: Alright my *one* and only one I bought from Hong Kong – the washable silk. [ICE-SG:S1A-003 #X323:1:C]

Apart from that, we include the annotation of *one* as a relativizer/nominalizer and phrase-final emphatic marker, as exemplified in (160) and (161). These functions are features of CSE *one*, but not of Standard English (see Chapter 5.4.2 and 5.4.3). The surface syntactic structure of *one* as a relativizer/nominalizer and emphatic *one* is identical, yet their semantic interpretations are different. While emphatic *one* is attached to sentences or phrases,

relativizer/nominalizer *one* forms an NP with the preceding modifier (see Chapter 5.4 on *one* and Chapter 6.5 on the Chinese counterpart *de*).

- (160) Relativizer/nominalizer *one*
- a. But all girls lah I mean those CUT group *one*. [ICE-SG:S1A-091#296:1:A]
‘All girls, I mean those girls who come from the CUT group.’
 - b. Those wear black *one*, is it? [ICE-SG:S1A-023#334:1:B]
‘Those people who wear black’
- (161) Emphatic *one*
- a. Our glasses also cannot shine *one*. [ICE-SG:S1A-065#71:1:B]
 - b. I always speak like this *one*. [ICE-SG:S1A-012#27:1:D]
 - c. Because I always wake up early *one*. [OHI-001953-LAS]

Lastly, the number of tokens where standard uses of *one* appear in phrase-final position was also counted. These include numeral *one* and pronominal *one* (see Chapter 5.4.1). This will help us answer the questions of (i) whether Chinese speakers use phrase-final *one* over a wider range of contexts as compared to speakers of other ethnicities, and (ii) whether there is a difference between OHI and ICE-SG in the ratio of substrate-influenced *one* as compared to its standard uses.

7.4 Summary

The purpose of this chapter was to elaborate on the database and design of the study as well as to provide an overview of the major methodological issues involved in extracting the data. Furthermore, the intention is to explore the change of *already*, *also*, *ever*, and *one* in CSE as a result of the historically continuous contact between English and the Chinese substrates over the past 100 years. In addition, the study attempts to discover how the functional changes of the four linguistic variables interact with social factors, i.e. ethnicity, educational level, age, and gender. As put by Siemund (2010:4), “language contact situations usually differ in terms of their social constellation.” Thomason (2001:21) also emphasizes that social factors are crucial in assessing stability of contact-induced language change (Thomason

2001:21). Therefore, we predict that the diachronic development of CSE is not only subject to the linguistic principles determining contact-induced language change (see Chapter 4), but also strongly influenced by social factors.

The following chapter continues with the analysis and the results based on the data extracted from OHI and ICE-SG. I will present the analysis of *already*, *also*, *ever*, and *one* individually before demonstrating how their changes in use are subject to social factors.

8 Results and analysis

This chapter examines the frequency distributions of *already*, *also*, *ever*, and *one* as well as their various semantic interpretations across OHI and ICE-SG (see Chapter 7). We also focus on their frequency distributions in various syntactic positions and semantic categories according to different social variables, i.e. ethnicity, educational level, age group, and gender.

The results suggest that CSE has remained relatively stable. In addition, we observe an ongoing grammaticalization of these four CSE markers, as the four expressions are becoming increasingly unrestricted in their functional range and syntactic position. The functional extensions of these CSE markers are, to a large extent, related to their Chinese counterparts. Apart from that, we observe significant individual differences in the uses of these CSE markers amongst speakers of OHI. Among the social factors, ethnicity and educational level appear to be the strongest predictors in the substrate-influenced uses of the four linguistic variables. In Sections 8.1–8.4, we examine the use of *already*, *also*, *ever*, and *one*. Section 8.5 gives a summary of this chapter.

8.1 Analysis of *already*

As presented in Chapter 7.2.1, 3,676 occurrences of *already* were found in the selection of OHI, while 489 cases were found in ICE-SG. Figure 8.1 provides a general impression regarding the difference in the relative frequency of *already* among three speaker groups from OHI, namely the Chinese and the British group as well as a group of mixed ethnic background referred to as “all other”. In addition, we compare their relative frequencies with those in ICE-SG and ICE-GB.⁶² The results in OHI show that the Chinese group attains the highest frequency of *already* in this dataset (0.31 pp), followed by the mixed group with a frequency of 0.18 pp, while the British group has a frequency of 0.08 pp. This result once again supports the assumption that the use of *already* is more likely to be influenced by the

⁶² The data of ICE-GB is drawn from Siemund and Li (2017:22).

Chinese language than other contact languages. The result is in line with the spoken sections of ICE-SG (0.74 ptw), in which the relative frequency of *already* is significantly higher than that in ICE-GB (0.33 ptw), though OHI focus more on the internal differences regarding the frequency of *already* among speakers of different ethnic backgrounds. Apart from that, the average frequency of *already* in ICE-SG (0.74 ptw) is between the average frequency produced by the OHI interviewees with a Chinese ethnic background (0.31 pp) and that of speakers with other ethnic backgrounds (0.18 pp).⁶³ It seems plausible to conclude that CSE has remained relatively stable amongst the non-British speakers. However, it should be remembered that the speakers in ICE-SG are anonymous and no details about their ethnic background are provided (see Chapter 7.1.3). Therefore, we can only assume that ICE-SG consists of speakers coming from various ethnic backgrounds, including not only Chinese, but also Malay and Indian speakers.

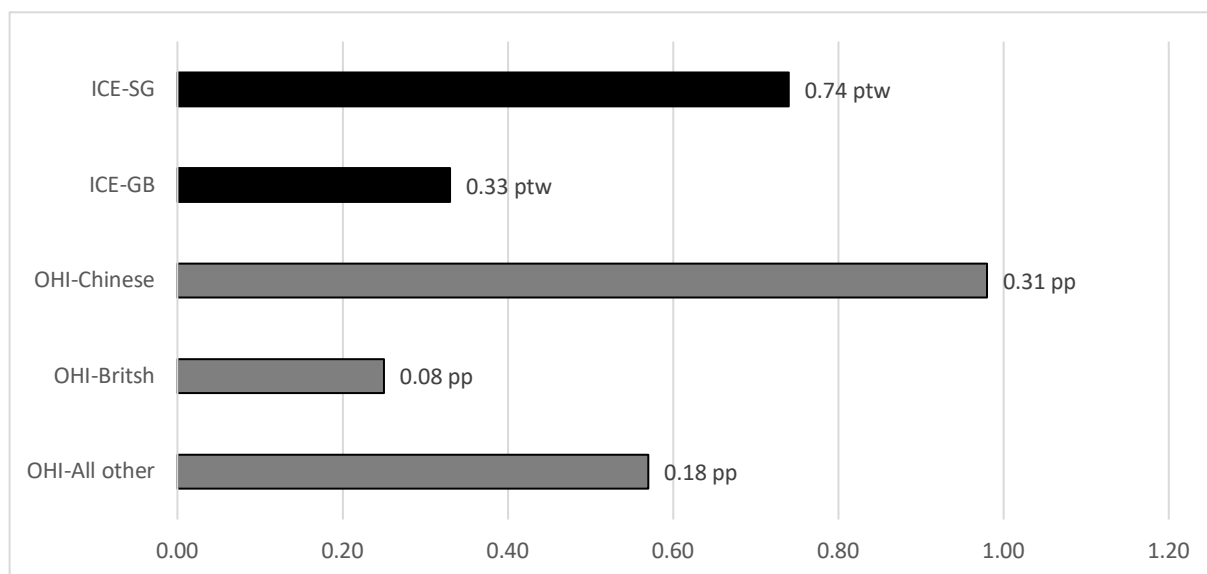


Figure 8.1: Frequencies of *already* in ICE-SG and OHI (ptw=per thousand words, pp=per page)

As mentioned earlier, the data in the Oral History Interviews represents a type of Singapore English at least five decades older than the English sampled in the ICE corpus. The 7

⁶³ The relative frequency of *already* in the Chinese group of 0.31 pp corresponds to 0.98 ptw: Relative frequency of *already* ptw = 2,474 tokens of *already* / (8,688 pages × 290.74 words pp) × 1,000. However, this method only yields a rough estimation (see Chapter 7.2.2).

interviewees with a British ethnic background in the Oral History Interviews were born in the UK but grew up in Singapore while the British component of the ICE corpus represents British English in the home country. What we can observe in Figure 8.1 is that British speakers in both corpora produced relatively low frequencies of *already* while other speakers, especially speakers with a Chinese ethnic background, produced higher frequencies of *already*. Judging from these results, it is not possible to detect a decline in the usage of *already* or a shift towards the British norm.

Calculating the frequency of *already* across the entire corpora neither reveals any information about syntactic and semantic differences of *already* among individual speakers, nor does it show the differences related to their sociolinguistic background such as ethnicity and educational level. Such information, however, would be highly desirable. The next sections demonstrate the syntactic position and the semantics of *already* in CSE as represented in OHI and ICE-SG in more detail, as well as how these are related to speakers' ethnicity and educational background.

8.1.1 Frequencies of *already* by sentence position

The preferred position of *already* in British English is sentence-medial, while CSE *already* favors sentence-final position, especially in less formal registers (see Chapter 5.1). If we look at the distribution of *already* according to their sentence position (initial, middle, and final), we can see that the Singaporean varieties represented by OHI and ICE-SG show a slightly different profile (see Figure 8.2). The proportion found in OHI, with a percentage of 39.6% of *already* in sentence-final position, is slightly smaller than the proportion of *already* found in ICE-SG, which accounts for 46.7% of the uses of sentence-final *already*. A Pearson's 2x2 chi-square test without Yate's continuity correction shows that the differences between OHI and ICE-SG in terms of the distribution of *already* in sentence-final and sentence-middle position are statistically significant ($p=0.0015$).

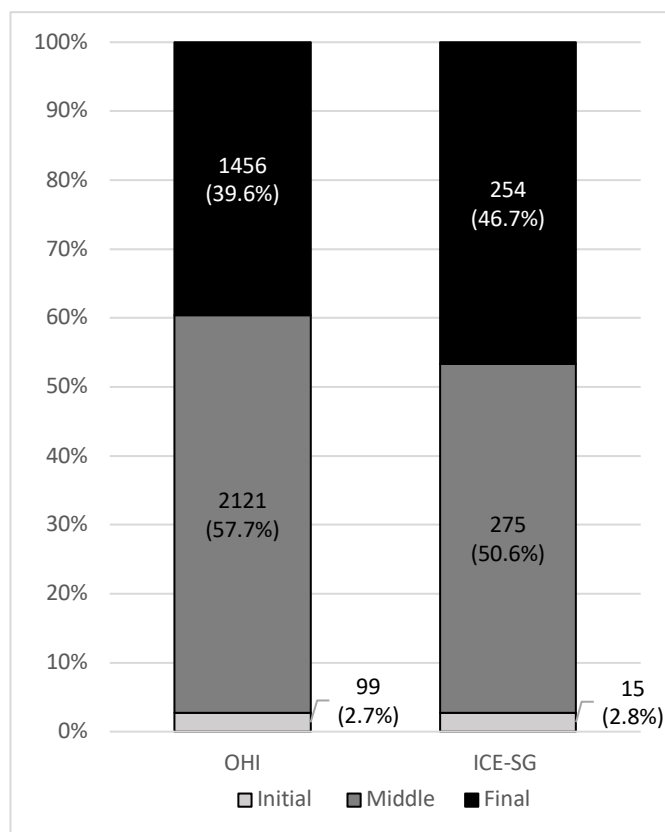


Figure 8.2: The proportion of *already* in initial, middle, and final position in OHI and ICE-SG

If we compare the distribution of *already* in different sentence positions in OHI with that in the private and public dialogues, and monologues of ICE-SG, a different pattern can be observed (see Figure 8.3). As mentioned earlier (see Chapter 5.1), the private dialogues represent a less formal register than the public dialogues and monologues, and informal speech is known to favor the use of sentence-final *already* (see Bao and Hong 2006:109). The result shows a preference of sentence-final *already* (64.3%) in the private dialogues in ICE-SG. Compared with this, OHI has a lower share of *already* (39.6%) in sentence-final position. A 2x2 chi-square test without Yate's continuity correction was performed to examine the difference between OHI and the private dialogues of ICE-SG in final and non-final use. Again, the difference is considered to be statistically significant (chi-square=79.272, $p < 0.0001$).

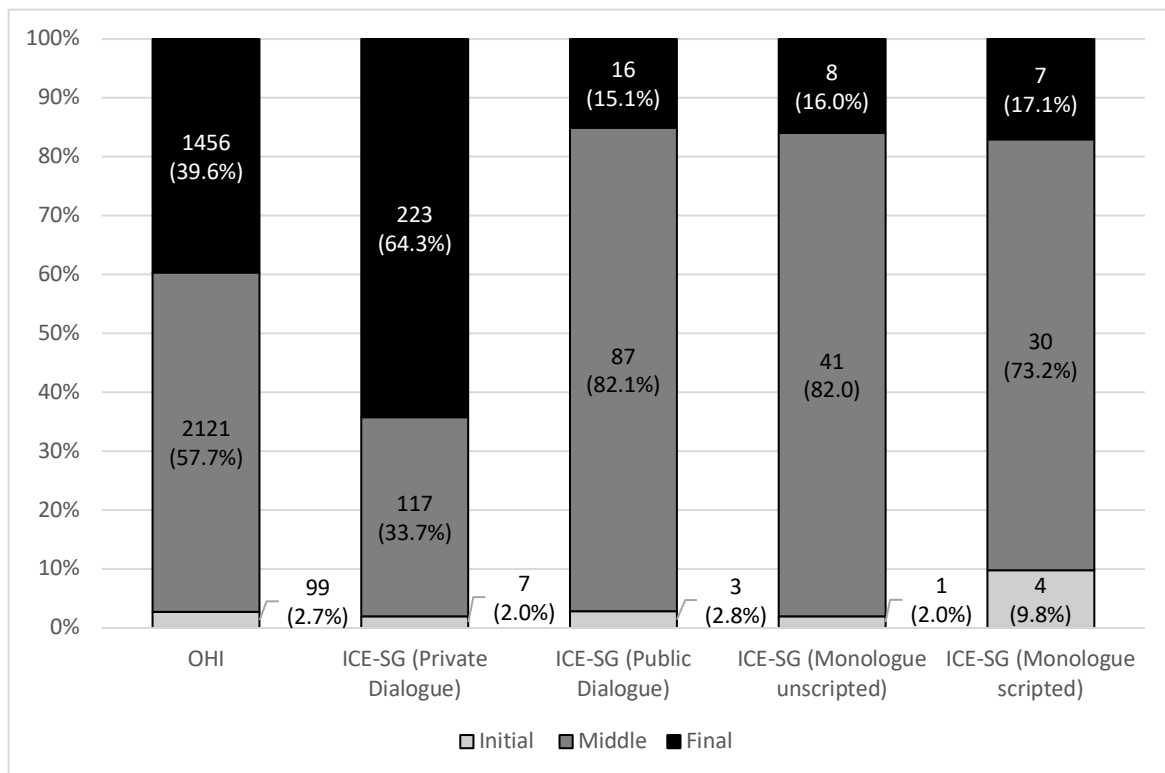


Figure 8.3: The proportion of *already* in initial, middle, and final position in OHI and ICE-SG (according to different text types)

In sum, the distributional difference in sentence position of *already* between OHI and the ICE-SG private dialogues seems to suggest an increase in sentence-final *already* over time in CSE. However, it is necessary to take into consideration the differences according to different text types.

Since register is an important factor in the frequency of sentence-final *already* in a sentence, I examined the proportion of *already* in different sentence positions according to lectal groups. It is difficult to measure the lectal level, as defining a person’s lectal group is a subjective decision. However, we can rely on the measurement of educational level, as a person’s socio-economic background and educational level determine the language variety that they speak (Platt and Weber 1980:108–35). Therefore, we can produce a rough estimate of the variety of speakers’ language use (basilect, mesolect, and acrolect) by measuring their educational level.

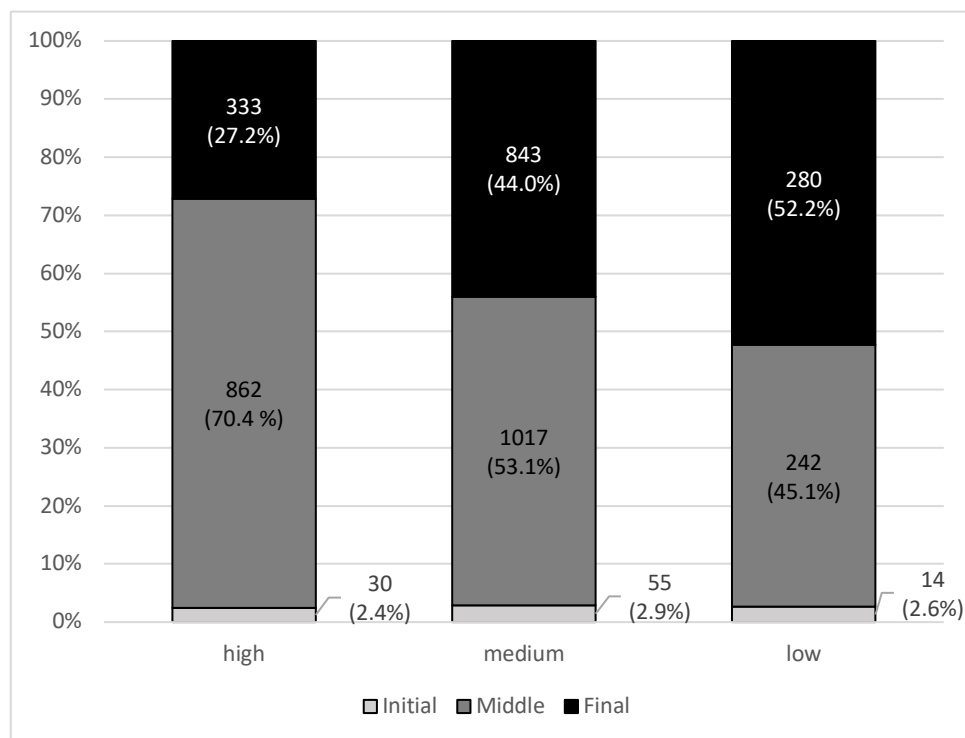


Figure 8.4: The proportion of *already* in initial, middle, and final position in OHI according to educational level

As can be seen in Figure 8.4, speakers who have a low educational level tend to use *already* more frequently in sentence-final position (52.2%) than speakers with a higher educational level, as represented by speakers with a medium (44.0%) and a high educational level (27.2%). This result confirms the assumption that the frequency of sentence-final *already* can serve as an indicator of individual lects (chi-square for 3x3 contingency table=134.9693, $p<0.00001$). This result is in line with Bao and Hong (2006) that the preference for *already* in sentence-medial position – in both British English and CSE – declines as register moves from formal to informal. In their words, “the more formal the register is, the higher the chance of *already* [...] being used in medial position” (Bao and Hong 2006:112). The phrase-final position of *already* parallels that of the Chinese counterpart – sentence-final 了 *le* (see Chapter 6.2 on *le* in Chinese).

Judging from these results, the pattern of change in CSE does not precisely agree with Schneider’s (2007) Dynamic Model on CSE (see Chapter 4.6). Schneider’s (2007) model hypothesizes that CSE was in stage 2 (exonormative stabilization) during 1867–1942, and proceeded with stage 3 (nativization) between 1945 and the 1970s, and reached stage 4

(endonormative stabilization) in the 1970s (Schneider 2007:155–61). Here, the dataset of OHI represents an earlier sample of CSE, as most speakers were born between the 1920s and 1940s (see Chapter 7.1.2). Therefore, CSE as presented by OHI can be assumed to belong to the earlier stages – stage 2 (exonormative stabilization) and stage 3 (nativization) – in Schneider’s (2007) Dynamic Model. However, both the old and the new CSE samples diverge from British English, which suggests that CSE is not moving towards British English. Apart from that, OHI and ICE-SG show a similar distribution of overall frequency of *already*. Furthermore, we observe that there are register and lectal variations in CSE, as frequencies of *already* and sentence-final *already* differ in different types of texts and among speakers with different levels of education. In ICE-SG, the less formal text type represented by the private dialogues favors sentence-final *already*, while the public dialogues and monologues – both scripted and unscripted – prefer sentence-medial *already*. On the other hand, as observed in OHI, the higher frequency of sentence-final *already* seems to be associated with a lower education level.

8.1.2 Substrate-influenced *already* vs. phasal polarity *already*

This section focuses on the distribution of *already* with substrate-influenced features compared to phasal polarity *already*. As discussed earlier in Chapter 5.1, *already* in CSE has extended its grammatical function from a phasal polarity expression (PhP) to an aspectual marker expressing (i) the completive (i.e. an event/action has been completed), (ii) the inchoative (i.e. an event/a state has just started), and (iii) the prospective aspect (i.e. an event/state is about to start). On the other hand, PhP *already* involves two contrasting temporal reference points, which typically result in the connotation of anteriority, i.e. an event/state happened or began earlier than expected. In contrast, aspectual *already* does not highlight two contrastive temporal points, nor does it necessarily involve the interpretation of anteriority and unexpectedness. Besides, aspectual *already* can appear in a negative environment, whereas PhP *already* commonly occurs in positive contexts. Apart from the grammatical functions of CSE *already* as an aspectual marker, the study categorized the

following features as substrate-influenced features: (i) lack of inflectional/perfect markers, (ii) double use of *already* in one single utterance, (iii) predicate ellipsis in the sentence *already* occurs in, (iv) word order change triggered by *already*, (v) *already* co-occurring with *still* and (vi) tokens of *already* with discourse functions (see Chapter 7.3.1). Tokens with these above features were annotated as substrate-influenced use, while examples without the above aspectual readings and substrate-influenced features were categorized as PhP *already*. The remaining tokens were counted as unclear examples.

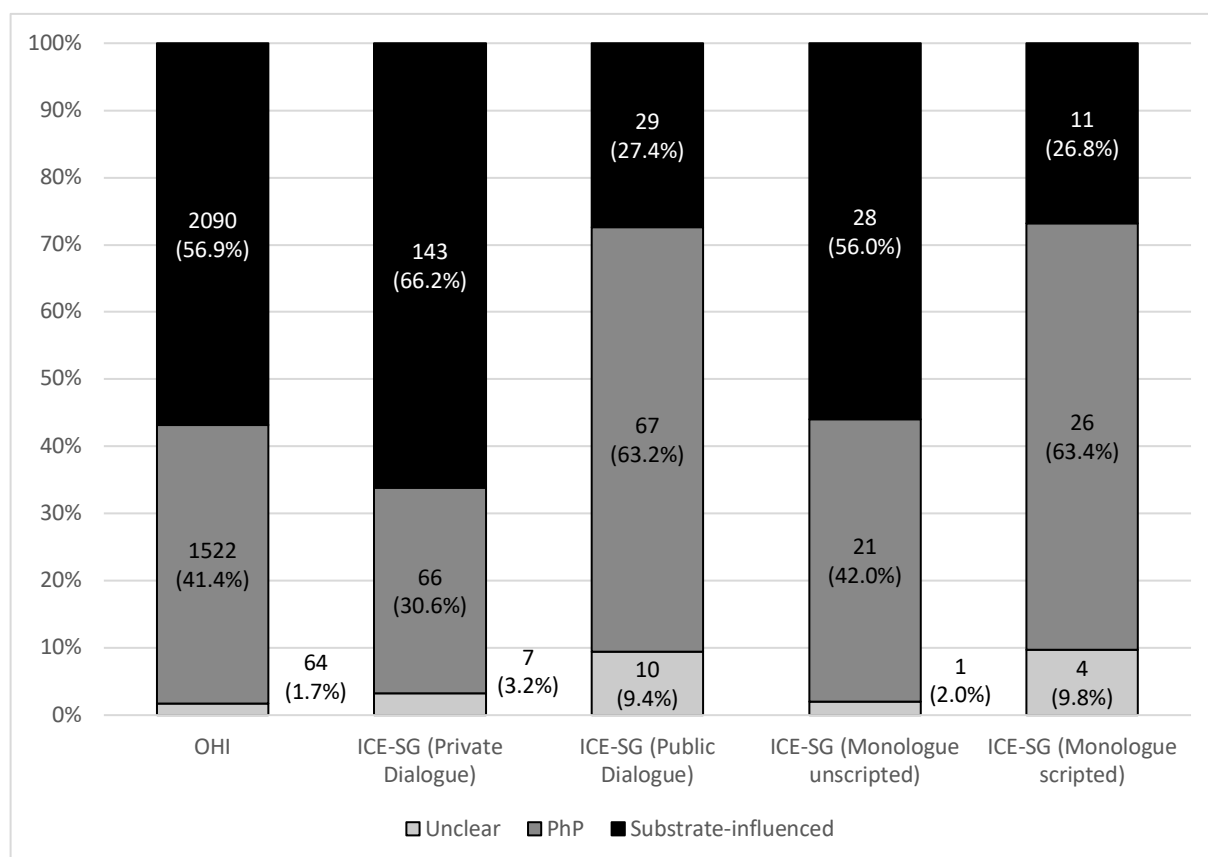


Figure 8.5: The absolute frequencies and proportion of substrate-influenced *already* vs. PhP *already* in OHI and ICE-SG

As revealed in Figure 8.5, more than half of the tokens of *already* in OHI (56.9%) exhibit substrate-influenced features, while 41.4% of the *already* occurrences belong to the category of PhP expression. On the other hand, ICE-SG demonstrates a similar distribution of *already* between the substrate-influenced (53.9%) and PhP expression (46.0%) categories. Apart from that, there is a preference for *already* with substrate-influenced features occurring in

the private dialogues (66.2%) and unscripted monologues (56.0%) in ICE-SG. In comparison, most PhP *already* occurrences were found in public dialogues (63.2%) and scripted monologues (63.4%). In sum, speakers tend to use *already* with substrate-influenced features in informal and spontaneous speeches while they prefer PhP *already* in formal and prepared speeches.

In the more qualitative part of this study, I analyzed the occurrences of *already* as aspectual markers as well as their tokens with substrate-influenced features in OHI (2,090 tokens) and ICE-SG (211 tokens). As Figure 8.6 and Figure 8.7 show, most of the *already* tokens function as an aspectual marker, either as a completive or as an inchoative marker. As a token of *already* can manifest multiple substrate-influenced features, the sum of these features is larger than the total sum of the *already* tokens in the substrate-influenced category. For example, inchoative *already* and *already* in negative sentences as shown in (162), or completive and lack of tense-aspect markers as in (163). In a majority of these aspectual cases, the predicate that *already* modifies remains in its bare, unmarked form (803 cases in OHI and 141 cases in ICE-SG). As shown in (163), it seems that aspectual *already* triggers the deletion of inflectional markers (the equivalent to English *-ed*), and it is used as an inflectional marker itself instead. It is worth mentioning that no examples show substrate-influenced meanings of *already* among the British speakers in OHI.

(162) Now, I think I don't want *already*. [OHI-001632-CHN]

‘Now, I think I begin to want nothing.’

(163) a. That time I have children *already*. [OHI-001632-CHN]

‘That time I have already had children.’

b. Oh yes, because you have *already* contact. You always asked questions.

[OHI-002198-TIT]

‘Oh yes. Since you got their contact information, you always asked them questions.’

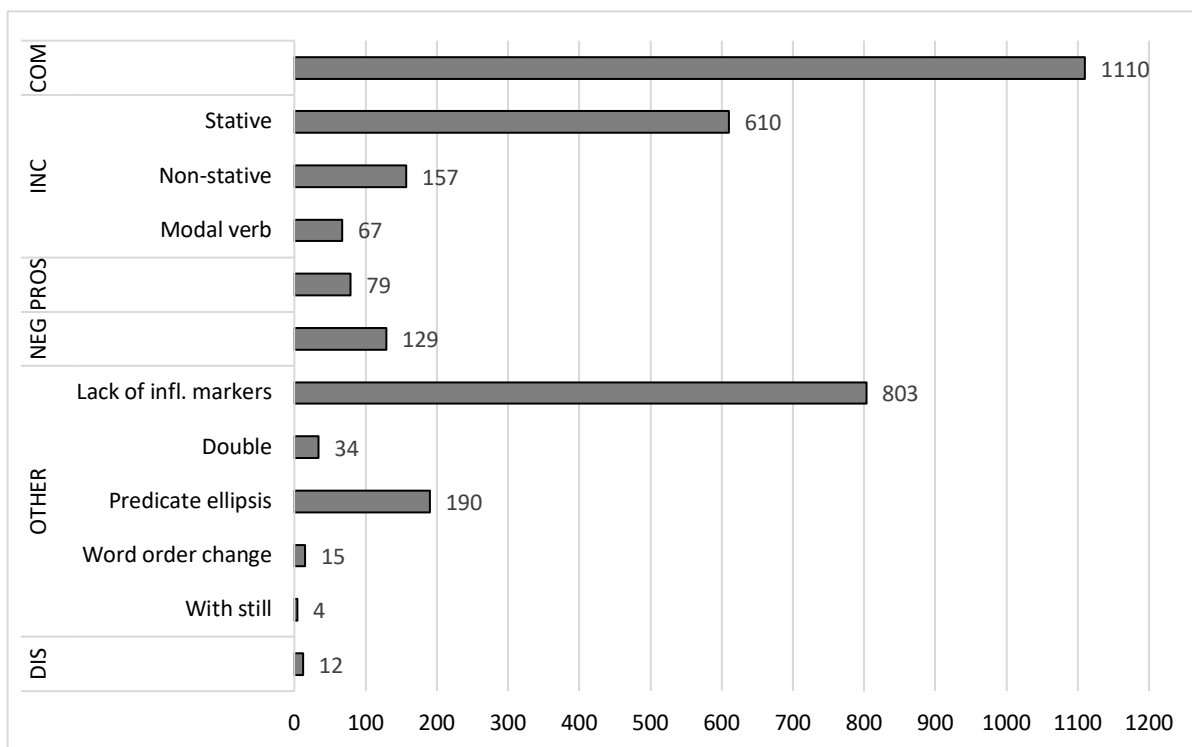


Figure 8.6: Categories of substrate-influenced uses of *already* in OHI (absolute figures)⁶⁴

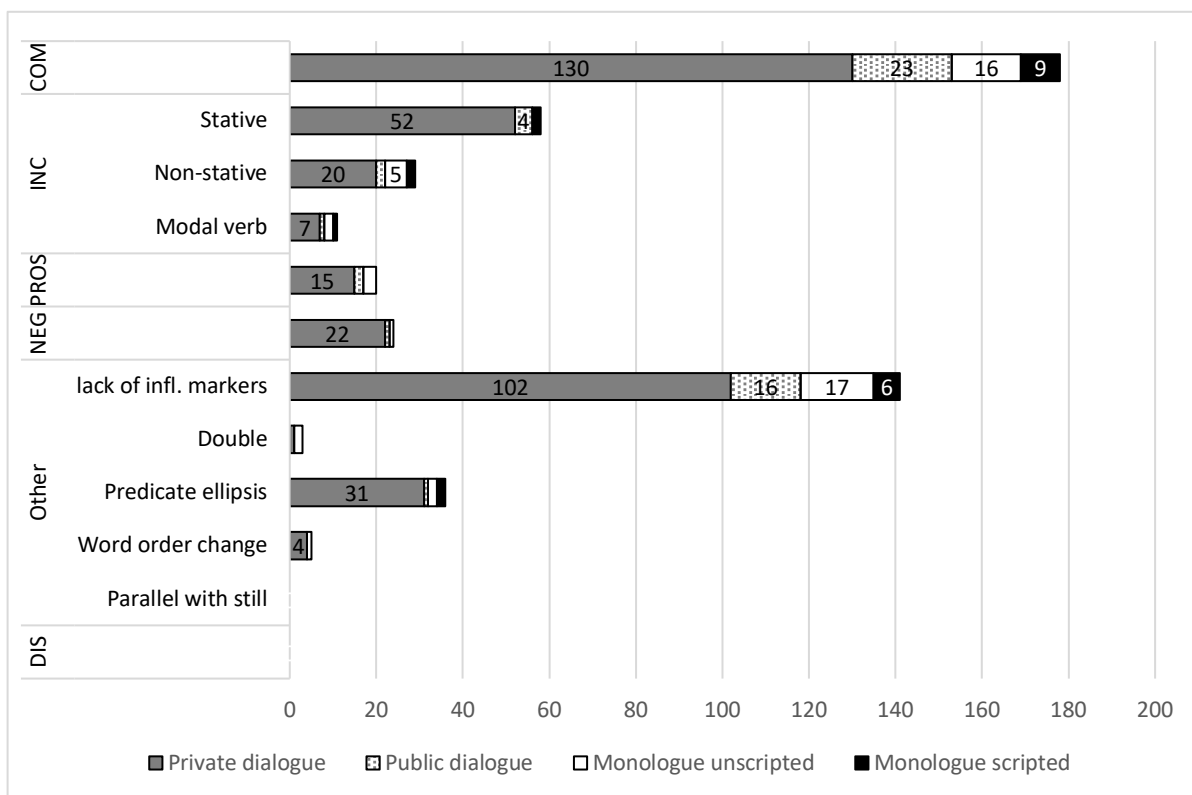


Figure 8.7: Categories of substrate-influenced uses of *already* in ICE-SG (absolute figures)

⁶⁴ COM=complete, INC=inchoative, PROS=prospective, NEG=in negative sentences, DIS=discourse marker.

If we compare Figure 8.6 with Figure 8.7, we can see that the distribution patterns of *already* with various functions do not differ significantly between OHI and ICE-SG. First, completive *already* accounts for 30.2% (n=1110) of all the *already* occurrences in OHI and 36.4% (n=178) in ICE-SG. This is followed by inchoative *already*, which comprises 22.7 % (n=834) of the *already* occurrences with substrate-influenced features in OHI, and 20.2 % (n=98) in the same category found in ICE-SG. The prospective aspect is not well-represented in OHI (2.1%, n=79) or in ICE-SG (4.1%, n=20). Additionally, 21.8% (n=803) of the cases of *already* in OHI occur with an uninflected verb. The proportion of the same phenomenon in ICE-SG is even higher, which accounts for 28.8% (n=141) of all cases of *already*. In both corpora, I found instances of *already* in negative contexts, with the proportions of their occurrences being moderately higher in ICE-SG (4.9%, n=24) than OHI (3.5%, n=129).

Register and formality are accountable for the competition between substrate-influenced *already* and PhP *already*. Most of the substrate-influenced features were found in the private dialogues in ICE-SG, as represented by the columns marked in grey in Figure 8.7. However, cases of *already* with substrate-influenced features were used infrequently in the public dialogues and monologues. The following section analyzes the examples of each of these substrate-influenced categories of *already*.

8.1.3 Aspectual *already*

Completive already

As introduced in Chapter 5.1, completive *already* marks the completion or termination of an event or a state, which is expressed in the preterit or the perfect in Standard English. In total, there are 1,110 occurrences in OHI, which account for 30.2% of all the *already* tokens in the dataset. In comparison, there are 178 occurrences of completive *already* in ICE-SG (130 in private dialogues, 23 in public dialogues, 16 in unscripted monologues, and 9 in scripted monologues), which make up 36.4% of all the *already* tokens in ICE-SG. Interestingly, the verb with which *already* occurs is mostly not inflected. Among all the tokens of *already* found in OHI, 21.8% of the occurrences manifest a lack of inflectional tense or aspectual

markers. In ICE-SG, the ratio of *already* occurring with a bare, non-inflected verb is generally higher (28.8%). The results suggest that the completive use of *already* co-occurring with a bare, non-inflected verb is one of the most salient features of CSE *already*.

In the completive use of *already*, the associative presupposition of the hearer's prior knowledge of the event and a potential implicature of "earlier than expected" are missing. Consider the following examples:

- (164) completive *already*
- a. 1937 and then after that, name *already* change. [OHI-E000284-CNS]
 - b. [...] JTC has built and two of them are *already* painted. He was very happy. [OHI-003223-CSS]
 - c. So obviously there were smell? Ah, now, it's cleared *already*. Since that time they have cleared a lot. [OHI-001953-LAS]
 - d. That time, once the fire start up *already*, all people said, "Wow, big fire." [OHI-001953-LAS]

For example, the time adverbial *1937* suggests that either the simple past or the present perfect should be used in (164)a, but the actual verb *change* remains in its bare, uninflected form, i.e. the auxiliary verb *have* and the inflectional suffix *-ed* on the actual verb are missing. The coordination sentence in (164)b provides further evidence that *already* is reanalyzed as a completive marker. In Standard English, clauses that are linked by the coordinating conjunction *and* are usually in the same tense and aspect. However, in the first half of the coordination sentence in (164)b, the speaker used the English perfect form (*build* – *has built*), but used the copula *be* in its present form in the presence of *already*. Moreover, the verb in the sentence *He was very happy* following the previous coordination sentence is inflected. It proves that the speaker was aware of the past tense context. It shows that the speaker used *already* as a completive marker and regarded *already* as an equivalent to the English preterit or the English perfect. The same applies to the other examples in (164), in which *already* is used with a non-inflectional verb. According to Eckhart (2011), changes in morphosyntax may cause semantic reanalysis. It is possible that the non-inflection of the main verb triggers a semantic reanalysis of CSE *already* from a PhP expression to an aspectual marker (also see the discussion on grammaticalization of *already* in Chapter 9.2.2).

Besides functioning as the equivalent to the English perfect form, the core semantic function of completive *already*, however, is to signal the completion or termination of a state or an event, as exemplified in (165):

- (165) a. They will use, I remember we have this booklet. Once you tear *already*, it disappear then you draw, its ‘cleanable’ they always play this [...] [OHI-002951-JY]
- b. Interviewer: What about the workers, they just stand by?
Interviewee: They don’t do anything. They stopped *already*. [OHI-001953-LAS]
- c. I got it and 1974 I was appointed a lecturer. '73 before I finished *already*, they gave me a senior tutorship. [OHI-002204-TWH]
- d. Ah, until I finish *already*, then went off. [OHI-003206-MA]
- e. A lot of people finish *already*. [ICE-SG-S1A-069#12:1:A]

In (165)a, the speaker described the use of an erasable drawing pad she used in her childhood. One can *tear* ‘clear off’ the sketches with a touch of the cleaning button. Here the speaker used *already* to denote that once the action of “tearing” had been completed, the previous sketches would disappear. In addition, *already* seems to denote that such an action is irreversible, as once the sketches were erased, they could not reappear. The context in (165)b is a discussion on people’s reaction after a fire accident. Without *already* the sentence *They stopped* is a mere description of the bystanders’ reaction to the accident. It may elicit further questions of the interviewer: *And what happened after that?* or *Why did they stop?* However, with the use of *already*, example (165)b denotes the finality of the story. The completive marker *already* implies that the previous utterance was all that the speaker wanted to contribute at that moment. When *already* occurs in the event of accomplishment, as exemplified by the verb *finish* in (165)c, (165)d, and (165)e, *already* emphasizes a situation where an action has been carried out to totality.

In (164), we observed that past tense/past participle affixation was deleted, even though the context made it clear that a past time reference was intended. Siemund (2013:117) also stated that the use of *already* occurs with untensed verb forms. I found cases in the OHI

data where *already* occurs with an uninflected verb, as exemplified in (165)d and (165)e (see also previous examples in (164)). However, our data also provide evidence that past tense forms are used alongside *already*, as shown in (165)b and (165)c. The co-occurrences of past tense inflection and completive *already* suggest a competition of the aspectual marking systems between the superstrate and the substrate languages, a phenomenon also observed in a Portuguese-based creole – the Korlai Creole Portuguese (KCP) (Clements 2006). In KCP, past tense can be marked redundantly with *ja* ‘already’ and verbal suffix *-or* or *-w* (Clements 2006:87). Similar phenomenon is found in Cape Flats English, Hong Kong English, Ghanaian Pidgin, and Uganda English, where past tense forms are equally abundant alongside ‘already’ (Li and Siemund 2021:531–534).

Apart from that, this study found *already* often co-occurring with verbs such as *finish(ed)* (n=12), *die(d)* (or *pass(ed) away*) (n=44), and *retire(d)* (n=22) in OHI. The result is in line with Ziegeler (2020), which observed that *already* was more frequently used with *finished* than with high frequency verbs such as *bought* and *went*. Ziegeler (2020:324) suggests there is a “semantic harmony” between verbs expressing lexical completion and the use of completive *already*. It is assumed that such a semantic harmony is influenced by verbal *le* in Chinese (see Chapter 6.2.2), which often co-occurs with verbs expressing lexical completion and presents a realization of a completive situation (see Xiao and McEnery 2004 on the actual aspect of Chinese *le*). Apart from co-occurring with verbs expressing completion, the study also found examples in the CCL corpus in which *le* co-occurs with adverbials meaning ‘all’ or ‘completely’, as exemplified in (166). In such cases, *le* tends to underline the completion of an action that has finally occurred.

- (166) 他们 的 新政 完全 打消 了
 tāmen de xīnzhèng wánquán dǎxiāo le
 they POSS new deal completely eliminate ASP
 ‘Their new deal has been completely eliminated.’
 [CCL-Contemporary-CWAC-AHB0018]

Inchoative already

As mentioned earlier in Chapter 5.1.3, inchoative *already* refers to the usage of *already* as a marker to indicate the beginning of an action (Bao 1995:183–84). Among all the instances of inchoative *already* (n=834), 74.3% (n=610) of them occur with a stative predicate as in (167) while 18.8% (n=157) of the cases occur with a non-stative predicate as in (168), and 8.03% (n=67) with a modal verb, as in (169). In ICE-SG, there are 98 occurrences of inchoative *already* (79 in private dialogues, 7 in public dialogues, 7 in unscripted monologues, and 5 in scripted monologues). Like OHI, the majority of the instances of inchoative *already* (59.2%, n=58) occur with stative predicates, followed by non-stative predicates (29.6%, n=29), and inchoative *already* co-occurring with modal verbs (11.2%, n=11).

(167) with a stative predicate

a. MH: We go and take the pineapple skin, the second layer, we eat with that.

Interviewer: That was during the war years? After the war?

MH: After the war, it was quite comfortable *already*. [OHI-002206-MH]

b. Sixties? No, that one is fifty something *already*. [OHI-001953-LAS]

(168) with a non-stative predicate

Saturday morning we go back *already*. That means Saturday night we didn't come back. [OHI-001953-LAS]

(169) with a modal verb

I was *already* Standard I. I could go on my own *already*. [OHI-001631-OCN]

As discussed in Chapter 5.1.3, although both English *already* and substrate-influenced *already* in CSE give rise to the inchoative reading when it occurs with a stative predicate, English *already* as a PhP marker is typically enriched by the connotation of anteriority, i.e. earlier than expected. Yet, CSE *already* as an inchoative marker signals the start of a new situation. Again, the context is crucial in determining this subtle semantic difference, as illustrated by example (167)a.

The context of (167)a is that during the war, the condition of life was extremely harsh, which was reflected in the scenario that people had to eat things like pineapple skin due to

lack of food supply. The interviewer proceeds to ask about life condition after the war. Speaker MH underlines that the situation has changed after the war by using inchoative *already*, i.e. that the post-war life started to return to normal. *Already* here was categorized as an aspectual marker of the inchoative, which signals the beginning of a new phase, i.e. post-war life begins to return to normal. We could also interpret a change of state, i.e. from ‘life being uncomfortable during the war’ to ‘life being comfortable after the war’. In either of the above readings, *already* involves no reading of “earlier than expected”.

Prospective already

It has been reported in the literature that *already* can serve as a prospective aspectual marker, which denotes an action that will happen shortly in the immediate future (see Chapter 5.1.1). Compared with the number of completive and inchoative *already*, the frequencies of prospective *already* found in OHI and ICE-SG are relatively small. Altogether there are 79 and 20 occurrences of prospective *already* found in OHI and ICE-SG respectively. Among the 20 occurrences of prospective *already* found in ICE-SG, 15 tokens occur in the private dialogues, 2 in public dialogues, and 3 in unscripted monologues. What is worth mentioning is that there are no occurrences of prospective *already* found in the scripted monologues. Again, this result shows that marking the prospective aspect is a feature that prefers informal and spontaneous speech. Two examples are listed below:

- (170) a. What? So I said, okay you report to police, I go home *already*. Then reached back here, the sister scolded me, “You must go in [...]”
[OHI-002206-MH]
- b. I am gonna bring all the jackets and wear *already* wear my jackets.
[ICE-SG:S1A-065#35:1:B]

Taken out of the context, both examples in (170) can have a completive reading: ‘I have left for home’ and ‘I have worn my jackets’. However, as indicated by the context, (170)a depicts a scenario where the speaker is going to leave for home, and in (170)b, the speaker is going

to wear the jackets in the immediate future. Again, the prospective meaning is parallel to the same aspectual function encompassed in Chinese *S-le* (see Chapter 5.1.1).

As we have seen so far, the attestations of the completive, inchoative, and prospective meanings of *already*, both in OHI and ICE-SG, provide us with evidence that *already* has grammaticalized from a phasal polarity expression to an aspectual marker. Furthermore, the parallel constructions between Chinese *le* and CSE *already* confirm that Chinese *le* is a possible source of cross-linguistic influence on CSE *already*.

8.1.4 Extended reading of aspectual *already*

Aspectual *already* may extend to imply notions such as irreversibility of an event and fundamentality of a change of state. Consider the examples in (171):

- (171) a. This is all changed *already*. Last time was not like that. It was like the Bugis Street area. Then later the government stepped in and then asked them to move. [OHI-002951-JY]
b. The Japanese money was of no use *already* after the surrender. [OHI-000462-CSC]
c. She telephoned me. She said, “Do you know that the ship has sunk *already*? Do you know that the ship has sunk?” [OHI-000123-NB]

Example (171)a is about the relocation of a famous street-food stall of the speaker’s grandmother. As requested by the government lead by Lim Yew Hock in the 1950s,⁶⁵ all the hawkers had to relocate their stall from Bugis Street to Malabar Street. Among these street stalls was the famous street-food stall held by her grandmother, which sold traditional Cantonese chicken porridge. According to the narratives of her grandmother, the stall used up to 60 chickens a day for the porridge during the wartime. The speaker fell into a train of melancholy and nostalgic thoughts about the childhood memory of her grandmother. Here,

⁶⁵ Lim Yew Hock (林有福 Lin Yǒufú, 15 Oct 1914 – 30 Nov 1984) was a Singaporean and Malaysian politician of Chinese descent, who served as a Member of the Legislative Council and Assembly from 1948 to 1963, and the second Chief Minister of Singapore from 1956 to 1959 (see OHI-000024-FSC).

already, which co-occurs with the adverbial *all*, as in *This is all changed already* denotes an irreversible change out of state. In addition, the use of *already* in such irreversible and fundamental changes would likely engender disappointment and regret, as the outcomes were beyond the speaker's control.

The reading of disappointment, regret, and nonvolitionality is also apparent in (171)b. The speaker depicted a scenario where the Japanese currency became useless towards the end of World War II. In the story, he only came to know the Japanese Surrender two or three weeks after it happened. Unfortunately, he had sold a grown pig for Japanese money in the market before he was informed about the Japanese surrender. Here, the aspectual marker *already* expresses the “inability to undo”.

In (171)c, the speaker informed the interlocutor about the sinking of the ship. Here, *already* is used in a sentence syntactically compatible with Standard English. From the context, however, it is clear that the hearer has no previous knowledge about the sinking of the ship. In other words, there is a lack of the connotation of anteriority of the event. Instead, *already* denotes the completion of the sinking of the ship, which is irreversible and beyond the speaker's control.

In sum, the study results show that the majority of aspectual *already* found in OHI and ICE-SG function as a completive marker. It is often used in a non-inflectional context equivalent to the English preterit or English perfect. Its core semantic function is to denote finality and totality of an event/action. In addition to events of accomplishment that signify the speakers' determination to achieve some desirable outcome, completive *already* found in OHI underlines certain types of completive events/actions or a change of state as irreversible and fundamental. Depending on the context, the sense of finality and totality may extend to the reading of disappointment, regret, and relief.

8.1.5 Aspectual *already* vs. the English perfect

Perfect and perfective are very different, albeit overlapping grams. As exemplified by the English perfect, the label “perfect” is typically used in relation to some specific

morphosyntactic forms (Schwenter and Cacoullos 2008:3). Semantically, perfects are relational, signaling a past situation related to discourse at speech time and is therefore currently relevant. In contrast, perfectives report an event “for its own sake” (Bybee, Perkins, and Pagliuca 1994:54), and indicate that a situation is viewed as bounded (Bybee and Dahl 1989:55). The change from perfect to perfective use involves a generalization of meaning with loss of current relevance (Bybee et al. 1994:86–87).

Four main uses are commonly distinguished in the literature on English perfect: (i) perfect of result (also resultative perfect), (ii) existential perfect, (iii) universal perfect (also perfect of persistent situation or continuative perfect), and (iv) perfect of recent past (Comrie 1976; Dahl and Wälchli 2016; McCawley 1971; Siemund 2013; Kroeger 2018). This is illustrated by the most salient interpretations of the sentences in (172), adapted from Kroeger (2018:428):

- (172) a. Perfect of result: I have lost my keys, so I can't open the door.
- b. Existential perfect: Have you *ever* seen the movie?
- c. Universal perfect: I have lived in Hamburg since 2011.
- d. Perfect of recent past: He has just arrived.

Intriguingly, there is a close semantic affinity between the English perfect and CSE *already*. Except for the existential perfect as shown in (172)b, all other perfects find equivalent expressions with CSE *already* in OHI. Some of the examples in each category are listed below:

- (173) Perfect of result
Interviewer: Did you still keep in touch with those who dropped out?
OCN: No. Because after the war all of us, lost each other *already*.
[OHI-001631-OCN]
- (174) Universal perfect
 - a. They, from young they learn *already*. [OHI-001109-MN]
 - b. After 14, we stay in Changi Road *already*. [OHI-002951-JY]
- (175) Perfect of recent past
 - a. Don't know what did we talk. I forgot *already*. [OHI-001631-OCN]

- b. Now you got a child *already*. It cannot be like one. Go and have your own home, stand on your own feet. [OHI-001663-MS]

In (173), the speaker Oh Choo Neo (OCN, female, born in Singapore in 1924) was a teacher in a Methodist Girls' School. She was recounting that when she was still a student, many of her female classmates suspended school due to lack of financial support.⁶⁶ The interviewer asked whether she still kept contact with those girls who suspended school. Oh Choo replied that they lost contact with each other after the war. Here, *already* corresponds to the “perfect of result” as the present state of “not being in touch” is the result of a past situation – the war (see Comrie 1976:56). In Standard English, the perfect of result is also used to imply the result of the event or action has an influence on the present situation (see “perfect of continuing result” in McCoard 1978). For example, in (172)a by using the present perfect, the sentence implies that losing the keys is still relevant at present, as one cannot use them to open the door. Likewise, in (173) *already* was used to signal that “we lost touch with each other” is relevant to the utterance time. Curiously though, Chen (2008) and Ziegeler and Lee (2019) claim that *already* may not be used in contexts expressing accidental or unintended events in CSE (**I broke my leg already*). However, this lexical constraint does not seem to apply to the earlier data of CSE.

The existential perfect in English is often expressed in a negative context (see Section 5.3), as illustrated by the example (172)b. The existential perfect refers to “a given situation [that] has held at least once during some time in the past leading up to the present” (Comrie 1976:58). *Already* in CSE does not express the existential perfect. The reason for the lack of *already* in the existential reading may be that CSE *ever* is preferred when expressing the experiential perfect, which is close to but not equal to the existential perfect (see Chapter 5.3.2). For example, we can potentially interpret *already* in (176) as ‘I have seen it once in

⁶⁶ Women were more likely to be excluded from receiving a comparable education to men in history, as the expense of rearing a girl was considered a loss to her family because she would eventually belong to her husband's family (Liu and Carpenter 2005:227).

the past'. However, it is more likely to function as a completive marker, which indicates that seeing an event that happened in Bugis Street occurred completely within topic time.⁶⁷

- (176) This is like last time Bugis Street. I see *already*. I am very scared.
[OHI-002951-JY]

With regard to universal perfect or perfect of persistent situation, it refers to “a situation [...] started in the past but continues (or persists) into the present” (Comrie 1976:60). It is important to note that the universal reading is only possible with atelic situations, as shown in (177). While (177)a and (177)b are compatible with a universal perfect reading, telic situations as in (177)c are not compatible with it (Kroeger 2018:436).

- (177) a. I have lived in Hamburg since 2011.
b. I have been working for three months.
c. *I have finished my dissertation for three months.

CSE *already* in (174) corresponds to the universal perfect reading. First, the verbs *learn* and *stay* belong to the atelic verbs, where the event has no natural endpoint. Moreover, the temporal adverbials *from young* and *after 14* are atelic-extent adverbials, which do not entail that the situation ended at a particular time (see Haspelmath 1997:120–26 on temporal extent of time adverbials). In contrast, a temporal adverbial that is presented as having an endpoint is a “telic-extent adverbial”, e.g. *The boy wrote the essay in five hours* (Haspelmath 1997:130). Example (174)a expresses that *learning* started with the Singaporeans at an early age and that continues on into the present, while (174)b indicates that the speakers started living in Changi Road at the age of 14 and that the state extends into the present. In both cases, *already* functions as a marker of the universal perfect.

The semantics of CSE *already* in (175), listed again in (178), is identical to the perfect of recent past. According to Comrie (1976), perfect of recent past refers to “the present relevance of the situation is simply one of temporal closeness, i.e. the past situation is very

⁶⁷ Topic time refers to “the time span to which the speaker’s claim on this occasion is confined” (Klein 1994:4). In other words, topic time is the time span about the topic that we are talking about.

recent” (1976:60–61). For example, *He has just arrived* indicates that the action *arriving* occurs shortly before the moment of utterance.

- (178) a. Don’t know what did we talk. I forgot *already*. [OHI-001631-OCN]
 b. Now you got a child *already*. It cannot be like one. Go and have your own home, stand on your own feet. [OHI-001663-MS]

In (178)a, the interviewee has just forgotten the topics of their previous discussion, as suggested by the preceding sentence *Don’t know what did we talk*. Similarly, *already* in (178)b conveys that the situation of *getting a child* is recent, which is suggested by the temporal adverbial *now*. The situation is relevant due to its proximity in time to the present. Clearly, *already* has acquired the grammatical function of the perfect of recent past.

One of the most well-known theories which attempts to capture all the usages of the English perfect is McCoard’s (1978) “current relevance” and “extended now” theory. According to McCoard (1978), the present perfect in English is a grammatical form used to describe a past event with present relevance (current relevance) and a present state resulting from a past situation (extended now). The use of aspectual *already* can be accounted for by both readings to varying degrees. For example, *after the war all of us, lost each other already* in (173) implies that the result of war is relevant to the present and that we don’t have each other’s contact at the utterance time. In (174), *We stay in Changqi Road already* conveys that this is a part of the experience of the interviewee at the age of 14 and that the state of *staying in Changqi Road* extends to the present. In (178)a, *I forgot already* implies that this memory loss is troublesome at the present moment as the interviewee cannot tell the interviewer the story of the past and that the state of forgetting extends to the time of the utterance. In short, both readings of the English perfect are applicable to CSE *already* as an aspectual marker.

8.1.6 *Already* in negative sentences

As previously discussed in Chapter 5.1.1, Standard English *already* is a typical example of “positively oriented polarity-sensitive items” (PPIs) which characteristically occur in positive clauses (Huddleston and Pullum 2002:710). The semantics of *already* involves the change from a negative state to a positive one (van der Auwera 1993:619). However, CSE *already* is able to appear in negative sentences which can be attributed to the influence of Chinese *S-le* with the negator 不 *bù* or 没有/没 *méiyǒu/méi* (see Chapter 6.2.4).

Of the 3,676 tokens of *already* in OHI, 129 tokens (3.51%) are used in a negative sentence. For the speakers in ICE-SG, the use of *already* in a negative sentence constitutes 4.91% of the total tokens of *already*. The examples in (179) illustrate the usages of *already* in negative sentences found in OHI.

- (179) a. I was not young *already*, *already* reaching the age of 27, 28.
 [OHI-000071-CKM]
 ‘I was no longer young/not young anymore at that time. I had reached the age of 27 or 28.’
- b. So I said, I would call it a day. At my age, I was not practicing that many hours *already*. [OHI-000838-SMK]
 ‘[...] At my age, I was no longer practicing that many hours.’

The construction of *not/no...already* in CSE is basically equivalent to English *no longer/no more/not anymore*. In Standard English, *already* is used to signal a change of state from a negative state to a positive state, *no longer* marks a transition from a positive state to a negative state. According to Löbner (1989:172), there is an “internal negation” between the semantic concept of English *already* and *no longer*. As illustrated by Figure 8.8, “internal negation” means that an element has the negation in its scope, e.g. (X NOT): *already* (not do/undergo x) = *no longer*. In contrast, there is an external negation between English *already* and *not yet*: NOT (X): not (*already* do/undergo x) = *not yet*. As pointed out by Kramer (2018:1), *no longer* implies the non-occurrence of a state at the reference time while such a state holds at a prior phase. However, the meaning of ‘no longer’ can be expressed with the

construction of *already* with the negator *no/not* in CSE. For example, in (179)a, *I was not young already* means that ‘I was no longer young’, and in (179)b *I was not practicing that many hours already* translates into ‘I was no longer practicing that many hours’ in Standard English.

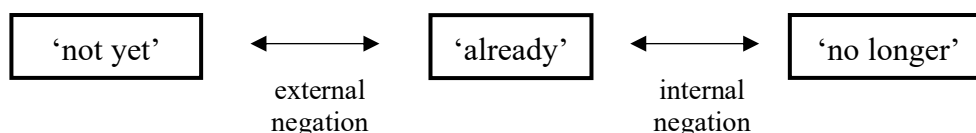


Figure 8.8: The system of semantic relations between *not yet*, *already* and *no longer* (adapted from Löbner 1989:172)

The observation here is that the construction of *not...already* in negative sentences found in OHI mirrors that of 不...了 *bù...le* in Chinese. As exemplified by example (180) found in CCL, Chinese allows the realization of ‘no longer’ based on internally negating *already* on the surface structure. Similar with example (179)a, example (180) marks the transition from a positive state of “being young” in a phase prior to the reference time to a negative state “not being young” at reference time.

- (180) 我 也 不 年 轻 了, 他 说。
 wǒ yě bù niánqīng le tā shuō
 I also *not* young *already* he say
 ‘I am no longer young either, he said.’
 (CCL Contemporary/Magazine/Reader)

However, the construction of *not...already* (179)a does not entirely copy the use of 不...了 *bù...le* as shown in (180): the copula *be* is present in (179)a but absent in (180). As mentioned earlier, the copula verb *be* (or a linking verb) is normally absent in a sentence with an adjective predicate in Chinese. It is assumed that the speaker added the copula verb *be* here to conform to English’s surface structure.

A further piece of evidence for the Chinese influence on the use of *already* in a negative sentence is its frequency difference between CSE speakers of a Chinese ethnic background and speakers of other ethnic backgrounds. Of the 129 tokens of *already* found

in a negative sentence, 118 tokens were used by the Chinese group in OHI, while the speakers of other ethnicities produced only 11 tokens. Of these 11 tokens, 5 tokens occur in the group of Peranakan speakers, and 5 tokens in the Malay group, the last one was produced by the speaker of an Iraqi ethnic background. As mentioned earlier in Chapter 5.1, Malay also has a construction of *sudah/dah* ‘already’. Combined with the negator *tidak*, the construction expresses the sense of ‘no longer’, but it is rarely used and limited to a change of state from a positive to a negative one. In conclusion, the ratio of *already* in a negative sentence can also serve as a good indicator of Chinese influence on CSE.

8.1.7 Double *already*

Interestingly, both datasets reveal examples of CSE *already* occurring twice in the same clause, as exemplified by (181):

- (181) a. After all this news going on, after the screening, the people were *already* living in fear *already*. [OHI-000265-LTS]
b. By the time they talked about globalization all of Singapore’s economy was *already* oriented towards a global marketplace *already*. [OHI-002715-CCB]
c. Uh you *already* have PC *already* right. [ICE-SG:S2A-055#148:2:B]
d. And then later on when I had my son, then we *already* wanted to leave for Endau *already*. [OHI-000462-CSC]

The appearance of *already* in two positions – one in preverbal/postverbal position and one in clause-final position – seems to violate the surface structural requirements of the lexifiers, which challenges Bao’s (2005) systemic transfer and lexifier filter hypothesis. However, it should be noted that the ratio of double *already* is relatively small. In OHI, there are 34 cases of double *already*, which add up to only 0.90% of all *already* occurrences. They are also present in the more recent CSE dataset – ICE-SG, though on an even smaller scale: 3 cases were found in ICE-SG, which only form 0.61% in the dataset. Besides, they only appear in less formal registers such as the private dialogues (1 token) and the unscripted monologues

(2 tokens) but do not appear in formal and prepared speech such as in the public dialogues and scripted monologues.

The construction of double *already* in (181) is not redundant. A close semantic analysis reveals that preverbal/postverbal *already* and clause-final *already* in a single clause serve different aspectual functions. For example, while the sentence-final *already* in (181)a gives rise to the inchoative reading, i.e. people started to live in fear after hearing the news. On the other hand, the preverbal/postverbal *already* indicates that the state of people living in fear started prior than expected.

The double use of *already* reveals a clear need to replicate the morphosyntactic conventions of the substrate languages. Chinese *le* may be found both in post-verbal and clause-final positions, serving similar functions as *already* in the above examples (see Chapter 6.2.1). However, the clause-medial *already* follows the auxiliary as in standard varieties of English and is not post-verbal as Chinese *V-le*. It is possible that the position of the antecedent *already* has gone through some modification to adhere to the morphosyntactic requirements of the superstrate language – English.

It is also possible that PhP *already* provides sources for the antecedent *already* while Chinese *le* provides the aspectual meaning for the clause-final *already*. In other words, PhP *already* and aspectual *already* can co-exist in a single clause. As shown in (181)c, the antecedent *already* is semantically and syntactically compatible with Standard English *already* (or PhP *already*), and sentence-final *already* can be interpreted as a completive marker, signaling that the action of acquiring a computer happened in the past.

A further possible source of the double *already* construction is the Chinese construction of 已经...了 (*yǐjīng...le*). Not only is *yǐjīng* semantically compatible with PhP *already*, but it is also syntactically compatible with *already* in Standard English. As shown in (182), the pre-predicate position of *yǐjīng* is identical to the antecedent *already* in the construction of double *already*. Consider the following examples from the CCL corpus:

- (182) a. 我发现我已经会唱这首歌了
 wǒ fāxiàn wǒ yǐjīng huì chàng zhè shǒu gē le
 I find I *already* can sing this CL song ASP
 ‘I found that I *already* started to know how to sing this song.’ (inchoative)
 (CCL-Contemporary-History-Biography-Chinese artists drifting north)
- b. 他已经想要出去了
 tā yǐjīng xiǎngyào chūqu le
 he *already* want go out ASP
 ‘He *already* wants to go out.’ (inchoative with the volitional verb *want*)
 (CCL-Contemporary-translation work-Crime and Punishment)

In either of the above explanations, we can see that Chinese *le* provides cross-linguistic influences on the double *already* construction. A further piece of evidence is that the use of double *already* is most prevalent amongst Chinese speakers of CSE: 32 out of 34 cases of double *already* construction in OHI occur amongst Chinese speakers while only 2 occurrences were found amongst speakers with a Malay background.

8.1.8 Other substrate-influenced morphosyntactic features of CSE *already*

Copula be ellipsis

In total, 5.2% (n=190) of the occurrences of *already* in OHI were found in a context without a copula. In ICE-SG, the ratio of *already* occurring in a context of zero copula is relatively higher (7.4%, n=36). It seems that besides functioning as an aspectual marker, *already* can assume the function of the copula *be*, as exemplified in (183):

- (183) a. The Government’s requisition *already* one million over.
 [OHI-000057-LYC]
- b. Where got the energy? Even our mind, all blank *already*.
 [OHI-001953-LAS]

Already with natural development predicate collocates

Interestingly, CSE *already* often collocates with adjectives such as *married*, *old* (also *not young*), and *dead*, as exemplified in (184).

- (184) a. [...] two sisters [were] married *already*. This is very interesting [...] [OHI-003036-MS]
 b. A: Have you made plans for the future?
 B: Old *already*, what plans for the future? [OHI-001880-LKN]
 c. I was not young *already*, *already* reaching the age of 27, 28. [OHI-000071-CKM]
 d. [...] him including my husband and Mr Wee Soo Bee, he's dead *already* [...] [OHI-000462-CSC]

As mentioned earlier, it is common for CSE speakers to omit a copular verb. In (184)a, the combination of *already* with the adjective *married* seems to assume the function of the predicate. The same applies to (184)b. In addition, *already* was found in predicates such as *grow up*, *get old*, *retire*, *die*, and *pass away*. These types of predicate resemble what Dahl and Wälchli (2016:328) refer to as “natural development predicate[s].” A natural development predicate will typically reach a certain phase, or an endpoint sooner or later under normal circumstances.

Again, Chinese *le* can account for the reconstruction of aspectual *already* with a natural development predicate. Like *already*, Chinese *le* can co-occur with one of the natural development predicates, such as 结婚 *jiéhūn* ‘married’, 老 *lǎo* ‘old’, and 死 *sǐ* ‘dead’ or ‘to die’ as shown in (185):

- (185) a. 他们 现在 全 结婚 了
 tāmen xiànzài quán jiéhūn le
 they now all marry ASP
 ‘They are all married now.’
 [CCL-Contemporary-Spoken-1982-PekingDialectSurvey]

- b. 这 十 几 年 来, 大 家 都 老 了
zhè shíjǐ nián lái dàjiā dōu lǎo le
these ten-about year come everyone all old ASP
'In the past ten years, everyone has grown old.'
[CCL-Contemporary-Spoken-Dialogue-LiAo]
- c. 他 病 大 重, 以 后 不 久 就 死 了
tā bìng dà zhòng yǐhòu bùjiǔ jiù sǐ le
he ill big heavy thereafter shortly then die ASP
'He was severely ill, and died shortly thereafter.'

Apart from Chinese *le*, Malay *sudah* may also be reconstructed for aspectual *already* with a natural development predicate. Example (186) is given by Olsson (2013:18), which suggests that the use of *sudah* with such predicates becomes systematic or even obligatory.

- (186) Kamu tidak bisa memakan-nya. itu sudah busuk. (Olsson 2013:18)
You not can eat-it that ASP rotten
'You cannot eat it. It is rotten.'

It is worth mentioning that the predicate 'be married' may not be considered so much as a natural development predicate as 'old' and 'dead'. However, marriage is still regarded in some cultures as a status that will be reached sooner or later under normal circumstances. A typical example is *sudah kawin* 'married' in Indonesian, which is printed on some Indonesian identity cards to indicate the citizen's civil status (Dahl and Wälchli 2016:328). Similar collocations such as *married already*, *marry already*, and *already married* occur 31 times in OHI. It seems that such a cultural notion was held in Singapore as well. It should be noted, however, Malay *sudah* occurs in preverbal position but rarely in sentence-final position (Teo 2019:358). Yet, 23 out of 31 tokens of these *already* collocations are cases where *already* occurs sentence-finally, which suggests a stronger Chinese influence. Apart from that, 22 out of 31 tokens were produced by Chinese speakers. Again, this result reveals a clear need to adhere to the semantic function of the Chinese substrates.

V-already Object

Curiously, *already* was found to occur between a verb and its object in OHI – [V-*already* Object]. Such a construction does not conform to the morphosyntactic requirements of English, which seems to minimize the role of the lexifier filter alluded to in Bao's (2005, 2015) studies (see Chapter 4.7.4). Consider the examples in (187):

- (187) a. In '66 I started in May, and by Nov '66, I engaged *already* two people first. Because our starting was little bit difficult. [OHI-000123-NB]
 b. [...] and visiting steel mills in other countries? Oh yes, because you have *already* contact. You always asked questions. [OHI-002198-TIT]
 c. I was a marine biologist never mind if outsiders thought I had lost *already* that art. And to do so I wanted to just publish one. [OHI-002204-TWH]

There are 15 occurrences of similar examples as above found in OHI, which only contribute to 0.4% of all *already* occurrences in OHI. The construction of V-*already* Object manifests a word-order shift of *already* from its pre-verbal position to post-verbal as in Chinese V-*le* (see Chapter 6.2.1).

As previously discussed in Chapter 6.2.1, V-*le* (or suffix *-le*) is placed directly after a verb or adjective. Its functions include receiving the English preterit interpretation and/or marking the completive/actual aspect. Parallel to the construction V-*already* Object in CSE, V-*le* can be followed by an object, as exemplified in (188):

- (188) a. 你们 实际上 鼓励 了 我
 nǐmen shíjìshàng gǔlì le wǒ
 you(pl.) actually encourage ASP me
 'You have actually encouraged me.'
 [CCL-Contemporary-Magazine-People'sDaily-1995-December]
- b. 我 和 几位 学者 有 了 联系
 wǒ hé jǐ wèi xuézhě yǒu le liánxì
 I and several scholars have ASP contact
 'I got in touch with several scholars.'
 [CCL-Contemporary-Application-SocialScience-RenYuzhong]

- c. 没有 比试, 就 丢失了 体育 之 魂
méiyǒu bǐshì jiù diūshī le tǐyù zhī hún
no competition then lose ASP sport POSS spirit
'Without competition, the spirit of sports is lost.'
[CCL-Contemporary-Magazine-1994 selections-10]

A comparison of the examples in (188) with the CSE examples in (187) shows a high similarity indicating a strong morphosyntactic influence from the Chinese substrates on the *V-already* Object construction.

The finding of the *V-already* Object construction violates the constraint of lexifier filter (see Chapter 4.7.4), as Standard English *already* is either pre-verbal or clause-final. However, the examples in (187) suggest that CSE *already* can be post-verbal. Bao (2015:59) claims that the novel exponents of the transferred system (here the aspectual system) must meet the grammatical requirements of the lexical-source language, at the expense of system transfer if necessary. In other words, the lexifier filter constraint was assumed to rank higher than the system transfer. However, in the case of the contact of pre-verbal *already* and post-verbal *le* where the lexifier and the substrate are in conflict, *already* in (187) adheres to both the morphosyntax and semantic functions of the substrate, at the cost of violating the lexifier morphosyntax.

However, such force of violating the lexifier filter constraint as in (187) appears to have a waning effect on the more recent CSE. With regard to ICE-SG, the construction *V-already* Object only appears two times. Consider both examples in (189):

- (189) a. So we represent *already* five departments and several subjects.
[ICE-SG:S2A-024#10:1:A]
b. I am gonna bring all the jackets and wear *already* wear my jackets.
[ICE-SG:S1A-065#34:1:B]

As shown in (189)a, *already* seems to build a stronger connection with the following object than with the preceding verb: the use of *already* here seems to modify the objects, emphasizing that *five departments and several subjects* are a remarkable number of properties to present. In other words, besides marking the completive aspect, post-verbal

already also functions to underline the number or the amount of the object/objects as significant and remarkable. Example (189)b suggests that the speaker was aware that using *already* post-verbally violates the morphosyntactic conventions of the lexifier language. Although the speaker initially used *already* post-verbally, he quickly corrected himself by discarding the use of *already*.

The co-occurrence of already and still

In addition to the above *already* constructions, a further piece of evidence of the grammaticalization of *already* is the co-occurrence of *already* with the adverb *still* in the same clause. Consider the following example:

- (190) a. But at that time *already* things were expensive *still*. [OHI-000001-TMK]
 b. They, from young they learn *already* but I'm *already still* small you know.
 [OHI-000265-LTS]

The above examples are interesting because the co-occurrence of *already* and *still* seems semantically contradictory. In standard varieties of English, *still* is also a phasal polarity expression, which signals that a state (e.g. things being expensive) is the case at reference time implying a subsequent phase where this state is not the case (e.g. things not being expensive) (see van der Auwera 1993). In other words, *still* notes that a certain event or state has not reached its endpoint, but we presuppose it had ended earlier. On the other hand, *already* signals a state is the case at reference time, but not the case at a prior state (see Chapter 5.1.2). Therefore, *already* and *still* only co-occur in a transitional complex sentence, either linked by *or* or *but*, as exemplified by the following examples from the Corpus of Historical American English (COHA):

- (191) a. Stallings, *already* wounded but *still* in the fight, took a machine-gun burst in his right knee, destroying the joint [...]
 [COHA-1993-MAG-AmHeritage]
 b. Was I *still* alive, or *already* dead? [COHA-1957-FIC-WapshotChronicle]

However, the use of *already* with *still* found in OHI does not resemble that in the above examples. Unlike Standard English *already*, CSE *already* can co-occur in a single simple clause. Apparently, it is not likely that both *already* and *still* function as phasal polarity expressions in a single simple clause as in (190). Therefore, it appears that CSE *already* functions as a preterit marker here, which denotes that the state/event happened or started in the past.

Potential pragmatic particle

It is found that *already* is used in imperative sentences in OHI. As shown in (192), the speaker tried to persuade the interlocutor to start a new business with him. Here, *already* denotes the verb *do* as a command, and entices solidarity at the same time. This finding suggests that *already* can serve as a pragmatic particle, as such usage is identical to one of the functions carried by the pragmatic particle *lah* in CSE (see Chapter 1.2 and Chapter 4.3.2). Like *already* in (192), the pragmatic particle *lah* attached to an imperative as in (193) expresses a commitment that an interlocutor is expected to act upon (Gupta 1992b:42).

(192) I got extra money. Then I called them. Do *already*, we profit sharing.
[OHI-002951-JY]

(193) You just go and do *lah* and go and ask any question. [OHI-003223-CSS]

It is worth noting that *already* could also be analyzed in this context as a prospective marker, in the sense of ‘about to start’. According to this interpretation, the sentence with *already* is interpreted as ‘let’s start to do it now’. However, the prospective aspect is usually interpreted in the declarative context (Bao 2005:241). The imperative context in (192) is more likely to give rise to the pragmatic particle interpretation.

8.1.9 *Already* vs. social variables

This section first displays the frequencies of *already* in OHI in relation to each social factor, such as ethnicity, educational level, and gender, before demonstrating the results of the

multiple linear regression, which considers all the social factors. The main reason for applying the multiple linear regression model is that certain social factors may cancel the effects of others (e.g. socio-economic factor or educational level may override the effect of gender) (Holmes and Wilson 2017:181). The results show that ethnicity and educational level have a strong effect on the frequency of *already*.

Frequency of already vs. ethnicity

The previous section analyzed how the Chinese substrate *le* asserts strong morphosyntactic and semantic influences on CSE *already*: CSE *already* is used as an aspectual marker in sentence-final position, which is identical to the Chinese counterpart *le*. If we assume that the speakers of different ethnic backgrounds in OHI speak a different ethnic language, then the factor of ethnicity is a particularly relevant indicator of cross-linguistic influence. In sociolinguistics, certain linguistic features often identify ethnicity, though variation will exist in both the ethnic community and in the mainstream community, they differ with regard to the frequency of variants (see Tagliamonte 2012). Therefore, this study predicts that the highest frequency of *already* with substrate-influenced features occurs in the Chinese group. In other words, it is expected that the ratios of *already* in sentence-final position and *already* as an aspectual marker serve as a good indicator of Chinese influence on CSE.

As predicted, the Chinese group produces the highest frequency of *already* on average. At the same time, it has the highest frequency of *already* as an aspectual marker and the highest frequency of *already* in sentence-final position (see Figure 8.9). The results confirm that CSE *already* is more likely to be influenced by the Chinese substrates than other ethnic languages. The vast range of variation amongst the Chinese speakers in terms of the relative frequency of *already* in all categories (i.e. overall frequency, *already* in sentence-final position, and *already* as an aspectual marker) suggests that there are substantial individual differences in this group.

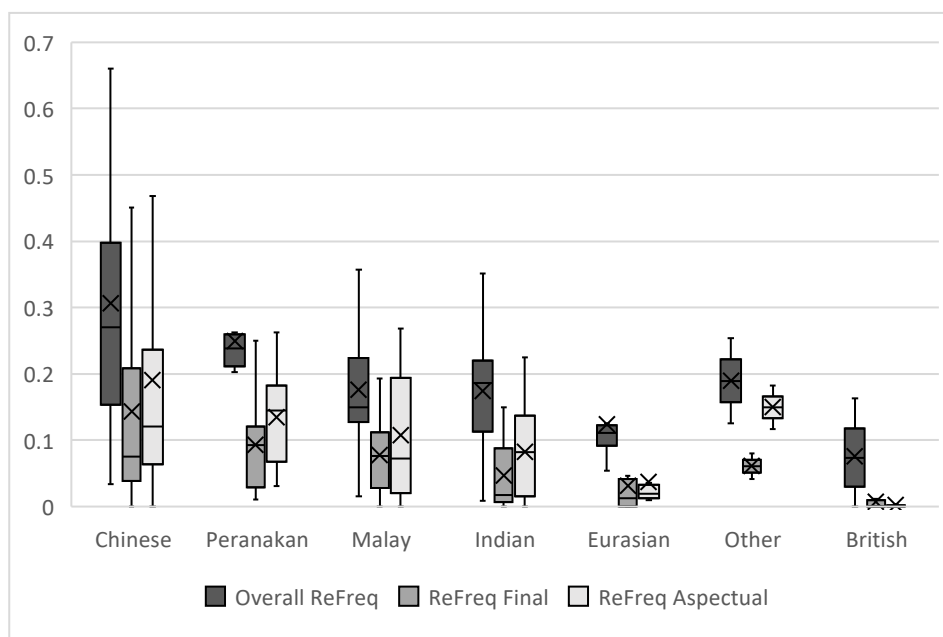


Figure 8.9: Frequency of *already* (pp) in different ethnic groups⁶⁸

On the other hand, the Peranakan, the Malay, the Indian, and the Other groups do not differ significantly from each other, as they produce similar ratios of *already* in all three categories. Not surprisingly, *already* with substrate-influenced features are hardly detected in the British and Eurasian groups. Both groups manifest a marginal frequency of *already* as an aspectual marker and sentence-final *already*. Apart from that, the British and the Eurasian groups produce the lowest overall frequency of *already*. The results are in line with the result of comparison between ICE-SG and ICE-GB: the frequency of *already* in the spoken sections in ICE-SG is significantly higher than in ICE-GB. It should be emphasized that while the comparison between ICE-SG and ICE-GB examines cross-variety differences (Singaporean variety and British variety), the focus here is on the internal differences regarding the frequency of *already* amongst Singaporean speakers with different ethnic backgrounds.

The fact that the Malay, the Indian, the Peranakan groups, and the group of “Other” show a homogeneous usage pattern of *already* suggests an indirect influence from CSE. Their frequency distribution of *already* falls between the Chinese and British/Eurasian

⁶⁸ The line inside the box represents the median, the cross corresponds to the mean. 50% of the data are within the box. The upper end of the box (also known as the upper quartile or 75th percentile) corresponds to the median of the upper half of the dataset while the lower end of the box (or 25th percentile) represents the median of the lower half of the dataset (see Levshina 2015:58).

communities. The result suggests that aspectual and sentence-final *already* has become a shared code embraced by different ethnic communities, not limited to the Chinese community.

Frequency of already vs. educational level

According to Platt's (1975) continuum hypothesis, speakers at the higher end of the social continuum have a wider range of available lexis, while those at the lower end have a more restricted range of linguistic choices (see Chapter 4.7.1). As sentence-final position and marking the perfective aspect belong to non-standard features of *already*, we expect that these *already* tokens occur more frequently in the speech of the interviewees with a lower educational level.

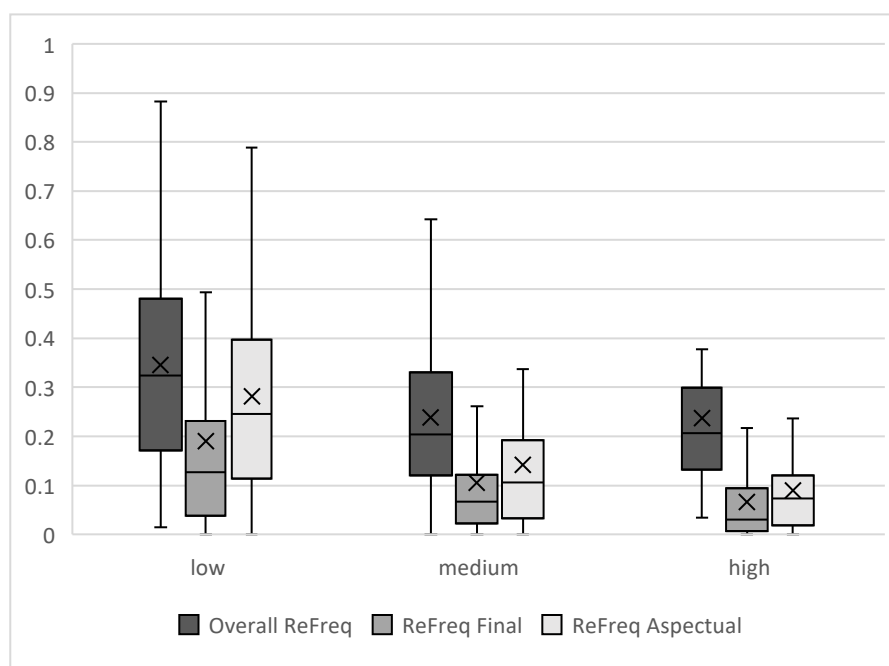


Figure 8.10: Frequencies of overall, sentence-final, and aspectual *already* pp according to educational level

As shown in Figure 8.10, sentence-final *already* and *already* as an aspectual marker occur indeed more frequently amongst the interviewees with a lower educational level. In other words, the frequencies of sentence-final *already* and aspectual *already* distinguish between

speakers socially. This result confirms the assumption that the frequency of *already* is indicative of individual lects.

To a large extent, the result agrees with Platt's (1975) continuum hypothesis, in which the frequency of substrate-influenced *already* is subject to social stratification. However, Platt (1975) also claims that speakers with a lower educational level have a narrow range of available lects. Yet, as shown in Figure 8.10, it is the speakers with an insufficient academic level (primary school education) who have a broader range of variation in the use of substrate-influenced *already*. In such cases, it may be more adequate to correlate language use with socio-economic background. For example, speaker LGS (accession number 000009), a politician with a medium level of education has a lower frequency of *already* (0.26 pp) than speaker LY, a teacher and later an inspector of an institute with a university education (0.46 pp).

Frequency of already vs. gender

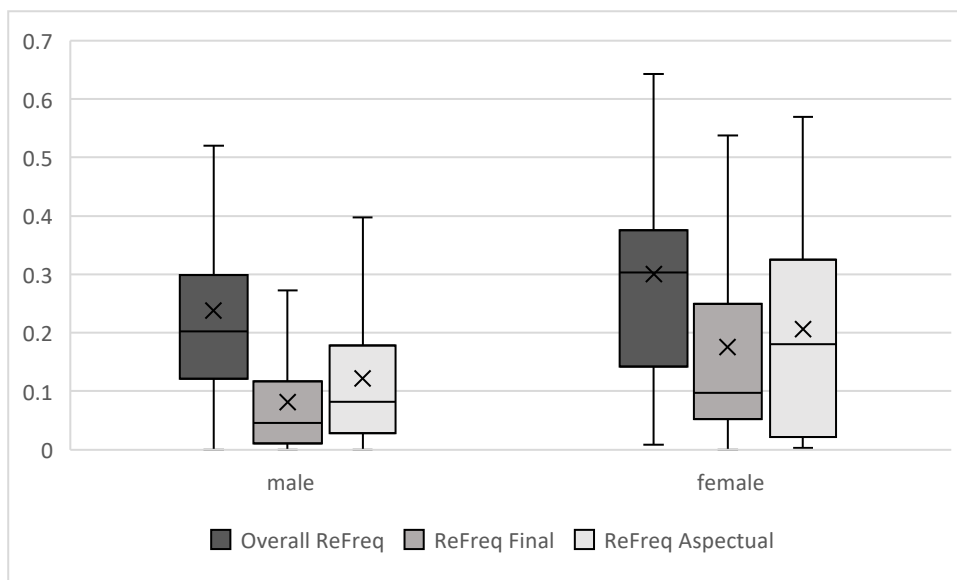


Figure 8.11: Frequencies of overall, sentence-final, and aspectual *already* pp vs. gender

As suggested by Figure 8.11, female speakers in OHI tend to use *already* more frequently than male speakers. Besides, higher frequencies of sentence-final *already* and aspectual *already* are found amongst female speakers than male speakers. However, the difference is

not very significant, as female speakers lead male speakers by only 0.1 token of *already* per page on average. It should be noted here that there is an unbalanced number of female speakers versus male speakers (1:4) in the OHI data.

Labov (2001) identifies that women adopt innovative variants at a higher rate than men in linguistic change from bottom (Labov 2001:266). Change from below denotes linguistic change in a speech community below the level of a speaker's conscious awareness. In this case, speakers are not consciously aware of a linguistic change in progress in a community (Meyerhoff 2006). The use of substrate-influenced *already* can be reckoned as a type of linguistic change from bottom as it can be seen as a covert marker of CSE (see Chapter 4.3). In this regard, the differences between women and men agree with Labov's (2001) hypothesis. However, it should be emphasized that gender generally interacts with other social factors, such as socio-economic status, ethnicity, and age (Holmes and Wilson 2017:181).⁶⁹ Therefore, this study draws on multiple linear regressions to include the joint impact of the social variables available in OHI. The multiple linear regression analyses will be shown after the next section on the frequency of *already* across age groups.

Frequency of already across age groups

Figure 8.12 visualizes the relative frequency of *already* across different age groups. Three age groups were distinguished, namely, speakers born in 1900–1919, 1920–1939, and after 1940.⁷⁰ The arrangement of the birth year of the speakers in ascending order exhibits the implication of the apparent-time finding: the study of the language used by speakers of different age groups is tantamount to studying language change (Labov 1994:197).

⁶⁹ Apart from socio-economic status, ethnicity, and age, other important social factors in sociolinguistic studies include the role of the speaker in an interaction and the formality of the context (Holmes and Wilson 2017:181).

⁷⁰ One speaker born in 1899 was placed in the group 1900–1919.

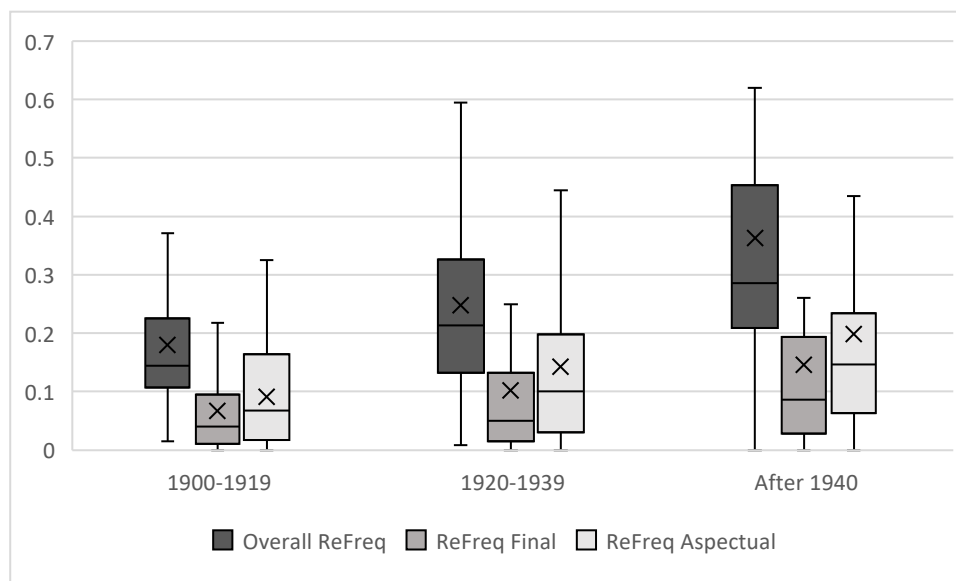


Figure 8.12: Frequencies of *already* overall, sentence-final, and aspectual *already* pp vs. age group

The result suggests a small but consistent increase in the use of *already* over the whole period. The same trend can be observed in the use of *already* in sentence-final position and aspectual *already*, as the frequencies of both categories produced by speakers born after 1940 are twice as high as those produced by speakers born between 1900 and 1919. Besides a general increase of *already*, we can also observe that each generation becomes more heterogeneous, as the gap between the highest and lowest frequencies of *already* grows more extensive than the previous period.

Multiple linear regression analysis

In order to calculate the joint impact of the factors discussed in the previous section on the frequency of *already*, this study draws on a multiple linear regression analysis. It is used when we want to predict the value of a variable based on the value of two or more other variables (Levshina 2015:139). For example, a multiple linear regression analysis is used to understand whether exam performance can be predicted based on revision time, test anxiety, lecture attendance and gender (Field 2017:528). For the multiple linear regression analyses, the frequency of aspectual *already* and the frequency of sentence-final *already* were defined

as the dependent variable while the predictors are ethnicity, education, age group, and gender. The analyses were conducted and computed in SPSS Statistics (IMB Corp 2020).

	Estimate	Std. Error	t	Sig.
(Intercept)	0.352	0.059	5.975	0.00000004
Ethnicity	-0.033	0.007	-4.982	0.00000278
Education	-0.106	0.019	-5.536	0.00000027
Age group	0.053	0.019	2.809	0.00602802

R-Square: 0.375

F-statistic: 19.177 on 3 and 96. p-value: 7.982E-10

Table 8.1: Multiple linear model on relative frequency of aspectual *already*

As shown in Table 8.1, a significant regression equation was found between the predictors (i.e. ethnicity, educational level and age group) and the dependent variable ($F(3, 96)=19.177$, $p<0.0001$) with an R^2 of 0.375. Four outliers were excluded from the model. Except for gender ($p=0.373$), all the other three variables added statistically significantly to the prediction. Apart from that, ethnicity and education level are better predictors of the frequency of aspectual *already* than age group.

	Estimate	Std. Error	t	Sig.
(Intercept)	0.275	0.054	5.116	0.00000162
Ethnicity	-0.023	0.006	-3.956	0.00014686
Education	-0.064	0.018	-3.607	0.00049675
Age group	0.036	0.017	2.135	0.03536651
Gender	-0.058	0.028	-2.082	0.04000318

R-Square: 0.309

F-statistic: 10.620 on 4 and 95. p-value: 3.7234E-7

Table 8.2: Multiple linear model on relative frequency of sentence-final *already*

Table 8.2 demonstrates the influence of the social factors including ethnicity, education, age group, and gender on the frequency of sentence-final *already*. Again, ethnicity and education

level are the better predictors of the frequency of sentence-final *already* than age group and gender. The results confirm our hypothesis that the frequency of sentence-final *already* indexes a speaker’s ethnicity and educational level. The next section continues with the analysis of substrate-influenced *also*.

8.2 Analysis of *also*

8.2.1 Frequencies of sentence-final *also*

Compared with the absolute frequency of *already*, the number of the occurrences of *also* in OHI is much higher – 11,293 tokens of *also* were found in OHI. The number of *also* is more than three times as high as the frequency of *already* (3,676) found in OHI. Among these tokens of *also*, 22.4% occur in sentence-final position. On the other hand, sentence-final *also* only accounts for 12.0% of all the occurrences of *also* in the spoken sections in ICE-SG (see Figure 8.13).

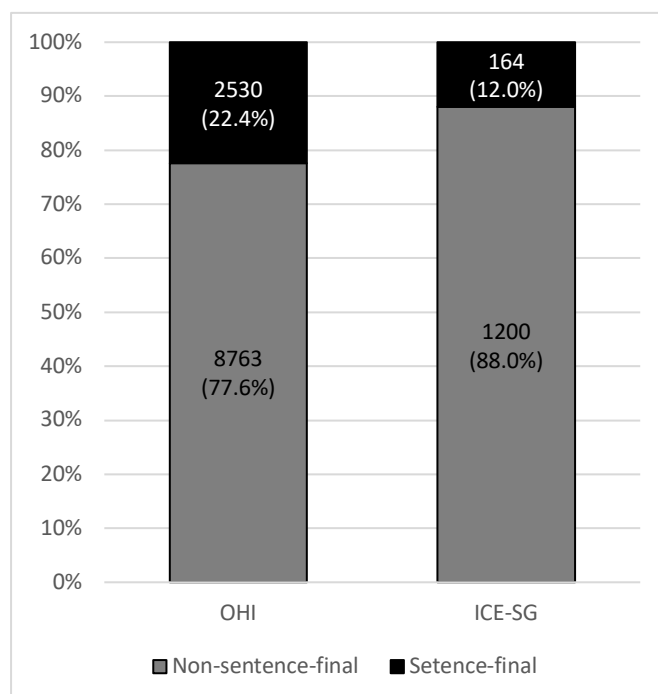


Figure 8.13: The proportion of *also* in non-sentence-final and sentence-final position in OHI and ICE-SG

The result in Figure 8.13 seems to suggest a decrease in the frequency of sentence-final *also* (chi-square=78.2855, $p < 0.00001$). However, as shown in Figure 8.14, the frequency of *also*

is register-sensitive: the private dialogues manifest the highest frequency of sentence-final *also* (37.81%) while more formal speeches represented by the public dialogues, unscripted monologues, and scripted monologues record only 4.25%, 4.03%, and 0.56% of sentence-final *also*, respectively. According to Bao and Hong (2006:110), sentence-final *also* does not occur in the private dialogues in ICE-GB, and the ratio of sentence-final *also* is 0.06 per thousand words in the public dialogues in ICE-GB (see Chapter 5.2.2). Hence, CSE – as represented by both OHI and ICE-SG – is marked by a much higher frequency of *also* in sentence-final position than British English.

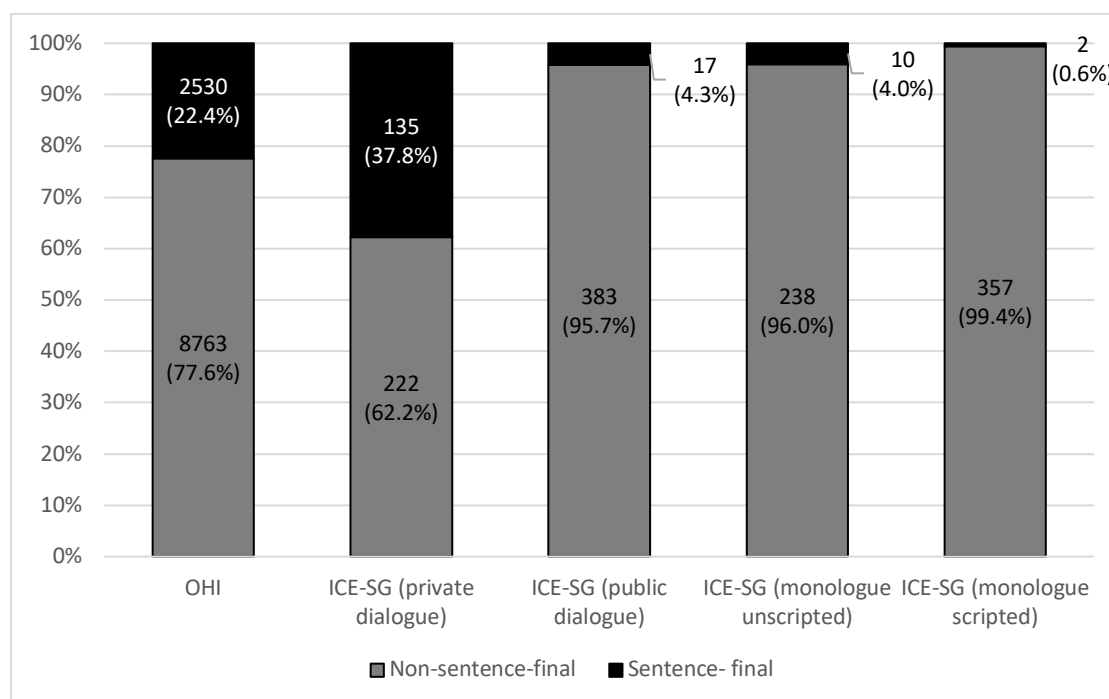


Figure 8.14: The proportion of *also* in non-sentence-final and sentence-final position in OHI and ICE-SG (according to different text types)

As indicated by Figure 8.14, register is an essential factor influencing the frequency distribution of sentence-final *also*. The primary genre in the OHI is narrative, as the interviewees were retelling past events which took place in their lives, which is perhaps closer to the speech of unscripted monologues in ICE-SG. Judging from the above figures, it is not possible to detect a shift away from sentence-final *also* to the more standard usage

of *also* in sentence-medial position. On the contrary, sentence-final *also* seems to be a stable feature of CSE.

8.2.2 Substrate-influenced *also*

Among the 2530 cases of sentence-final *also* found in OHI, 500 occurrences (19.8%) were interpreted as having acquired novel meanings from the Chinese substrates, in addition to the additive function meaning ‘too’ or ‘as well’ (see Chapter 5.2.1). As illustrated in Figure 8.15 and Figure 8.16, *also* most frequently co-occurs with universal quantifiers (i.e., *every** (*everything*, *everyone*, and *everybody*), *all*, *always*, and *any** (*anything*, *anyone*, and *anybody*), functioning as a maximality operator (see Chapter 6.4.2 on the substrate *dōu* with universal quantifiers). Apart from that, 158 cases of *also* with novel grammatical meanings occur in negative sentences. There are 65 instances of *also* co-occurring with the concessive adverb *even*. Similar to *already*, *also* was sometimes found to occur twice in the same clause.

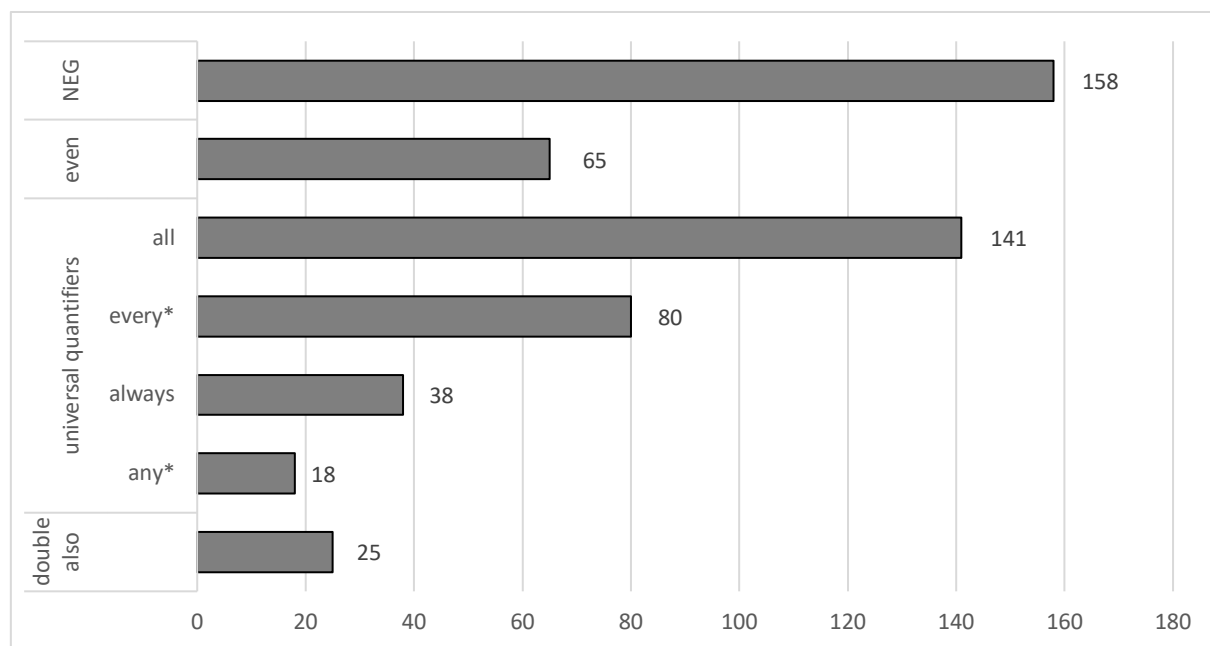


Figure 8.15: Frequencies of sentence-final *also* with novel grammatical meanings in OHI

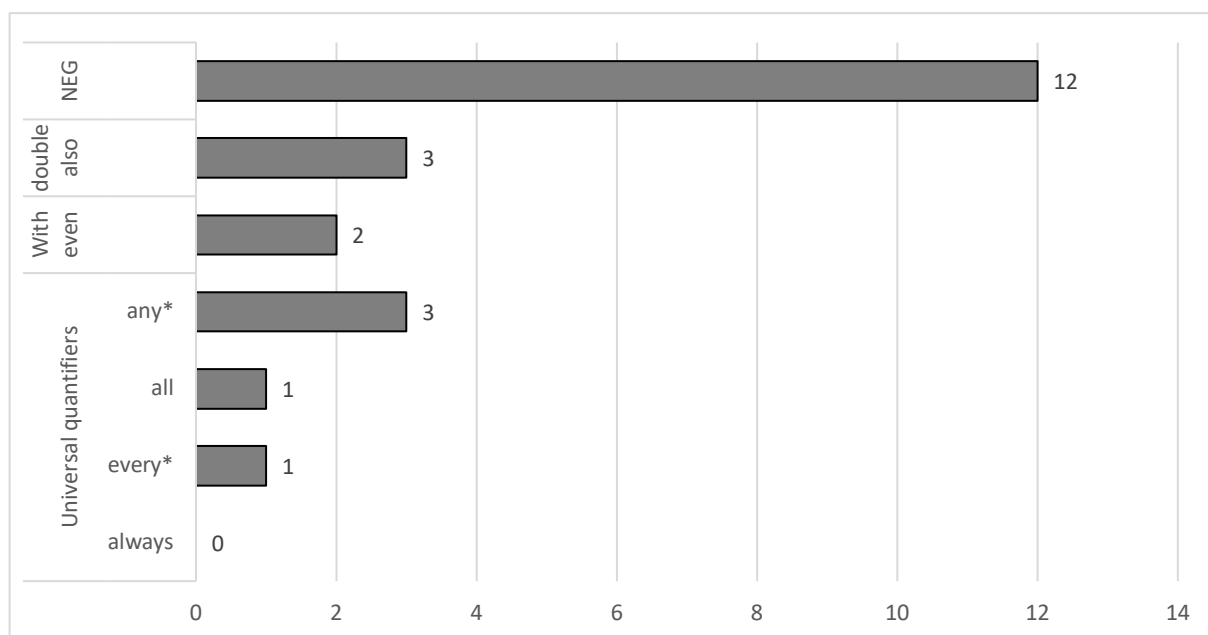


Figure 8.16: Frequencies of sentence-final *also* with novel grammatical meanings in ICE-SG

Similar to the dataset in OHI, all the substrate-influenced categories of *also* mentioned above were found in ICE-SG, though in a slightly smaller amount: 22 occurrences out of 164 sentence-final tokens of *also* (13.41%) were found to have acquired Chinese-influenced grammatical meanings. Here we focus on the sentence-final tokens of *also*. Apart from that, there are cases found in ICE-SG, where *also* has acquired these substrate-influenced meanings, but occurs in sentence-medial position, as exemplified in (194). The influence of Bao's (2005, 2015) lexifier filter can be an explanation for the lower frequency of sentence-final *also* relative to non-sentence-final use in ICE-SG, as the morphosyntax of Standard English may have a constraint on *also* in adopting the sentence-final feature while acquiring the substrate-influenced meanings.

- (194) a. *All of them also have to pass the same bar exam.*
 [ICE-SG:S2B-024#89:2:B]
- b. *Even Mandarin is also not standard.* [ICE-SG:S1A-083#12:1:B]
- c. *Even officers also they wear plain shorts also right.*
 [ICE-SG:S1A-065#1:1:A]
- d. *Are you telling me the private clinic are also subsidising us also and for the ward charges let me tell you.* [ICE-SG:S2A-022#34:1:A]

The results suggest that, first of all, sentence-final *also* with Chinese-influenced grammatical meanings is a stable feature in CSE. It had been used at least from the early 20th century (OHI) to the late 1990s (ICE-SG). It confirms our hypothesis that at least some functions of *also* were undergoing contact-induced grammaticalization. We will now turn to the semantic analysis of these substrate-influenced usages of *also*, namely (i) sentence-final *also* in the negative expression of negation, (ii) sentence-final *also* co-occurring with the concessive *even*, (iii) sentence-final *also* with universal quantifiers, and (iv) double *also* in a clause, to be treated one after the other.

Also in negative sentences

In OHI, cases are found where the additive marker *also* is used sentence-finally instead of being substituted by the disjunctive additive marker *either*, as shown in in example (195):

- (195) *also* used in place of *either* in final position:
- a. They don't have the time to have a drink *also*. [OHI-001953-LAS]
 - b. So your lunch time wasn't fixed *also*? [OHI-001953-LAS]
 - c. And Su Inn Su Inn may not be happy *also*. [ICE-SG:S1A-057#31:1:A]
 - d. Uhm I haven't had my lunch *also*. [ICE-SG:S1A-091#7:1:A]
 - e. Ya he's doing that now but what what I find problem is he's he has not enough time *also* you know. [ICE-SG:S1A-015#44:1:A]

As mentioned earlier, the natural choice of Standard English in negative scope as in the above examples is clause-final *either*. According to Ziegeler (2017:187), the use of the additive *also* preceded by negation in the same clause (*Neg.V > also*) is not used at all in the ICE-GB data, while the *Neg.V > either* construction contributes to 69.4% of all the occurrences of focus particles under negation in ICE-GB.⁷¹ In contrast, there are 12 occurrences of clause-final *also* co-occurring with negation in the ICE-SG data, as shown in (195)c, d, and e. On the other hand, there are 10 occurrences of the *Neg.V > either*

⁷¹ The other construction included in the category is *also > neg. V*.

construction found in ICE-SG. In the OHI data, 6.25% (n=158) of the sentence-final *also* tokens occur in a negative sentence. These results suggest that the additive *also* in CSE – represented by both OHI and ICE-SG – has generalized across the positively and negatively polarized contexts, a sign of grammaticalization of CSE *also*.

Also with concessive even

The additive adverb *also* has developed subtle grammatical meanings when used with the concessive adverb *even* in the same clause. According to Bao (2014), the correlating use of *also* alongside the concessive *even* is ungrammatical in Standard English. In CSE, however, *also* appears to reinforce the concessive meaning of *even*, which is a clear replication of the Chinese *lián* and *yě/dōu* construction (see Chapter 6.4.3 on concessive *dōu* and *yě*). There are ample examples of such usage presented in both OHI and ICE-SG, as listed in (196) and (197):

(196) Examples from OHI

- a. Malays? I think, just...not only the Malays, *even* the Chinese *also*, even everybody during the Japanese, they suffered a bit. [OHI-000242-II]
- b. Most of them understand Malay. *Even* the Indians *also*, some of them Muslim Indians, some of them non-Muslim Indians. [OHI-000242-II]
- c. Sometimes *even* they marry a Tamil girl *also*, they *never* care their parents *also*, they go to the girl's... [OHI-000081-K]
- d. [...] of course dental health is very low in priority. *Even* the government *also*, they spent money, but when it comes to dentistry, they will think [...] [OHI-002130-LKH]

(197) Examples from ICE-SG

- a. *Even* dea[th] *also*, I don't know. Uh I mean if I'm the one who's going to die it's not so bad you know. [ICE-SG:S1A-028#23:1:A]
- b. Aye come on lah, *even* officers *also*, they wear plain shorts *also*, right? [ICE-SG:S1A-065#1:1:A]

The adverb *also* is analyzed as a concessive adverb in Bao (2014), and Siemund and Li (2017) also list one of the functions of *also* as concessive 'even'; other functions mentioned in Siemund and Li (2017) are addition, universal quantification (see the following subsection),

and a fourth adverbial meaning of ‘may as well’. Gast and van der Auwera (2011; 2013) demonstrated with extensive data that the scalar additive meaning ‘even’ is a well-established cross-linguistic feature of a general additive marker meaning ‘also’.

The co-occurrence of *even* and *also* resembles closely the Chinese substrates. Unlike English *even*, which serves as a stand-alone scalar additive marker, the additional scalar component of *lián* in Chinese has to be licensed by *yě* or *dōu* (see Chapter 5.2.2 and Chapter 6.4.3). Examples drawn from the CCL corpus syntactically match the construction in (196) and (197), as shown below:

- (198) 连 说话 也 没有 力气
lián shuōhuà yě méiyǒu lìqì
 even talk also no strength
 ‘I don’t even have the strength to speak.’
 [CCL-Contemporary-History-Li WenCheng]
- (199) a. Mandarin (personal knowledge)
 **lián shuōhuà méiyǒu lìqì*
 even talk no strength
- b. Cantonese
 **lin4 gong2ye5 mou5 hei3*
 even talk no strength
 intended meaning ‘I don’t even have the strength to speak.’

Example (198) matches the syntactic structure of the CSE example. An additional scalar reading of *lián* ‘even’ without *yě/dōu* ‘also’ is ungrammatical in Chinese, be it Mandarin, or Cantonese, as exemplified in (199). It suggests that the construction of *even* + *also* is modeled on *lián* + *yě/dōu*, in which *also* is used to license the additional scalar component of *even*.

Also with universal quantifier and dual/plural noun phrase

In addition to the use as an additive marker with an extension to negatively polarized contexts, and as a scalar adverb, tokens of clause-final *also* were found to co-occur with a

universal quantifier in both OHI and ICE-SG. As previously discussed in Chapter 5.2, such use of *also* appears to reinforce the universal quantifier (see Bao 2014).

- (200) a. [F]or three years during Japanese time, we *all* suffered a lot *also*. And no going out, no nothing. And very frightened of the Japanese.
[OHI-001663-MS]
- b. And this song you can hear *everywhere* in cinemas *also*, very patriotic song. [OHI-000483-TTC]
- c. Not just the lab lah, anywhere *also*. [ICE-SG:S1A-052#164:1:C]

However, examples were found in OHI, where sentence-final *also* does not co-occur with universal quantifiers, but with plural noun phrases, especially dual sets as in (201). Such usage cannot simply be analyzed as additive marker, as it does not point to the existence of an alternative to the associate of the additive (see Chapter 5.2.1).

- (201) a. Interviewer: Are you talking about the European nurses or both local and European?
MH: Locals and Europeans *also*. [OHI-002206-MH]
- b. It was not spoken of. Neither did we ask. By *both* mum *also*. Yes. No, they don't quibble about good, old days. [OHI-003223-CSS]
- c. Now like my neighbour and I, *both* sides and opposite *also*. When I have my fruits I say, "You want, take it, you're welcome." [OHI-002017-FKS]
- d. I think for *both* of us we need some kind of cupboard *also*, right.
[ICE-SG:S1A-054#236:1:A]

Yang and Wu (2019) proposed that Chinese *dōu* is a distributivity operator, which serves as a predicate over a plural noun phrase preceding it. Similarly, Xiang (2008) suggests that Chinese *dōu* is a maximality operator, which yields maximal plural individuals (see Chapter 6.4.2). Drawing analogy from these proposals on Chinese *dōu*, we see a similar pattern with the use of *also* found in OHI and ICE-SG. CSE *also* does not only serve to reinforce a universal quantifier, but also operates on a dual/plural noun phrase as in (201). As such, it can be moved even to the right periphery following the dual/plural noun phrase, functioning as post-determiner meaning 'both' or 'all'. As exemplified in the response of speaker MH to

the Interviewer's question in (201)a, *locals and Europeans also* can be interpreted as 'both the locals and the Europeans'. Again, this result provides strong evidence that CSE *also* has a greater range of functions than in Standard English.

Double use of also

Interestingly, as with double *already*, there are cases where *also* occurs twice in the same clause. Both uses of *also* conform to the lexifier morphosyntax though with one redundant use in each case, as shown in (202):

- (202) a. I think he's experienced *also* because he has his own music party, Panchara Muda. Yusoff *also* played mandolin *also*, musical instrument. [OHI-000242-II]
- b. Malaysia *also* found a way out. They *also* wanted to stop *also*, less Malaysians come here. [OHI-002198-TIT]
- c. So she *also* studied at CHIJ *also*?⁷² [OHI-002827-EC]
- d. So apart from visiting your company she *also* visited other companies *also*. [OHI-002827-EC]

The double use of *also* could be interpreted as a result of transfer from the local Chinese substrates (Bao and Hong 2006:110). However, double *also* does not appear to represent cases of present-day Chinese. We may perhaps interpret the double use of *also* as an attempt to adhere to the morphosyntax of both the lexifier and the substrate languages: the antecedent *also* follows the auxiliary as in Standard English, while the clause-final *also* is used as an additive scalar marker as in Chinese. Interestingly, it is not only amongst Chinese speakers that the double use of *also* is frequent, but also amongst Malay and Tamil speakers (see the analysis on frequencies of *also* vs. ethnicity in Section 8.2.3 below).

⁷² CHIJ Secondary (Toa Payoh), founded in 1854, is an autonomous all-girls Catholic school in Singapore.

8.2.3 *Also* vs. social variables

Frequencies of also vs. ethnicity

Both the highest frequencies of *also* and clause-final *also* occur amongst speakers in the Chinese group (speaker JN, accession number 002598, ethnicity: Chinese, frequency of *also*: 1.96 pp; speaker LAS, accession number 001953, ethnicity: Chinese, frequency of clause-final *also*: 0.64 pp). However, these were excluded from Figure 8.17 as outlier points. Curiously though, it is not the Chinese group that attains the highest average frequencies of *also* or sentence-final *also* (ReFreq *also*: 0.73, ReFreq CF-*also*: 0.14), but the Other group (ReFreq *also*: 0.84, ReFreq CF-*also*: 0.34). As shown in Figure 8.17, the frequencies of *also* and sentence-final *also* do not differ substantially among the Chinese, the Indian, the Malay, the Peranakan and the Other group. On the other hand, the Eurasian group and the British group generally have a lower frequency of *also* overall and a lower frequency of *also* in clause-final position.

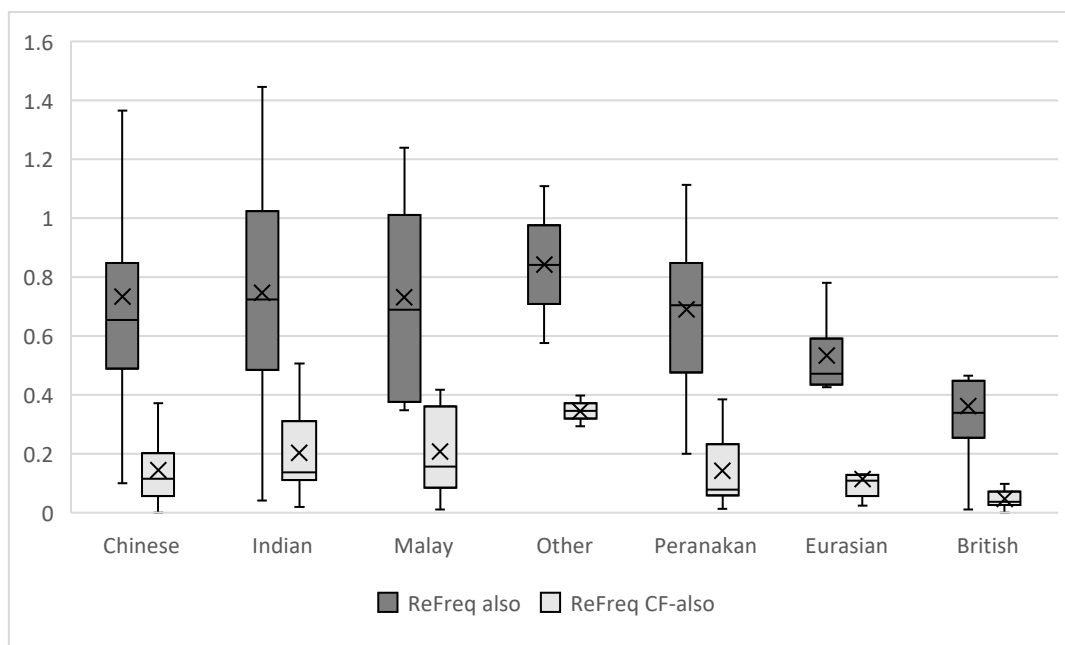


Figure 8.17: Frequencies of *also* and clause-final *also* pp in different ethnic groups

Different from what has been observed from the frequencies of *already* according to ethnic groups in Section 8.1.9, ethnicity seems to play a less significant role in affecting the frequencies of *also* (average frequency and clause-final frequency).

The result may suggest an indirect influence from CSE rather than a direct influence from the Chinese substrates on the use of *also* amongst Malay and Tamil speakers. However, it is also likely that Chinese is not the only source of transfer, as Malay and Tamil provide similar constructions of ‘also’ – Malay *pun*, Tamil *mēlum* (see Sharma 2012; Ziegeler 2017). A further study of the distribution of these constructions in Malay and Tamil would supply more information on the use of *also*, but this is beyond the scope of the present study. Apart from assessing the effect of the relevant substrates on the grammaticalization of *also*, we could integrate the recent discussion on vernacular universals in the investigation of variation in postcolonial varieties of English (Chambers 2004; Kortmann 2004; Sharma 2009). Sharma (2009), for example, examined three candidates (past tense omission, over-extension of the progressive, and copula omission) for English universals and found that past tense omission is genuinely similar in Indian English and Singapore English. There, she explains the similarity with typological parallels in the substrates. It is assumed that the functional extension of *also* may also be a shared feature among some varieties of English. Fuchs (2012) also investigated focus marking in Indian English and found that the usage of *also* in Indian English differs significantly from British English: it follows its focus immediately, has developed scalar and presentational use (or emphatic), and is often used in negative contexts, as shown in (203). He attributes the innovative uses of *also* to the additive clitics spread among indigenous Indo-European and Dravidian languages, such as Hindi *bhī* and Tamil *-um* (see Fuchs 2012:47). Like the Chinese *dōu*, the clitics *bhī* and Tamil *-um* share the additive function with *also* and the additive scalar meaning ‘even’. For more on the Hindi clitic *bhī* and its equivalents in other Indian languages, interested readers are referred to Koul (2008) on Hindi, Asher and Annamalai (2002) on Tamil, Smith and Paauw (2006) on Sri Lanka Malay and Sri Lanka Muslim Malay.

- (203) a. Postponement: We should have all the revision *also*.
[ICE-IND:S1A-087 #99:1:B]
- b. Additive scalar: You can do well or better than them *also*.
[ICE-IND:S1A-052#19:1:A]
- c. Presentational: That day *also* he has told us that in the middle of the
copper wire we put a magnet.
[ICE-IND:S1B-019#296:1:A]
- d. Negative context: [A]t the first move they don't like to talk with people
also. [ICE-IND:S1A-090#73:1:B]

In sum, *also* is expanding its functions rapidly in CSE amongst speakers with a Chinese, Indian, Malay, Peranakan background as well as amongst other ethnicities except for Eurasian and British. Presumably because of the additive function of *also* shared by the substrates spoken in the area, other uses (additive scalar, post-determiner, emphatic, *either*) were transferred to *also* (see Matras 2009:240–42 on pivot-matching in pattern replication; also see Chapter 4.2).

Also vs. educational level

As with *already*, speakers with a medium or high education level tend to use *also* less frequently than speakers with a low educational level. There is no significant difference between speakers with a medium educational level and a high academic level. In Figure 8.18, the categories of educational level were ordered on a cline from low to high. As can be seen, the innovative uses of *also* in clause-final position are most frequent among speakers with a low educational level, followed by speakers with a medium educational level, while speakers with a high academic level seldom use *also* in clause-final position. In other words, the frequencies of clause-final *also* decrease steadily with increasing educational level.

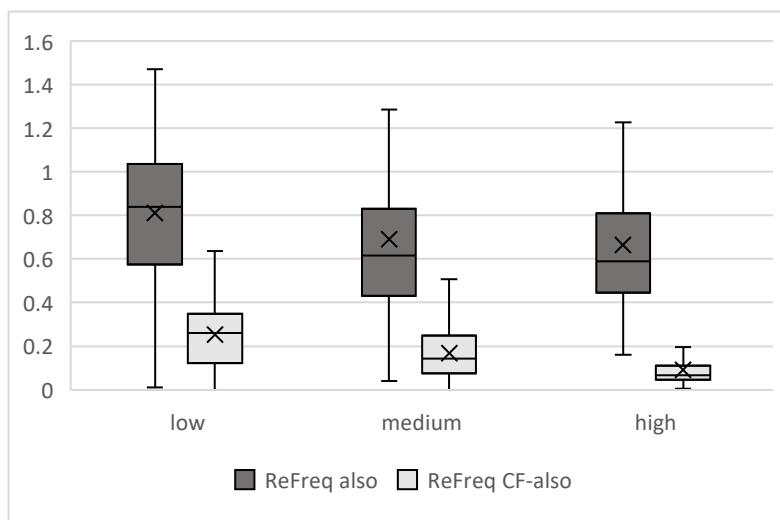


Figure 8.18: Frequencies of overall and clause-final *also* pp according to educational level

Frequency of also vs. age group and gender

Overall, age seems to have an influence on the frequencies of *also* while gender does not have an impact on the ratio of *also* or clause-final *also*.

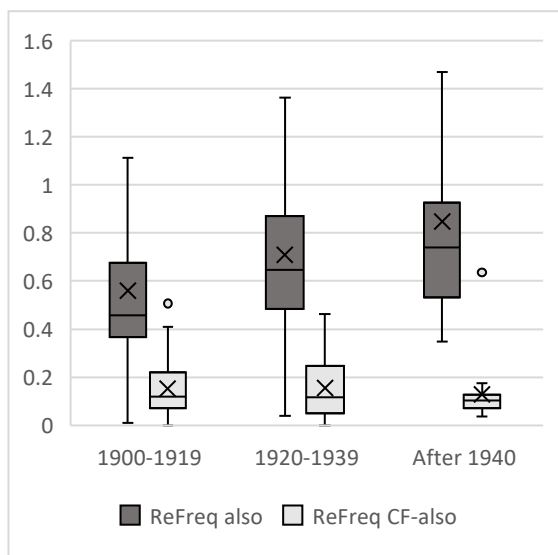


Figure 8.19: Frequencies of overall and clause-final *also* pp according to age group

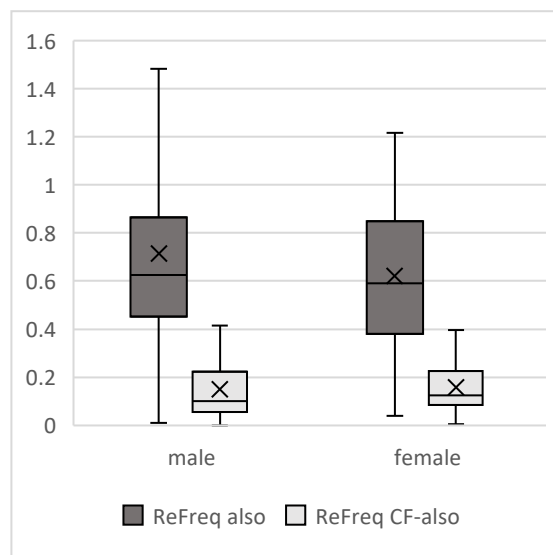


Figure 8.20: Frequencies of overall and clause-final *also* pp according to gender

From Figure 8.19, we can see a clear pattern of age grading with an increased use of *also* overall. However, an increased frequency pattern of clause-final *also* between speakers born in 1920–1939 and after 1940 is hardly recognizable. Although the highest frequency of

clause-final *also* occurs in the youngest group (with speaker LAS, accession number: 001953, born in 1940, ethnicity: Chinese, educational level: low), the data point counts as an outlier. The average frequency of *also* in the group of after 1940 is lower than the groups of 1900–1919 and 1920–1939. The reason is that the majority (61.1%) of the interviewees in the youngest group have a high educational level, while only 5.7% of the speakers have a low educational level. As such, the speakers' educational level may be a more significant factor influencing the frequency of clause-final *also*.

Multiple linear regression analysis

	Estimate	Std. Error	t	Sig.
(Intercept)	0.332	0.042	7.969	2.95E-12
Education	-0.081	0.018	-4.533	1.60E-05

R-Square: 0.173

F-statistic: 20.547 on 1 and 98, p-value: 0.000016

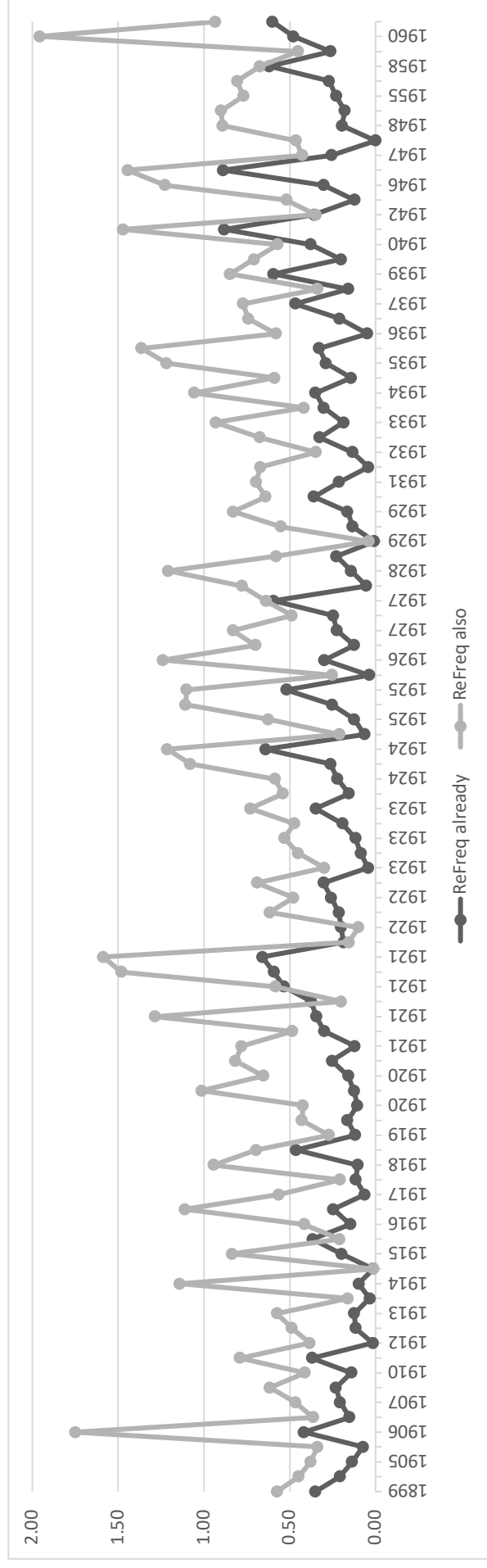
Table 8.3: Multiple linear model on relative frequency of clause-final *also*

A multiple linear regression was calculated to predict the frequencies of clause-final *also* based on the social factors: ethnicity, education, age group, and gender. As predicted, education level is the best predictor of the frequency of clause-final *also*. A significant regression equation was found ($F(1, 98) = 20.547, p = 0.000016$), with an R^2 of 0.173 (see Table 8.3). Except for education, all the other three variables did not add statistically significantly to the prediction and were therefore excluded from the model (gender: $t = 0.692, p = 0.491$; ethnicity: $t = -0.966, p = 0.337$; age group: $t = 0.143, p = 0.887$).

8.2.4 Positive correlation between frequencies of *already* and *also*

Figure 8.21 shows the frequencies of *already* and *also* measured as per page in the interviews delivered by the 100 speakers selected in the study. We can see there are substantial differences in the frequencies produced by the individual speakers, with high numbers of

already and *also* pointing to the more basilectal speakers. Curiously enough, the columns (black vs. grey) in Figure 8.21 seem to suggest that the frequencies of *already* and *also* are positively correlated. For a significant number of speakers, high frequency of *already* seems to imply high frequency of *also*, and vice versa.

Figure 8.21: Frequencies of *already* and *also* of individual speakers in OHI born between 1899–1983 (pp)

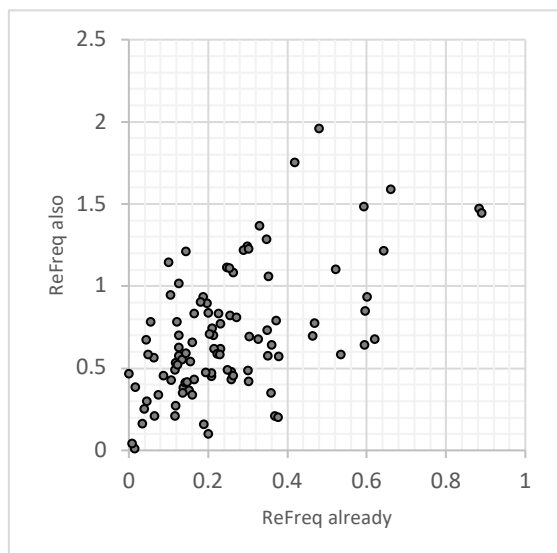


Figure 8.22: Scatter plot of Relative frequency of *already* vs. *also* pp

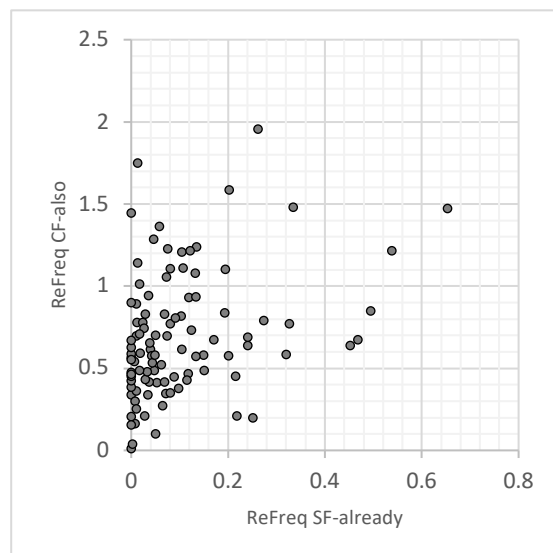


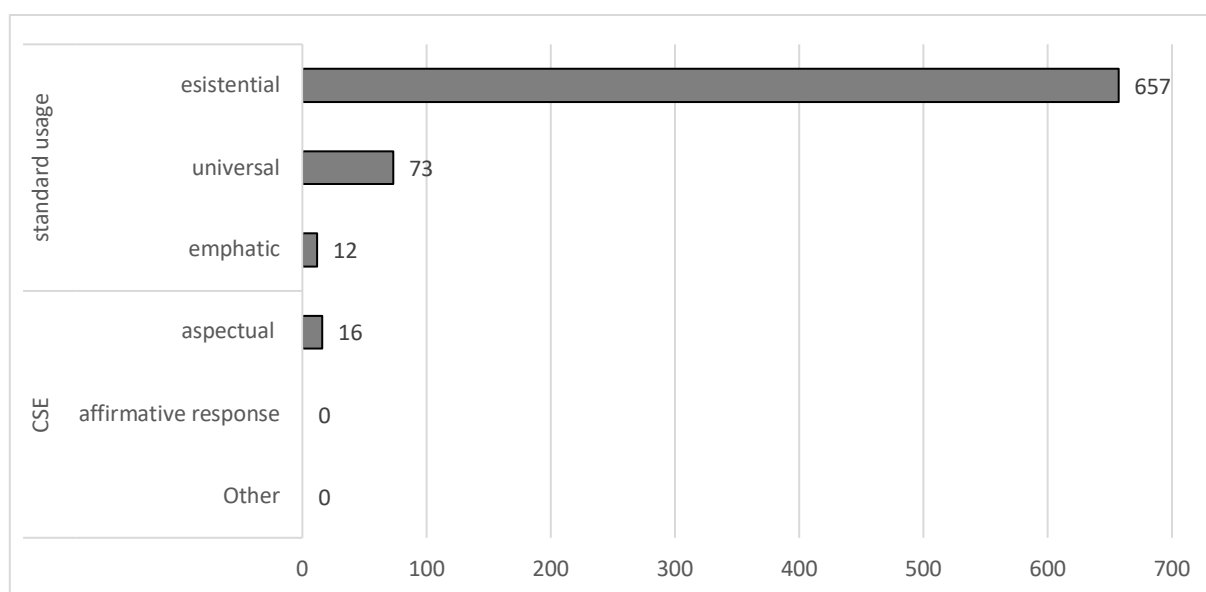
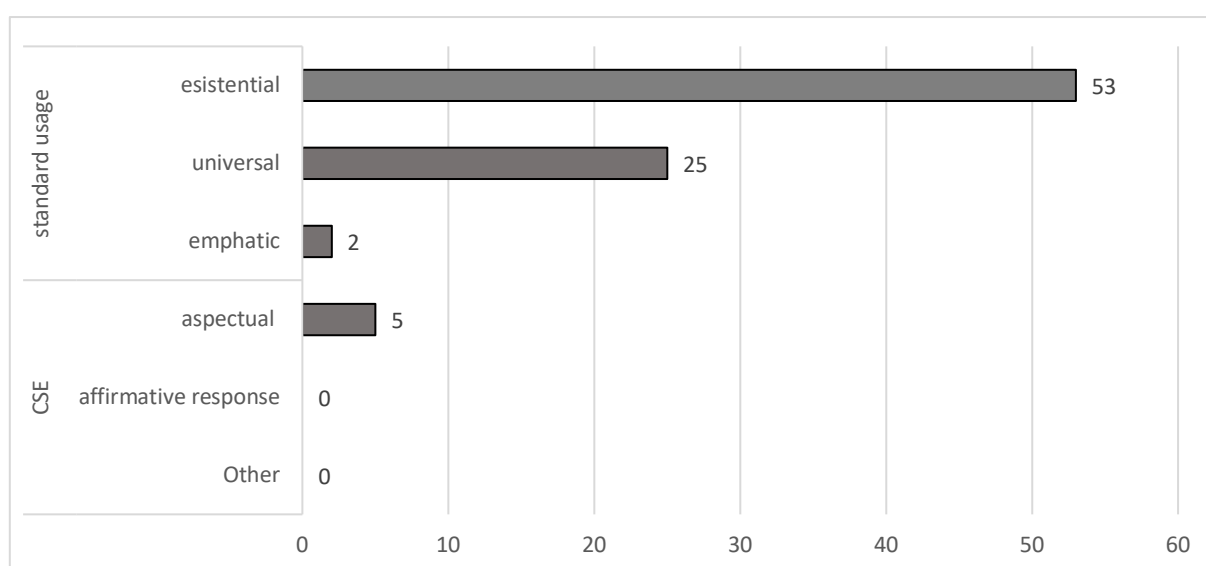
Figure 8.23: Scatter plot of Relative frequency of CF-*already* vs. CF-*also* pp

By applying the Pearson's correlation coefficient, it was found that the frequencies of *already* and *also* are positively correlated, $r(98)=0.562$, $p=1.1398e-9$. There is also a positive correlation between the frequencies of clause-final *already* and clause-final *also*, $r(98)=0.491$, $p=2.1079e-7$. The results suggest that higher ratios of *already* and *also* are indicative of higher substrate/basilectal influence.

8.3 Analysis of *ever*

8.3.1 Frequencies of *ever* according to semantic categories

Compared with *already* and *also*, the frequency of *ever* is much lower: 760 occurrences of *ever* were found in OHI and 85 occurrences were found in ICE-SG. Among these, only 16 cases (2.11%) in OHI and 5 cases in ICE-SG (5.89%) were found to have developed novel meanings parallel to the Chinese experiential marker 过 *guò* (see Chapter 6.3.1). Apart from that, there was no occurrence of *ever* in affirmative responses to polar questions in OHI or ICE-SG (see Ho and Wong 2001, also Chapter 5.3.3).

Figure 8.24: Absolute frequencies of *ever* according to different functions in OHIFigure 8.25: Absolute frequencies of *ever* according to different functions in ICE-SG

As mentioned earlier, the only substrate-influenced use of *ever* found in OHI and ICE-SG is its aspectual usage as an experiential marker. Example (204)a describes a scene after a fire accident happened in the neighborhood of the interviewee. Here the sentence can be interpreted as ‘We have seen a big dog’. The same applies to (204)b, which can be glossed as ‘I have seen that being mentioned in the media’. Both examples imply a sense of ‘at least once’ in addition to the present perfect, which conveys the “indefiniteness in specification of the anterior event in the time period leading to the moment of speaking” (Ziegeler

2015:120). Interestingly, *ever* and *already* co-occur in the same context in (204)a. In (204)c, *already* and *ever* even co-occur in the same clause. In both examples, the experiential meaning conveyed by *ever* is distinctive from the completive meaning expressed by *already*.

- (204) a. All you see is black. We *ever* saw one big dog, all in black colour, turned like that. They roasted *already*. [OHI-001953-LAS]
- b. It is said that, I think, his ability to judge people is very, very strong. I mean I *ever* see that being mentioned in the media. [OHI-003223-CSS]
- c. And then that time my own son had diarrhea or what. And then he recommend Done's cholera mixture, I still remember. That old man, now I think out of date *already*. And I myself have *ever* taken *already*. Like chili so hot and it really cured. [OHI-001631-OCN]
- d. And he *ever* ask me uh luckily uh when I was young. [ICE-SG:S1A- 065#115:2:F]

In (204)b and (204)d, *ever* occurs with a verb in its bare form, which unequivocally shows that *ever* functions as an aspectual marker. However, there are also cases of *ever* co-occurring with present perfect in non-negative sentences, as exemplified in (205):

- (205) a. What big deal delivering a breech. I have *ever* delivered a breech in the house alone. And you have so many helpers besides you. [OHI-002206-MH]
- b. And Senu related to me that had never met a man like Sukarno who liked women very much. He said that he had *ever* accompanied Sukarno. [OHI-000133-OW]
- c. The man said that the woman has ever compere[d]. [ICE-SG:S1A-006#197:1:A]

The examples show that there is a double marking of the function of the existential/experiential aspect in CSE, i.e. there is a layering of the means of expressing the sense of 'happening once' or 'experiencing once' in the indefinite past by *ever* and the morphological means of the English perfect. For example, in (205)a, the speaker MH, an

experienced midwife, depicted an incident in which the breech of the baby came out first. She calmed the doctor down by saying that she had once encountered the same situation and asked him not to panic. Here, both the experiential *ever* and the present perfect were used to denote the meaning of ‘once at least’ in the indefinite past period leading to the utterance time.

According to Hopper (1991), “layering” refers to the co-existence of the old and new forms in the process of grammaticalization, before the new form eventually replaces the older one (Hopper 1991:22). Lehmann (2015:23–24) emphasizes that the replacement of an old construction takes its time. An example he gives is the competition between the new analytic and the old synthetic perfect in the Romance language (*passé composé* vs. *passé simple*, e.g. *elle s’est évanouie* replacing *elle s’évanouit* ‘she fainted’). During this phase, the two constructions are not functionally identical, and therefore there exist two distinctive categories instead of one (Hopper and Traugott 2006:9; Lehmann 2015:24). Hopper and Traugott (2006:9) note that there is a tendency for periphrastic forms to replace morphological ones over time – known as “renewal” – a process can be seen to occur repeatedly. They gave the example of the etymological source of the French future form, as illustrated in Figure 8.26.

Pre-Latin		Latin		French
*?				
*kanta b ^h umos	>	cantabimus		
		cantare habemus	>	chanterons
				allons chanter > ?

Figure 8.26: Grammaticalization chain of the French future form (adapted from Hopper and Traugott 2006:9)

According to Hopper and Traugott (2006:9), the etymological source of the French future form *chanterons* ‘we will sing’ can be traced back to the earlier periphrastic construction **kanta b^h umos* from Indo-European. It was later reconstructed as *cantabimus*, an older morphological future form in Latin. A later periphrastic future form *cantare habemus* had

been competing with it for several centuries, before the latter eventually replaced the former. The French future form *chanterons* ‘we will sing’ is an inflectional form, and is being replaced by (*nous*) *allons chanter*, literally ‘we are going to sing’.

As mentioned earlier, the proportion of *ever* functioning as an aspectual marker is relatively small (2.11% in OHI and 5.89% in ICE-SG). Judging from these figures, it is difficult to predict whether *ever* will continue the process of further grammaticalizing into an aspectual marker, with the concomitant loss of the morphological marking of the English perfect. The other examples are parallel to the canonical usages of *ever* in Standard English. These correspond to the following usages of *ever*: (i) the existential use meaning ‘at any time’ in interrogative and negative sentences, as in (206); (ii) the universal use meaning ‘constantly’, ‘at all times’ or ‘always’ in *ever since*, with an adjective, as well as in the comparative form *as adj. as ever*, as in (207); and (iii) the emphatic use with *so*, as in (208) (also see Chapter 5.3.1 and Chapter 7.3.3).

(206) Existential *ever*

- a. [D]id you *ever* approach the Japanese for anything? [OHI-000284-ARKS]
- b. I don’t think you could *ever* afford to live by the sea. The prices of property by the sea just went up and up. [OHI-002044-RF]
- c. I mean like there was one section on uhm what you call meaning-centred uh motivate meaning-centred programme or approach uhm which she hardly *ever* touched which she <unclear> word </unclear> hardly ever touched on it. [ICE-SG:S1A-075#69:1:A]

(207) Universal *ever*

- a. Oh, antique house. Well, *ever since* I knew Peter, he already started the antique house, but I still think it’s the post war period. [OHI-002186-LLH]
- b. It has been maintained *ever since*. [ICE-SG:S2B-035#92:3:C]
- c. The second point to say about him is that he is a very strong supporter of Peranakan culture you see. So he is *ever willing* to share his knowledge of Peranakan culture with people who are interested, who take the trouble

- to
 come to his place and chit chat with him. [OHI-002186-LLH]
- d. The burden of the *ever increasing* expenditures ultimately falls back on the people. [ICE-SG:S2B-048#X124:2:B]
- e. They were not allowed to tidy their hair, or put on lipstick, powder, nothing at all, just made yourself *as dowdy as ever*. [OHI-000237-HCY]
- (208) Emphatic *ever*
- a. Another thing, for example, last year when they had this communal problem in Sri Lanka, how my children... not only here, we called them *ever so* often to find out how they were. [OHI-000345-AST]
- b. Congratulations indeed to Suriati Marijar who did *ever so* well. [ICE-SG:S2A-011#25:1:A]

Lehmann (2015:24) notes that the functional similarity of the two competing grammatical constructions to a large extent depends on their formal similarity. It is very likely that a complete replacement of the old function will never be obtained if the path towards grammaticalization does not lead to formal similarity to the earlier construction. A case Lehmann (2015) mentioned is the partial replacement (referred to by Lehmann (2015:24) as “renovation”) of Latin case inflection by prepositional construction. Also argued by Schøsler (2008:428), a general replacement of the genitive does not exist, but rather a renewal of specific uses of the genitive by different prepositions or syntactic functions. For instance, the genitive argument in Latin is expressed in different ways in Modern French, e.g. for the verb *accusare* ‘accuse’ the genitive argument is expressed by a Prepositional Phrase (PP) headed by *de* in Modern French, but for *damnare* and *condemnare*, it corresponds to a PP headed by the preposition *à* (Latin: *capitis condemnare* / Mod. Fr. *condamner à mort* ‘sentence to death’), and for *oblivisci* ‘forget’ the genitive argument was substituted by an accusative in Modern French (Carrier, Goyens, and Lamiroy 2013).

Although situations in contact-induced grammaticalization may not be the same as in monolingual grammaticalization, Heine and Kuteva (2005:169–70) found similar examples of partial replacement of the inflectional means of expressing the comparative of inequality (‘more than’) in Luxembourgian German by the analytical means of French *plus*.

Similarly, CSE *ever* does not share formal similarity with the English perfect despite their functional similarity. Therefore, we predict that the co-existence of *ever* and the English perfect in marking the experiential/existential aspect may remain stable in the near future. Apart from that, Ziegeler (2015:44) found that 6.67% of *ever* functions as an experiential marker based on a corpus of 85,909 words available from a local Singaporean Internet forum: Flowerpod (<http://flowerpod.com.sg>). One example from the Flowerpod corpus is shown in (209):

- (209) [I] *ever* thought of joining this industry, but heard lots of people saying tt
[‘that’] u play the bad guy in the co[mpany] and stuff.
(Posted by diamonds Aug 21, 2007, cited in Ziegeler 2015:44)

The Flowerpod corpus was constructed in 2007–2009 and therefore represents a more current version of CSE than OHI and ICE-SG. The ratio of *ever* as an aspectual marker in the Flowerpod corpus (6.67%, 2007–2009) is higher than that in OHI (2.11%, early 20th century) and ICE-SG (5.89%, 1997). These figures seem to suggest an ongoing grammaticalization of *ever* as an aspectual marker in CSE.

8.3.2 Frequencies of *ever* vs. social variables

In the following section, we will again first look at the frequency of *ever* according to each individual social variable (i.e. ethnicity, educational level, age group, and gender) before estimating the total effect of them on the frequencies of *ever* based on the multiple linear regression analysis.

Ever vs. ethnicity and educational level

Different from *already* and *also*, which serve as good indicators of substrate/basilectal influence, the frequency of *ever* alone seems to be an index of lexifier influence. From Figure 8.27, we can see that the British and Eurasian groups use *ever* more frequently,

while the Chinese, the Indian and the Other groups have relatively lower frequencies of *ever*. On the other hand, the Malay and the Peranakan groups behave similarly, having a frequency of 0.048–0.057 per page.

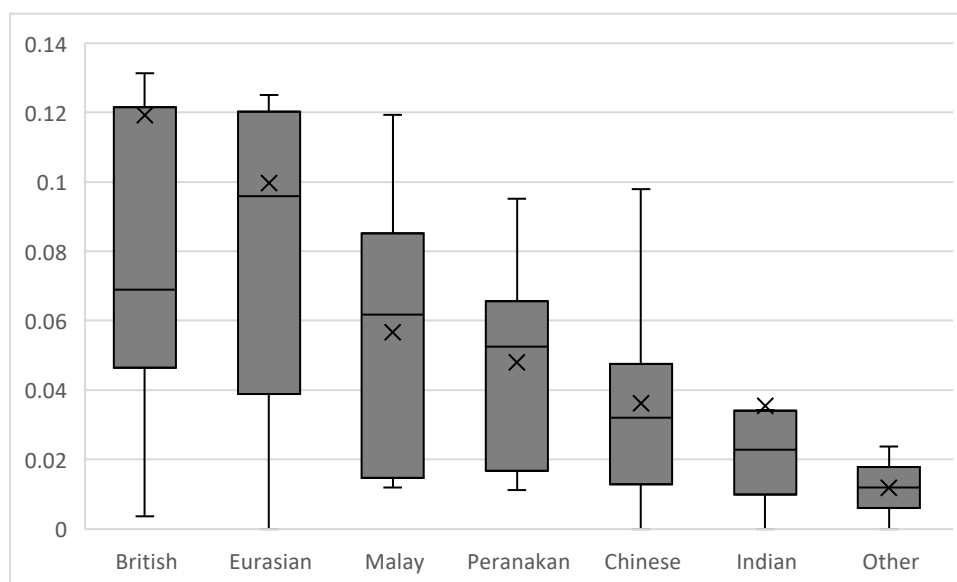


Figure 8.27: Frequency of *ever* pp in OHI according to ethnic communities

As mentioned earlier, the frequency of *ever* is much lower than *already* and *also* in OHI and ICE-SG. In addition, 97.89% of the occurrences of *ever* in OHI and 94.11% of those in ICE-SG do not differ from Standard English. Therefore, it is not surprising that the British and Eurasian groups demonstrate the highest frequencies of *ever*. The Chinese group attains the third lowest frequency of *ever* (0.036 pp) on average, with two outliers having a frequency of 0.15 and 0.11 per page.⁷³ These two speakers have an educational level from medium to high. As such, we can observe that the frequency results of *ever* in relation to social factors such as ethnicity and educational level are almost opposite to *already* and *also*: while higher ratios of *already* and *also* suggest higher substrate/basilectal influences, higher ratio of *ever* indicates higher superstrate/acrolectal influences.

⁷³ Two outlier points are not represented by Figure 8.27.

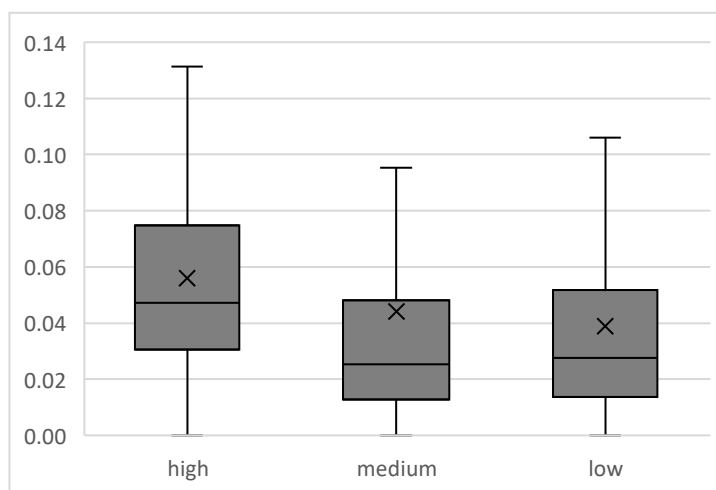


Figure 8.28: Frequency of *ever* according to educational level

However, the result of *ever* in relation to educational level only partially agrees with the above observation. As shown in Figure 8.28, the frequencies of *ever* produced by speakers with a high educational level (0.06 pp on average) are higher than those by speakers with a medium or low educational level (both 0.04 pp on average). Curiously, the frequencies of *ever* produced by speakers with a medium educational level do not differ significantly from those produced by speakers with a low educational level.

Ever vs. gender and age

Neither gender nor age seems to have an influence on the frequencies of *ever*. As shown in Figure 8.29, male speakers behave similarly to their female counterparts. Apart from that, there are no significant differences among speakers coming from different age groups. Speakers belonging to the age group 1900–1919 produced an average frequency of *ever* of 0.052 pp, the group 1920–1939 have a frequency of 0.047 pp, while speakers born after 1940 attain a frequency of 0.044 pp on average.

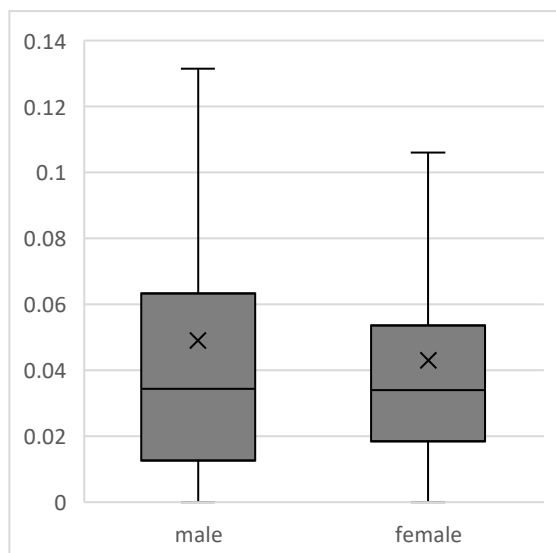


Figure 8.29: Frequencies of *ever pp* according to gender

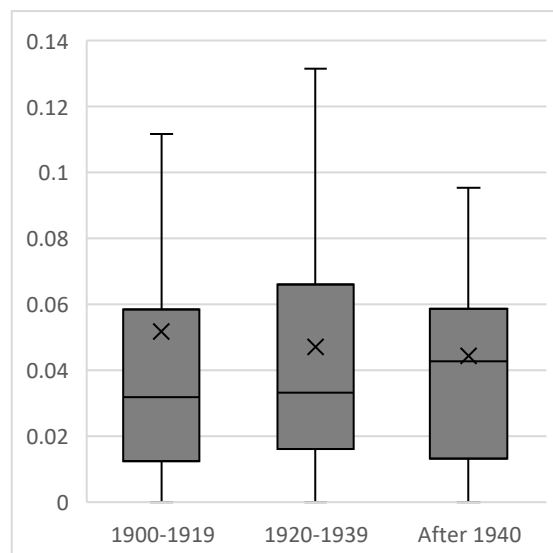


Figure 8.30: Frequencies of *ever pp* according to age group

In sum, we observe that only ethnicity has a strong impact on the overall frequency of *ever*. With regard to the correlation between frequency of *ever* and educational level, although speakers with a high educational level have a higher frequency of *ever* than speakers of a medium or low educational level, the latter two groups do not differ from each other significantly. Apart from that, there seems to be no correlation between the use of *ever* and speakers' gender or age group.

Multiple linear regression analysis

The multiple linear regression analysis confirms the above observations. Except for ethnicity ($t=3.545$, $p=0.000604$), all the other three variables did not add statistically significantly to the prediction. Therefore, the factors age, gender and educational level were excluded from the model, as shown in Table 8.5.

	Estimate	Std. Error	t	Sig.
(Intercept)	0.024	0.008	2.859	0.005193
Ethnicity	0.010	0.003	3.545	0.000604

R-Square: 0.114

F-statistic: 12.567 on 1 and 98, p-value: 0.000604

Table 8.4: Multiple linear model on relative frequency of *ever*

	Estimate	t	Sig.	Partial Correlation
Education	0.165	1.729	0.087012	0.173
Age group	0.015	0.157	0.875327	0.016
Gender	0.008	0.079	0.937117	0.008

Table 8.5: Excluded variables of multiple linear regression model on relative frequency of *ever*

The low frequency of substrate-influenced *ever* could be explained by the constraints of Bao’s (2005) System Transfer and Lexifier Filter theory. As mentioned in Chapter 4.7.4, an entire grammatical subsystem is involved in substrate transfer in the process of system transfer. Therefore, the grammatical subsystem of the transferred items in the new language typically resembles that of the substrate language. This process actually implies a constraint, which favors a transferred grammatical category that conforms to the substrate. On the other hand, the lexifier language also imposes a constraint on the transferred item in the new language, which “stipulates that the exponence of the transferred grammatical system conform to the morphosyntax of the lexical source language.” (Bao 2005:259)

Based on the above rivalry between the constraints of the lexifier language and the substrate language, Bao (2005) outlines four possible scenarios in a contact situation depending on whether the morphosyntax of a transferred category is compliant with the lexifier, with the substrate, with both the lexifier and the substrate, or with neither of them. These possibilities are listed below:

	Lexifier-compliant	Substrate-compliant	Example
a.	yes	yes	<i>already</i>
b.	yes	no	<i>ever</i>
c.	no	yes	not transferred
d.	no	no	non-existent

Table 8.6: Transfer and morphosyntactic compliance (adapted from Bao 2005:259)

The use of *ever* in CSE belongs to scenario b, as shown in Table 8.6. The syntax of CSE *ever* is preverbal, which mirrors that of Standard English, but the substrate Chinese *guò* is post-verbal (see Chapter 6.3 on *guò*). Here the syntax of the lexifier source *ever* is in conflict with the substrate *guò*, which may explain the low frequency of CSE *ever*, especially among speakers with a Chinese background. In comparison with aspectual *ever*, aspectual *already* has achieved a much higher frequency in CSE (see Section 8.1.3). Different from aspectual *ever*, the syntactic position of *already* in Standard English overlaps with that of the Chinese *le*: both English *already* and Chinese *le* can be clause- or sentence-final, though English *already* prefers sentence-medial (preverbal) position. The concordance between the syntax of *already* and *le* may have facilitated the transfer process, and therefore, aspectual *already* occurs substantially more often than *ever* in CSE.

However, it is also essential to assess how frequent the need for the experiential function is in Standard English and in Chinese, i.e. how often we use the perfect aspect or Chinese *guò* to express an experience that happened at least once in the past. It is perhaps more often that we need to express the negative, *never*, but the zero-marked positive experiential in English is less frequent (see Chapter 5.3.1). A further study of the distribution of zero positives versus *never* in Standard English would supply more information on the frequency of the experiential function in Standard English. For instance, we could search for a restricted lexical range in a corpus like the BNC (e.g. been to X) and look for the frequency of zero positives versus *never* in Standard English. It would give us a better idea of the frequency to expect in another corpus, but such an investigation is beyond the scope of the present study.

8.4 Analysis of *one*

8.4.1 Frequencies of phrase-final *one* in OHI and ICE-SG

As previously introduced in Chapter 7.3.4, this study focuses on the use of phrase-final *one* in OHI and ICE-SG. Altogether 2,469 and 522 occurrences of phrase-final *one* were elicited in OHI and ICE-SG. The pronominal and numeral usages of *one* belong to Standard English usages (see Chapter 5.4.1). On the other hand, CSE *one* functions as a relative/nominalizer and an emphatic marker modeled on the Chinese substrate 的 *de* (see Chapter 5.4.2 on relative/nominalizer *one*, Chapter 5.4.3 on emphatic *one* and Chapter 6.5 on Chinese 的 *de*). Apart from that, tokens of CSE *one* were found following a proper noun or a possessive pronoun (see Chapter 7.3.4). These were categorized as substrate-influenced uses as the pronominal function of *one* in Standard English generally follows an adjective (see Chapter 5.4.1).

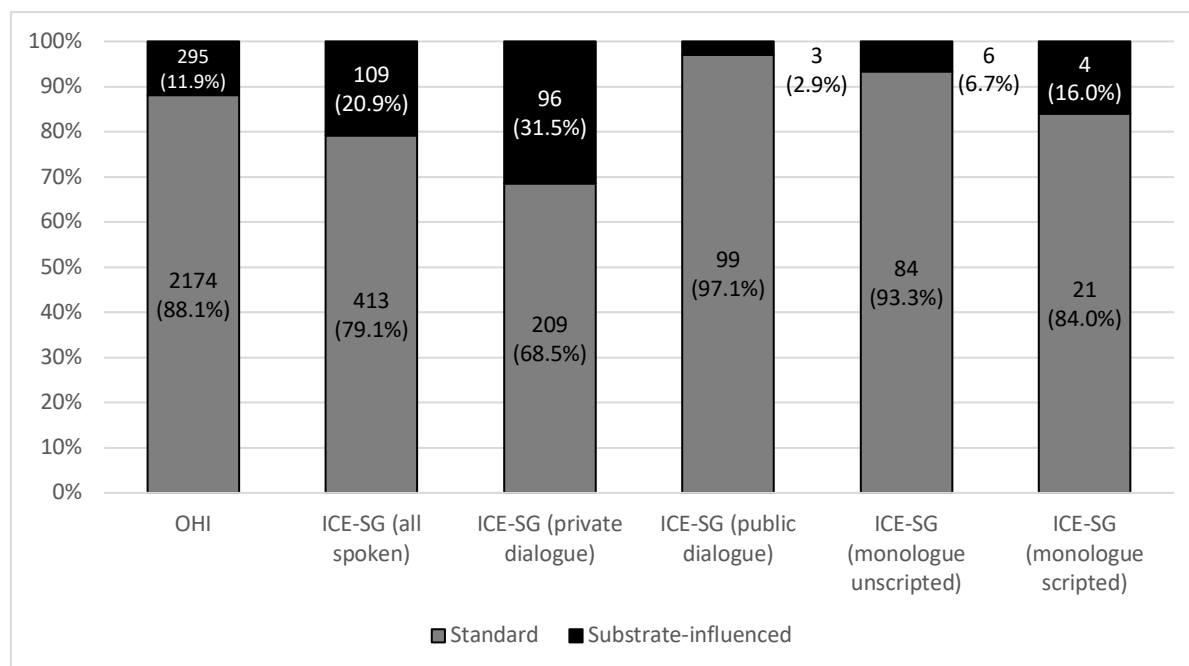


Figure 8.31: The proportion of substrate-influenced *one* vs. standard *one* in OHI and ICE-SG

The overall data presented in Figure 8.31 shows how the proportions of substrate-influenced *one* and standard uses of *one* in OHI differ from those in ICE-SG (chi-square=29.4328,

$p < 0.00001$). Substrate-influenced *one* constitutes 11.9% of the total tokens of *one* in phrase-final position in OHI, while the same category constitutes 20.9% of all phrase-final *one* in ICE-SG. As with the other three CSE expressions, it is necessary to consider the differences according to different text types. Again, there is a preference for using *one* with substrate-influenced features in the private dialogues (31.5%), while there is a marginal percentage of substrate-influenced *one* in the public dialogues (2.9%) and the unscripted monologues (6.7%). The proportion of substrate-influenced *one* in the scripted monologues, however, is not representative, as only a small number of *one* (25 in total) were found in phrase-final position in this text type. The following subsection will take a look at the frequencies of phrase-final *one* according to its semantic categories in more detail.

8.4.2 Frequencies of phrase-final *one* and its semantic categories

The majority of *one* tokens in phrase-final position in both datasets were used in a standard way (88.1% in OHI; 79.1% in ICE-SG). These occurrences of *one* exhibit either a pronominal or numeral function: 55.0% (n=1,358) of phrase-final *one* were found to be pronominal such as in (210) and 33.1% (n=816) serve a numeral function as in (211). In ICE-SG, pronominal *one* and numeral *one* account for 43.7% (n=247) and 31.8% (n=166) of all occurrences of phrase-final *one*, respectively.

(210) pronominal *one*:

- a. It was a very closed party. It was a very powerful *one*, about half a million members as I remember rightly, and with a powerful influence both inside Parliament and outside it. [OHI-000166-DB]
- b. And then halfway before she reaches home, she has to take off her mourning flower, the white *one*, and she has to wear a red *one*. [OHI-001631-OCN]
- c. This is the only *one*. [ICE-SG:S1A-069#28:1:B]

(211) numeral *one*:

- a. One group of several girls learnt Book *One*, and then another group Book *Two*, and then another group Book *Three*, like that. [OHI-000237-HCY]

- b. So on the spot, we pick up all their problems *one* by *one*, and then we talk to Housing Board, *one* by *one* also. [OHI-000454-ST5]
- c. It is seen in P thirty-five and thirty-*one*. [ICE-SG:S1B-070#70:1:C]

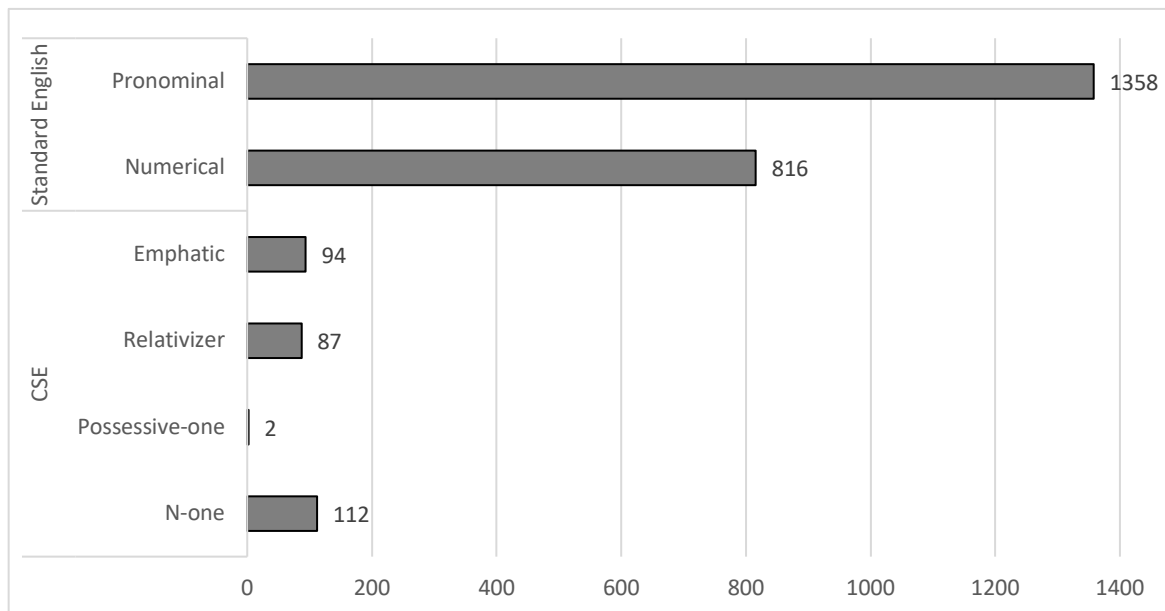


Figure 8.32: Frequencies of *one* according to different semantic categories in OHI (absolute figures)

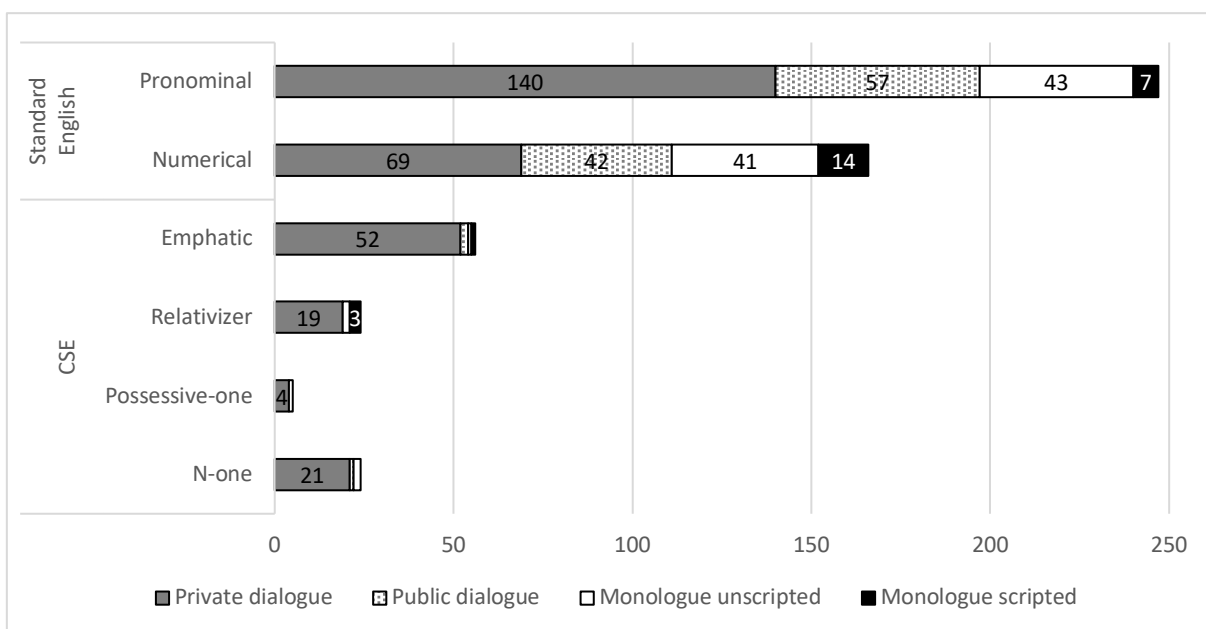


Figure 8.33: Frequencies of *one* according to different semantic categories in ICE-SG (absolute figures)

The major difference between OHI and ICE-SG occurs in the use of substrate-influenced *one*, especially in the emphatic use, as exemplified in (212). As shown in Figure 8.32 and

Figure 8.33, there is a sizeable increase in the use of emphatic *one* from 3.8% (OHI, n=94) to 10.7% (ICE-SG, n=55). Apart from that, there is a slight increase in the use of *one* as a relativizer/nominalizer by comparing OHI (3.5%, n=87) and ICE-SG (4.6%, n=24), as exemplified in (213). It is worth mentioning that the vast majority of emphatic *one* (92.9%, n=52) in ICE-SG occur in the private dialogues.

- (212) emphatic *one*
- a. This type of women, they can stand hardship *one*. [OHI-001953-LAS]
 - b. You know, men are always very simple *one*. [OHI-001953-LAS]
 - c. Because I always wake up early *one*. As for me I won't late *one*.
[OHI-001953-LAS]
 - d. Everytime PAP also win *one*. [ICE-SG:S1S-004#191:1:A]
 - e. Ya what the lab the lab things spread very fast *one*.
[ICE-SG:S1A-052#163:1:B]
 - f. I always use microwave *one*, you know mix. [ICE-SG:S1A-006#150:1:A]
 - g. Two metre crab can eat *one*, you know <unclear> word </unclear> shiok
ah. [ICE-SG :S1A-085#392:1:A]
- (213) relativizer/nominalizer
- a. That means bottom is shop, the top is people staying *one*.
[OHI-001953-LAS]
 - b. But this one is push *one*. But they have the four railings.
[OHI-001953-LAS]
 - c. But Fuji Steel idea is to have very low capital, high loan because this is
government supported *one*. [OHI-002198-TIT]
 - d. Whoa all those uh from London *one* nuh [...] [ICE-SG:S1A-096#290:2:E]

As discussed in Chapter 5.4.3, emphatic *one* serves a pragmatically oriented function. In the examples in (212), *one* can be deleted without making the syntactic structure incomplete. However, it is semantically obligatory as it underlines an entity as particularly salient in a category, e.g. *the ability to stand hardship* in (212)a, *always being simple* in (212)b, and *always waking up early* in (212)c, and *being able to win every time* in (212)d. Chapter 6.5.2 relates emphatic *one* to sentence-final *de* as part of the cleft structure *shì...de*. The argument applies to the examples of emphatic *one* in (212) as well. For instance, (212)a can be interpreted as 'it is this type of women who can stand hardship'. The same is relevant in (212)b, 'it is men who are always very simple'. Another piece of evidence to support the

argument for a Chinese influence on emphatic *one* is that the majority (81.9%) of the occurrences of emphatic *one* were produced by speakers amongst the Chinese group (also see Section 8.4.3).

In addition to the emphatic function, *one* functions as a nominalizer/relativizer. The structural frame of nominalizer *one* (XP-*one*) is similar to emphatic *one*. Yet, different from emphatic *one*, which can be omitted without making the sentence structurally incomplete, nominalizer/relativizer *one* is obligatory both from a syntactic and a semantic perspective. It converts any phrase into an NP (see Chapter 5.4.2). For example, in (213)a, *one* attaches to the VP *people staying*, turning it into an NP, meaning ‘the floor in which people live’. In (213)b, *this one is push one* refers to ‘a type of cart that pushes the coal’. The same applies to (213)c, where *one* transforms the VP *government supported* into an NP, meaning ‘the project that was supported by the government’. What is more, the PP *from London* in (213)d is nominalized by *one* into ‘the people from London’.

Pronominal *one* in CSE extends to co-occur with possessive nouns and nominal phrases, which are ungrammatical in Standard English. As exemplified in (214), it co-occurs with a possessive noun such as *Gazali’s*, *Tang’s* and even the possessive pronoun *my* to form a nominal referent: *Gazali’s one* refers to ‘a book written by Gazali’, *Tang’s one* refers to ‘a kind of lifestyle of the Tang family’, and *my one* refers to ‘my phone number’. On the other hand, *one* co-occurs with a noun, e.g. *credit card one*, *the red type one*, *the Paya Lebar one*, *the Romeo and Juliet one* in (215). Like *one* co-occurring with a possessive noun, N-*one* is also used to form a nominal referent that was already mentioned in a previous discourse, e.g. *credit card one* refers to ‘the payment of credit card’. There is a slight increase in the use of possessive N-*one* from 0.8% (OHI, n=2) to 1.0% (ICE-SG, n=5). On the other hand, the proportion of *one* forming an NP with a previous noun has remained relatively stable (4.5%, n=112 in OHI vs. 4.6% , n=24 in ICE-SG). It is worth noting that the public dialogues and monologues do not favor *one* with the above substrate-influenced features.

- (214) possessive N-*one*
 a. I haven’t read *Gazali’s one*, but I will now. [ICE-SG:S1A-047#189:1:A]

- b. Do they follow the Tang's *one*? [ICE-SG:S1A-082#12:1:B]
 - c. No, I never put down my *one*. [ICE-SG:S1A-079#234:2:D]
- (215) N-*one*
- a. Credit card *one*. [ICE-SG:S1B-073#85:1:B]
 - b. But this one is the blue type one, not the red type *one*. They are very tough. [OHI-001953-LAS]
 - c. And Joo Chiat and the Paya Lebar *one*. [OHI-000374-LLL]
 - d. Oh it's the Romeo and Juliet *one*. [ICE-SG:S1A-093#345:1:C]

8.4.3 Frequencies of *one* vs. social variables

Similar to *already* and *also*, the frequency of clause-final *one* identifies ethnicity (see Section 8.1.9 and Section 8.2.3). As shown in Figure 8.34, the group of “Other” has the highest frequency of clause-final *one* on average (0.19 pp), followed by the Peranakan group (0.18 pp). However, as there are only two speakers in the group of “Other”, the number here may not be representative. The Chinese group attains the third highest frequency of clause-final *one* on average (0.17 pp). As with the other CSE expressions, there is a large variation in the frequency of clause-final *one* amongst the Chinese group ranging from 0.03 pp to 0.38 pp. Moreover, the highest frequency of clause-final *one* occurs in the Chinese group: speaker LAS (accession number 001953) attains the highest frequency of *one* (1.21 pp).⁷⁴ The Malay group ranks the fourth on average (0.15 pp), followed by the Indian group (0.12 pp). Again, the British and the Eurasian groups manifest the lowest frequency of clause-final *one*.

⁷⁴ The datapoint for speaker LAS was regarded as an outlier. It was counted in the average frequencies but is not shown in Figure 8.34.

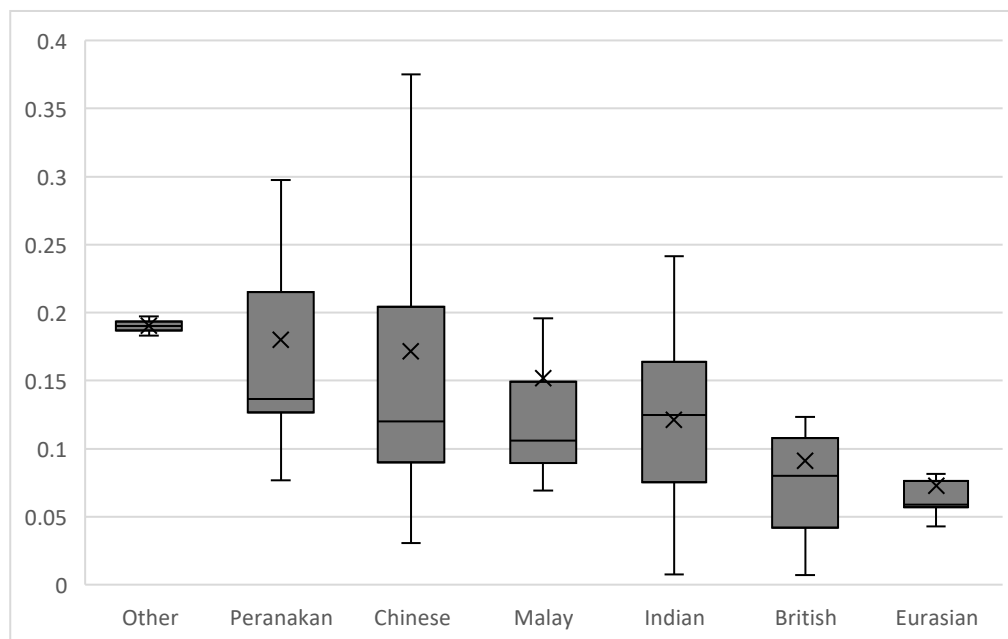


Figure 8.34: Frequency of clause-final *one* in different ethnic communities in OHI

Figure 8.35 provides the distribution of the frequencies of clause-final *one* and substrate-influenced clause-final *one* according to speakers' educational level. The result suggests a negative correlation between educational level and the frequencies of *one* in both categories, i.e. the frequencies of clause-final *one* and clause-final *one* with substrate-influenced functions descend when educational level moves from low to high. Speakers with a high educational level use *one* infrequently in clause-final position (0.10 pp). The same applies to clause-final *one* with substrate-influenced functions: speakers with a high educational level have the lowest frequency of *one* in this category (0.0065 pp). On the other hand, speakers with a medium educational level generally use a higher number of *one* in clause-final position (0.17 pp) and *one* with substrate-influenced features (0.019 pp). Finally, speakers with a low educational level produce the highest frequencies of clause-final *one* (0.23 pp) and clause-final *one* with substrate-influenced functions (0.057) on average.

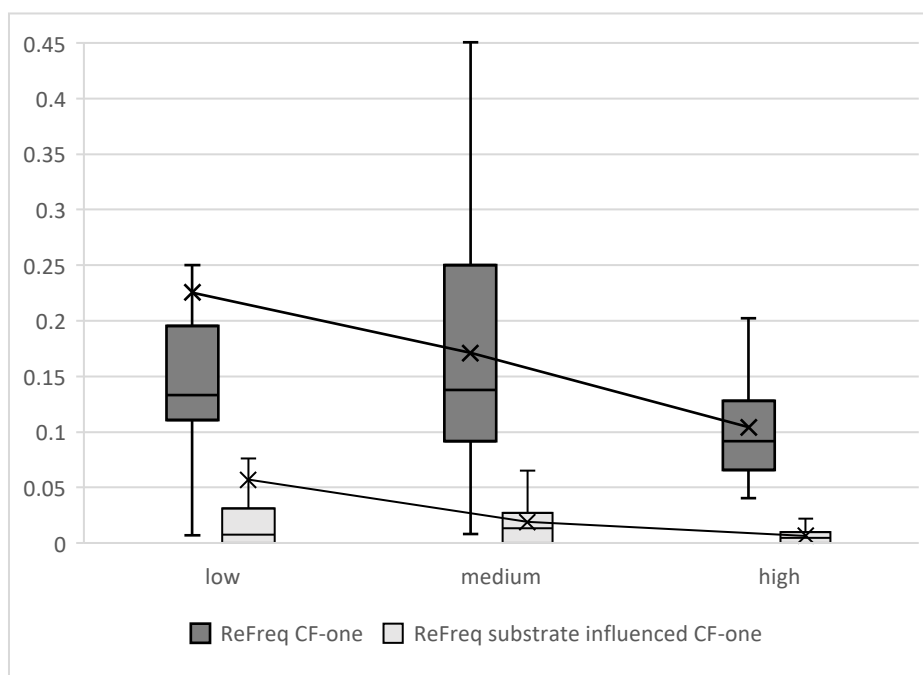


Figure 8.35: Frequencies of CF-one and substrate-influenced CF-one according to educational level

Apart from ethnicity and educational level, the impact of gender and age group on the use of *one* is less clear. As shown in Figure 8.36, female speakers tend to use clause-final *one* more frequently than their male counterparts. However, female speakers also have a wider range of variation in the frequency of clause-final *one*. Apart from that, there is no significant difference in the frequency of substrate-influenced *one* in clause-final position between the two groups differentiated by gender. Neither are there any apparent differences among speakers of different age groups, as suggested in Figure 8.37.⁷⁵

⁷⁵ Speakers born after 1940 have a higher frequency of clause-final *one* and clause-final *one* with substrate-influenced features on average due to an outlier that appearing outside the graph (speaker LAS, CF-one: 1.2 pp, CF-one with substrate-influenced features: 0.51 pp).

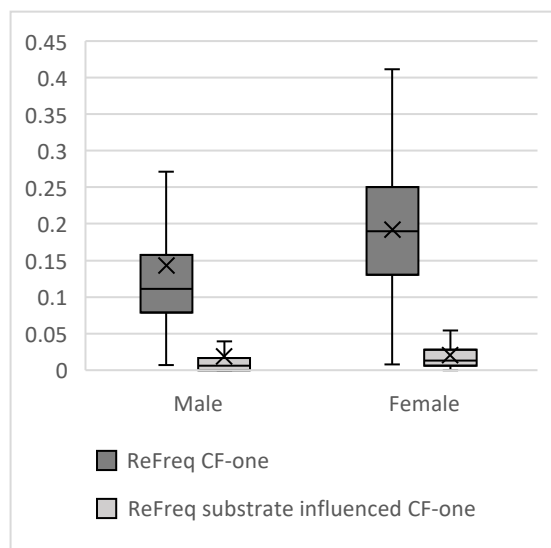


Figure 8.36: Frequencies of overall and clause-final *one* pp according to gender

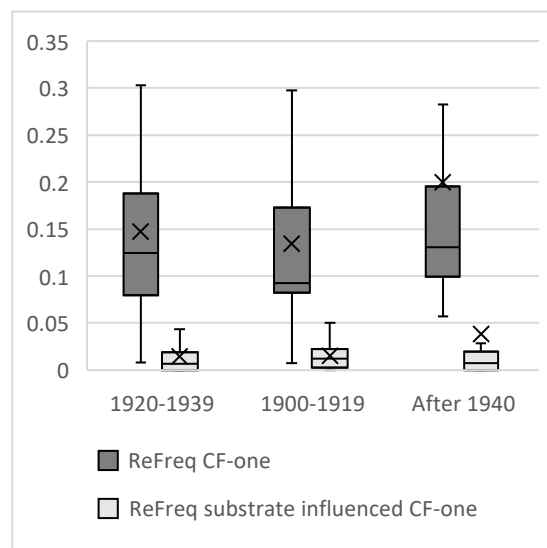


Figure 8.37: Frequencies of overall and clause-final *one* pp according to age group

Multiple linear regression analysis

Again, a multiple linear regression analysis was performed to calculate the impact of ethnicity, education, age group, and gender on the frequencies of clause-final *one* and substrate-influenced *one*. The results confirm the above observation: education level and ethnicity are the best predictors of the frequency of clause-final *one*: $F(2, 97)=7.860$, $p=0.000686$, $R^2=0.139$, as shown in Table 8.7:

	Estimate	Std. Error	t	Sig.
(Intercept)	0.354	0.053	6.628	1.9231e-9
Education	-0.070	0.021	-3.395	0.000997
Ethnicity	-0.018	0.007	-2.537	0.012761

R-Square: 0.139

F-statistic: 7.860 on 2 and 97, p-value: 0.000686

Table 8.7: Multiple linear model on relative frequency of CF-*one*

On the other hand, only educational level is the best predictor of clause-final *one* with substrate-influenced features: $F(1, 98)=7.322$, $p=0.008035$, $R^2=0.070$ (see Table 8.8). Except for educational level, the other variables did not add statistically significantly to the

prediction (ethnicity: $t=-1.620$, $p=0.108394$; gender: $t=0.482$, $p=0.631$; age group: $t=1.817$, $p=0.072$).

	Estimate	Std. Error	t	Sig.
(Intercept)	0.067	0.018	3.627	0.000457
Education	-0.021	0.008	-2.706	0.008035

R-Square: 0.070

F-statistic: 7.322 on 1 and 98, p-value: 0.008035

Table 8.8: Multiple linear model on relative frequency of *CF-one* with substrate-influenced features

8.5 Some other findings

Interestingly, the frequency of *already* in general is higher in the second half of the interview than in the first half in OHI.⁷⁶ On average, the absolute frequency of *already* is 16.87 in the first half and 20.01 in the second half on average for the 100 interviews. A paired t-test was performed to examine whether the means of the two results are significantly different. The result suggests that the interviewees in OHI produced statistically significantly higher number of *already* during the second half than the first half of their interviews ($t=3.01$, $p=0.00166$).

Figure 8.38 shows the frequency difference between the number of *already* in the first half and in the second half of the interview amongst all speakers. It mirrors that in Figure 8.39, which includes only the Chinese speakers. However, when we exclude the Chinese group, the difference between the first and the second half of the interview among speakers of other ethnic backgrounds is not significant, as shown in Figure 8.40. Therefore, we can conclude that mainly the Chinese speakers contribute to the differences here.

⁷⁶ The result was achieved by dividing the interviews in half. The middle page of each interview was used as a middle point.

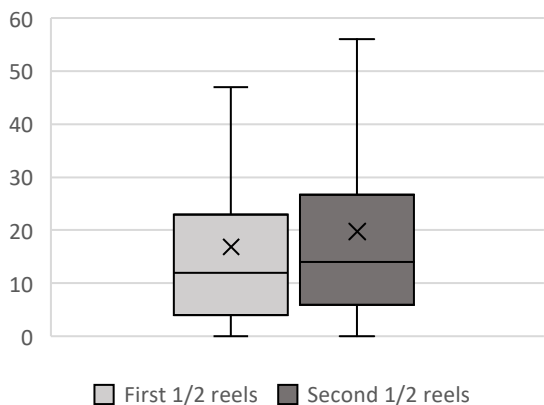


Figure 8.38: Frequency of *already* in 1st half vs. 2nd half of interview (all speakers)

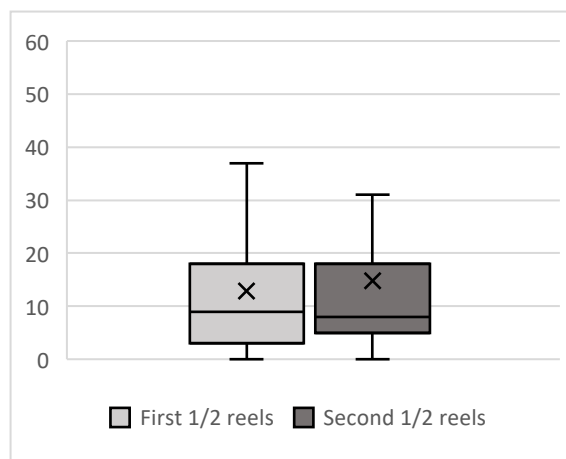


Figure 8.40: Frequency of *already* in 1st half vs. 2nd half of interview (all other groups)

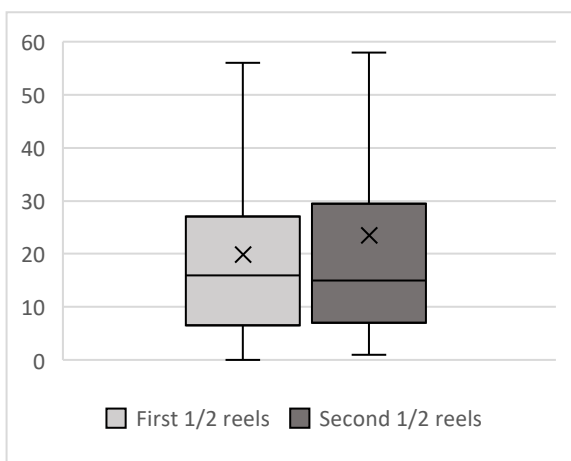


Figure 8.39: Frequency of *already* in 1st half vs. 2nd half of interview (Chinese group)

The interview setting of OHI is rather official, as they mostly took place in the National Archives of Singapore. Such an official setting may have encouraged interviewees to speak a variety of English that is as close as possible to the standard norm in the beginning of the interview. We assume that as the interviewees got more familiar with the interviewer and the environment, they started to feel more relaxed. Hence, the register of the interview shifted gradually from relatively more formal to less formal.

8.6 Summary and conclusion

The data surveyed in the above sections reveal – in accordance with Bao and Hong’s (2006) study – that the frequencies of *already* and *also* are generally higher than in Standard English. Apart from that, there are register and stylistic variations: the frequencies of *already*, *also*, and *one* with substrate-influenced features – either phrase-final or with novel grammatical meanings – are higher in the private dialogues than in the other text types, including the public dialogues and monologues (both scripted and unscripted). Contrary to our expectation, the frequency of *ever* is much lower in CSE (as represented by both OHI and ICE-SG) than Standard English. Compared with the distribution of *ever* in Ziegeler (2015), which represented an even more recent type of CSE, the result suggests an increase in the distribution of experiential *ever*. We can therefore conclude that *ever* has gradually shifted from its standard usage meaning ‘any time’ in a negative sentence to an aspectual marker expressing the experiential aspect meaning ‘at least once’ in the indefinite past in both affirmative and negative sentences.

The results in this chapter also suggest that substrate-influenced grammatical functions and syntactic structures have been a stable feature of the four CSE variables. In OHI and ICE-SG, both standard and substrate-influenced usages were identified among these four expressions. For example, *already* has been used as an aspectual marker that gives rise to the completive, inchoative, and prospective reading, but it has maintained its usage as a PhP expression in both corpora. The double use of *already* is another case in point, in which the antecedent *already* follows the auxiliary as in Standard English, indicating that an event occurs earlier than expected, while the sentence-final *already* functions as an aspectual marker. Apart from that, the functional distributions of these CSE expressions across the two corpora are not particularly divergent. There is also no apparent evidence showing a decrease in the frequencies of these CSE markers. In addition, we can observe an ongoing functional extension of these four CSE markers, as they are becoming increasingly unrestricted in their functional range and syntactic position.

As far as the social variables are concerned, the most important factors in determining the frequencies of the substrate-influenced tokens are ethnicity and educational level. There are also differences in the frequencies of these expressions between male and female speakers, as well as across different age groups, but these differences are not as significant.

Overall, the Chinese group in OHI attain generally higher frequencies of *already*, *also*, and *one* but a lower frequency of *ever* than speakers of other ethnic backgrounds. They also use a higher number of the first three CSE markers in phrase-final position. Different from the Chinese group, the British and the Eurasian groups have the lowest frequency of *already*, *also* and *one*, but the highest frequency of *ever*. Besides, hardly any substrate-influenced tokens were found in these two groups. The Peranakan group, conventionally subsumed under the Chinese group (see Chapter 3.2.3), produce lower frequencies of *already*, *also*, and *one* than the Chinese group, but they produce higher frequencies of these CSE expressions than the Malay and the Indian groups. On the other hand, the frequency of *ever* is higher in the Peranakan group than in the Chinese group, followed by the Malay and the Indian speakers.

There is a negative correlation between the frequencies of *already* (overall, sentence-final position, and aspectual) and the level of education. The same applies to the frequencies of *also* and *one*. However, there is a positive correlation between the frequency of *ever* and the level of education. The results, to a certain extent, support Platt (1975), which proposed a post-creole continuum subject to social stratification (see Chapter 4.7.1). However, the data also suggest a wider variation amongst speakers with a low or a medium level of education. The next chapter discusses the implications of the above findings and answers the research questions outlined in Chapter 2.

9 Discussion

This chapter examines the results with respect to the research questions outlined in Chapter 2. Section 9.1 first addresses the question whether CSE is diachronically stable. Section 9.2 explores the possibility of a replica grammaticalization of *already*, *also*, *ever*, and *one* in CSE in relation to the Chinese as. Section 9.3 elaborates on the sociolinguistic findings of this study, namely the correlation between frequencies of the CSE expressions and two social factors – ethnicity and educational level, before discussing further theoretical implications with regards to the previous CSE models. Section 9.4 summarizes this chapter.

9.1 The emergence of CSE and its stabilization

The study discusses the questions concerning the developmental trend of CSE during the last five decades as well as whether the OHI data provide empirical evidence to support the Dynamic Model (Schneider 2007, also see Chapter 2.1). The research questions outlined in Chapter 2 were (i) whether the uses of the prominent markers of CSE (*already*, *also*, *ever* and *one*) have increased over the last five decades, and (ii) whether these markers of CSE obeyed usage principles and distributions different from today.

Stable development of CSE

Within Schneider's (2007) Dynamic Model, CSE was considered to be in stage 2 (exonormative stabilization) during 1867–1942 and stage 3 (nativization) between 1945 and the 1970s (see Chapter 7.1.2). Within these two stages, it was assumed that there is a shift from standard codes of English serving as the norm to the nativization of innovative linguistic expressions introduced by the local communities. Currently, CSE is assumed to have reached stage 4 (endonormative stabilization) and potentially stage 5 (differentiation) (Schneider 2007:155–61). In these current stages, CSE is believed to be characterized by the

change of orientation towards the local norm and the acceptance of a local model of English rather than looking towards the British norm.

As the OHI data represents CSE in the early 20th century, it can be assumed to belong to stage 2 (exonormative stabilization) and stage 3 (nativization) in the Dynamic Model. On the other hand, ICE-SG can be subsumed under stage 4 (endonormative stabilization) and potentially stage 5 (differentiation). However, the results of this study suggest that CSE has been in this phase for an extensive period of time, apparently undergoing no substantial changes. Neither are there any suggestions of a strong shift away from substrate-influences towards Standard English. These two conclusions are evident in that the year of birth is not significantly correlated to the frequencies of the four CSE markers. Apart from that, there has been substantial individual variation across all speakers, irrespective of their year of birth. However, although there is no significant correlation between the frequencies of the CSE markers and the year of birth of the interviewees, speakers who were born after 1940 produce slightly higher frequencies of *already*, *also*, and *one* overall and these expressions with substrate-influenced features (e.g. clause-final *already*, *also*, and *one*). High frequencies of substrate-derived features generally point to the more basilectal speakers, and vice versa. However, it is important to note that the number of speakers is not equally distributed among the three age groups, as there are fewer interviews with younger speakers sampled in OHI – the average birth year of the selected speakers being 1928 (see Chapter 7.1.2).

Frequency differences among the four CSE expressions

Among all linguistic variables explored in this study, there are considerable differences in the frequency of their occurrences. CSE, as represented by both OHI and ICE-SG in this study, is marked by a higher frequency of clause-final *already*, *also*, and *one* but a lower frequency of *ever* than in Standard English. Apart from that, these CSE expressions have shown a different pace of functional extension. For example, more than half of the tokens of *already* found in OHI and ICE-SG can be argued to represent substrate-influenced features, e.g. as an aspectual marker expressing the completive, inchoative and perspective meanings

(see Chapter 8.1.3). *Already* has also extended its functional range to negative polarity contexts. It is used with negators such as *not*, *no*, or *never* to denote the discontinuity of a habit or earlier state, substituting another phasal polarity expression – *no longer* (see Chapter 8.1.6). *Already* has developed other substrate-influenced features, such as (i) double occurrences in one single clause, (ii) assuming the function of the copula *be*, (iii) co-occurring with natural development predicates, (iv) occurring between a verb and its object, (v) co-occurring with *still*, and (vi) developing pragmatic functions (see Chapter 8.1.7 and Chapter 8.1.8).

In comparison with *already*, other CSE markers have manifested lower proportions of substrate-influenced meanings and non-standard syntactic positions. For example, only 22.4% of *also*-tokens were used in clause-final position in OHI, and among these, 19.8% have developed novel grammatical meanings which are parallel to its counterparts in the Chinese substrates. ICE-SG shows an even smaller percentage of clause-final (12.0%) and substrate-influenced *also* (13.4%). Similar patterns can be observed in the use of *one*, as only 11.9% of *one*-tokens in phrase-final position show non-standard features in OHI. The ratio of *one*-occurrences of the same category in ICE-SG is 20.9%. *Ever* manifests the lowest proportion of tokens with novel grammatical meanings. Thus, the results suggest that the frequencies of *already*, *also*, and *one* serve as good indicators of substrate and/or basilectal influences, while *ever* is less susceptible to substratum interference.

The distributional differences between the former three expressions and *ever* may be subject to the so-called “frequency effect” (see Ellis 2002; Terassa 2017). It has been observed that high-frequency linguistic items typically adopt marked usages and complementary patterns in language contact (Schneider 2007:46). Frequency is also a bridging variable that binds the scholarly circles of language acquisition research, sociolinguistic variation and language change (Ellis 2002:143). In SLA studies, frequency has been proven as a key determinant because it is the learners’ lifetime analysis of the distributional characteristics of the language input that forms the analysis (from phonology, through syntax, to discourse) of language “rules” (Ellis 2002:144). Diessel (2007:109) also argues that usage frequencies fundamentally influence language structure and use. Bybee

(2007:8) adds that frequency or repetition is important for language development and change because mental representations of language are influenced by repetition. Although less has been researched concerning the effect of usage frequency on substratum transfer,⁷⁷ it is hypothesized that grammatical nativization of postcolonial English typically starts with high frequency grammatical innovations, which are intuitively more acceptable than low frequency grammatical items, and thus become firmly rooted (Schneider 2007:46).

Register, formality, and genre as important factors of frequency variation

The results also show that the register, formality, and genre of the corpora are important factors influencing the distributions of the CSE markers. It is evident that the percentages of all the CSE variables are significantly higher in the private dialogues than the other text types, i.e. public dialogues and monologues in ICE-SG. In other words, there are consistent decreases in use of these CSE markers when the text types move from less formal to more formal ones (private dialogues vs. public dialogues; dialogues vs. monologues), as well as from unprepared to prepared speeches (unscripted monologues vs. scripted monologues). Apart from that, the average frequencies of these CSE markers in OHI are much lower than those in the private dialogues but slightly higher than the monologues in ICE-SG. It is worth noting that the main genre in OHI is narrative, as speakers recounted their personal recollections during the course of their lives (see Biber 1992 on conversation text types). It seems plausible to conclude that the OHI data represents a type of CSE more formal than the private dialogues but less formal than the public dialogues and monologues in ICE-SG. On the other hand, in the private dialogues in the ICE-SG corpus, there are various discussions and conversations, with perhaps less chance for narratives than in OHI.

These register differences are also visible in OHI, as speakers tend to use a higher number of *already* when they became more relaxed and familiar with the interview settings. Moreover, the frequencies of the non-standard variations with extended functional and

⁷⁷ Terassa (2018) studies the substratum transfer functioning as a constraint on frequency effect, but not the other way around.

structural range are contingent on speakers' ethnic background and level of education, as demonstrated in OHI.

In sum, CSE, as represented by OHI and ICE-SG, shows consistent stylistic and internal sociolinguistic variation. As such, this study provides evidence that stylistic as well as other sociolinguistic variation may be present from very early on in the development of CSE.

These findings are noteworthy as the emergence of internal differentiation would not have been expected until stage 5 (differentiation). Schneider (2003, 2007) hypothesizes that stylistic differentiation is theoretically expected in phase 5 (differentiation) of the Dynamic Model. In Schneider's words, phase 5 "marks the onset of a vigorous phase of new or increased, sociolinguistically meaningful internal diversification" (Schneider 2003:254). In his later monograph, Schneider (2007) specifies that in phase 5, "differences within a society and between individuals with respect to their economic status, social categories, and personal predilections come to light and can be given greater prominence" (Schneider 2007:53). More importantly, it is the internal differentiation that marks the turning point of "the stage of dialect birth" (Schneider 2007:54):

New varieties of the formerly new variety emerge, as carriers of new group identities within the overall community: regional and social dialects, linguistic markers (accents, lexical expression, and structural patterns) which carry a diagnostic function only within the new country emerge.

(Schneider 2007:54)

The emphasis on these sociolinguistic factors is not new to studies of new varieties of English. Huber (2014), for example, discovers consistent stylistic and gender-related differentiation in Ghanaian English by comparing data from sociolinguistic interviews and from the Ghanaian component of ICE. There, he found a consistent decreasing rate of the zero relativizer (instead of using e.g. *that* and *who*) with increasing formality of the text type. Such a preference for zero relativizer in informal texts parallels British English. Yet, the distribution of relative *that* in Ghanaian English shows no significant grading across the stylistic continuum (Huber 2014:104). Apart from that, Ghanaian English has created genuinely new sociolinguistic variables, e.g. the affricated variant of /t/ and the social

prestige that attaches to it. Yet, Ghanaian English was assumed to only have arrived at stage 3 according to Schneider's Model. Although the Dynamic Model does allow for some sociolinguistic variability as early as stage 3, differences were anticipated as social class variation, with higher social classes of the local communities approaching Standard English (Schneider 2007:44–45). However, the stylistic and gender-related variation found within the group of educated Ghanaian speakers was apparently not caused by unequal access to English due to class-based reasons (Huber 2014:91). Huber (2014) also predicts that such internal variation is unlikely to disappear in the near future, but appears to be a fundamental characteristic of a developing New English.

Wee (2014), arguing for stage 5 (differentiation) of today's CSE, underlines the importance of the sociolinguistic factors, i.e. the understanding of community, identity, and culture in a rapidly globalizing world. He also suggests that there is a need to reconsider these concepts given recent theoretical changes of these in late modernity, with constant inward and outward migration as well as commodification which lends a different dynamic to language practices (Wee 2014:128). For example, the presence of CSE in social media such as YouTube videos and Singaporean movies, a sign of CSE becoming increasingly commodified, can influence the relationship between language and identity. Thus, the impetus for retaining and expanding the use of a language variety becomes less conditioned on solidarity considerations.

Similar findings have been presented in more recent studies on CSE. For example, Buschfeld (2021) finds that there are different CSE varieties spoken by ethnically Chinese and Indian children. Cavalloro and Ng (2021) suggest that the changes in the Singaporean linguistic landscape have different degrees of impact on the three major ethnic groups.

In sum, the findings of this study suggest that CSE has remained remarkably stable, with a slight increase in the use of the grammatical variables showing substrate-influenced features. CSE has been in the stages of nativization and/or endonormative stabilization for an extensive period of time, apparently undergoing no substantial changes. Apart from that, the degree of nativization and stabilization crucially depends on the individual speaker, contingent on their different ethnic background and level of education. Lim (2007) suggested

that CSE did not vary significantly across different ethnic groups in the past. Schneider (2007) did not expect meaningful sociolinguistic diversification amongst different ethnic strands of speakers until stage 5 (differentiation). However, it is evident that internal differentiation, formerly believed to be a characteristic of fully-developed new varieties, has to be considered from earlier stages of an emerging variety of English, perhaps even tracing back to its genesis. The findings suggest that further studies on New Englishes – both synchronically and diachronically – should not neglect or generalize away these social variables of the individual speakers.

9.2 CSE in relation to its Chinese substrates

The second research question focuses on the CSE expressions *already*, *also*, *ever*, and *one*. This study explores to what extent the usages of *already*, *also*, *ever*, and *one* are related to their Chinese substrates.

To a large extent, Chinese provides the extended semantic meanings and syntactic features, which are “merged” with the already existing usages in English (see Ziegler 2015:253 on “Merger Constructions” and Matras 2009:240–42 on “Pivot Matching”). All the extended grammatical functions, which are not compatible with Standard English have found their equivalents in the Chinese substrates. For example, different from Standard English *already* as a phasal polarity expression (PhP) typically enriched by the connotation of unexpectedness and anteriority (see Chapter 5.1.2), CSE *already* has acquired additional aspectual functions, marking the completive, inchoative and prospective aspects, resembling the Chinese substrates (see Chapter 5.1 and Chapter 8.1). It is also used in negative contexts, which is atypical in Standard English. In addition, it has developed novel usages such as occurring twice in the same clause, with the preverbal/postverbal *already* serving as a PhP expression, and the clause-final *already* as an aspectual marker (see Chapter 8.1.7). *Also* has acquired subtle grammatical meanings when used with universal quantifiers (e.g. *all*, *everything*, *everyone*, etc.) and concessive *even*, modeled on the Chinese counterpart 也/都 *yě/dōu*. Unlike Standard English *also*, which is restricted to affirmative contexts, CSE *also*

extends its use to negative contexts, again mirroring that of its Chinese substrate equivalent 也/都 *yě/dōu*. *Ever* is used to express existential meaning in negative contexts in Standard English, whereas CSE *ever* is an experiential aspectual marker which means ‘to experience at least once’. Chapter 5.3 argued that these novel usages are influenced by Chinese 过 *guò* (see also the results and analysis of *ever* in ICE-SG and OHI in Chapter 8.3). Last but not least, the pronominal *one* in Standard English is limited to replacing a general, vague noun phrase, or a noun phrase that was already mentioned in a previous context to avoid repetition (see Chapter 5.4.1). However, CSE *one* has extended its usages in CSE as a nominalizer and as an emphatic marker (see Chapter 5.4.2 and Chapter 5.4.3, as well as the results and analysis of the corpus data in Chapter 8.4).

In some cases, the influence from Chinese in the CSE expressions is so strong that the entire CSE sentence no longer meets the surface structural requirements of English. The results here seem to challenge Bao’s (2015) lexifier filter (see Chapter 4.7.4). For example, *already* was found occurring between the verb and its object, e.g. *I engaged already two people, I stopped already working*, just like *V-le* in Chinese. Here, *already* adheres to both the morphosyntax and semantic functions of the Chinese substrate, at the cost of violating the lexifier morphosyntax. However, similar examples do not occur very frequently, and they appear less frequently in the more recent corpus of CSE – ICE-SG – than in the older type of CSE represented by the OHI data (see Chapter 8.1.8).

The findings of the strong Chinese influence on the CSE expressions and the stability of CSE are not very surprising. First of all, from a demographic perspective, the Chinese became the largest ethnic group, which surpassed the Malay population in the 1830s (see Table 3.1 and Figure 3.2 in Chapter 3). By the 1930s, the Chinese grew to make up 75% of the total population, and that proportion has remained more or less unchanged to this day (see Chapter 3). As explored in Chapter 4.8, one of the most important external factors in contact language formation is the size of the communities, along with types of intra- and intersocietal network, multilingual practices and language policies (see Ansaldo 2009, 2019; Lim 2009). In addition, as revealed in Chapter 3, Mandarin Chinese started to gain significance in the linguistic ecology of Singapore in the 1910s (after the Chinese revolution

in 1911 in China), much earlier than many have estimated (Ansaldo 2004; Gupta 2001; Leimgruber 2013; Lim et al. 2010). This also provides evidence for the argument that the emphasis on English and Mandarin Chinese in the language-related policies, established in Singapore in the 1960s, did not appear suddenly upon its independence, but have their historical roots dating back to the precolonial and colonial periods. It appears plausible to attribute the stability of CSE and its high resemblance to its Chinese substrates – both semantically and syntactically – to the stable composition of the population of the Chinese ethnic group since the 1930s and the emphasis of teaching and learning Mandarin Chinese since the 1910s.

Apart from that, “parallel constructions” (Teo 2020) between CSE and the substrate languages appear to be a strong motivation of cross-linguistic influence. Aspectual *already* and *already* in clause-final position are used most frequently by Chinese and Peranakan speakers, probably due to the influence of the parallel constructions of *le* in Chinese. Following the Chinese and Peranakan Speakers, the Malay group tend to use aspectual *already* more frequently than speakers from other ethnic backgrounds. The findings confirm the assumption that Malay *sudah* may also be reconstructed for aspectual *already*, though to a smaller extent than Chinese, as Malay *sudah*, unlike Chinese, rarely occurs sentence-finally or in negative sentences (see Chapter 5.1.1 and Chapter 8.1.8).

Moreover, the tendency of using aspectual *already* and clause-final *already* decreases with increasing level of education (see Chapter 8.1.9). Similar patterns apply to the use of clause-final *also* and clause-final *one*. For example, speakers with Indian, Peranakan, Malay and Chinese ethnic background use *also* and clause-final *also* more frequently than other speakers. However, compared to the level of their education, the factor ethnicity has a less explanatory power in predicting the frequencies of *also* (see Chapter 8.2.3).

9.2.1 Contact-induced grammaticalization

Within the framework of contact-induced grammaticalization, Heine and Kuteva (2003; 2005) propose two types of grammaticalization process due to cross-linguistic influences: (i) ordinary grammaticalization, and (ii) replica grammaticalization (see Chapter 4.4). They differ in whether speakers replicate a grammaticalization model into the replica language that exists in the model language. If yes, it follows the process of replica grammaticalization, and if not, it belongs to the process of ordinary grammaticalization.

Ordinary grammaticalization is based on a strategy that a speaker uses to transfer a grammatical concept from the model language to the replica language. This type of grammaticalization is applicable to all four CSE expressions in this study. Let us take CSE *already* as an example. CSE speakers notice that there is a grammatical category (aspectual marker *le*) in Chinese. They use the available material (PhP *already* in Standard English) to develop an equivalent category (aspectual *already*). To that end, they grammaticalize *already* from a phasal polarity expression to an aspectual marker. As such, ordinary grammaticalization applies to contact-induced grammaticalization of *already* from a phasal polarity expression to an aspectual marker in CSE. However, the account of ordinary grammaticalization does not offer any explanation as to why speakers choose one conceptual source item in the lexifier language over another in the first place (Ziegler 2015:123).

Bao (2015), on the other hand, proposes that CSE markers were first introduced as a result of transfer from the substrate languages. As a sequential step, these usages became stabilized by frequent and common use in the local communities. By integrating the theory from second language acquisition by Siegel (2008), Bao (2015) argues for post-transfer stabilization of the CSE markers. Bao (2015:4) describes the process of transfer or substratum transfer as the process in the mind of “the erstwhile creator-developer of the contact language” (i.e. the speakers of the contact language) by which the morphosyntactic resources of one language (the lexifier) are used to represent the grammatical construction of another (the substratum). Bao’s study (2015) compares the distribution of substrate-derived *one* between two corpora: ICE-SG and the Singapore Corpus of Research in

Pedagogy (SCoRE), which recorded classroom discourse in Singaporean school in the mid-2000s. Bao (2015) focuses on the section of English-lessons in SCoRE, which amounts to 1.2 million words. The result shows that the ratios of relativizer *one* and emphatic *one* are approx. 15% lower in SCoRE than ICE-SG. According to Bao (2015:27), the decline of substrate-derived *one* supports the transfer analysis, in which the lexifier filter inhibits a further development of the Chinese-derived frames of *one*. However, it poses a challenge to the grammaticalization analysis, as one would expect that the frequency of substrate-derived *one* increases as the substrate grammaticalization deepens in the contact language.

However, the two corpora in Bao's (2015) study are very different in nature. ICE-SG consists of a series of register-based sub-corpora, including the private dialogues which recorded informal conversations amongst students outside of classrooms. On the other hand, SCoRE compiled classroom dialogues between teachers and students, which are closer to the text type of public dialogues in ICE-SG. Therefore, the result there does not show a real decline in substrate-derived *one* but manifests a register difference across the two separate corpora.

In contrast, the findings in this study show an increase of the substrate-derived variables over time by comparing OHI with ICE-SG. The results do not clearly support Bao's (2015) idea of an inhibition of further grammaticalization of the substrate-influenced CSE markers. On the other hand, the findings support the grammaticalization analysis, as the CSE expressions become increasingly unrestricted in their functional range and syntactic position (see Chapter 8). According to Himmelmann (2004), semantic extension is a necessary component of grammaticalization. Clearly, *already*, *also*, *ever* and *one* in CSE are undergoing semantic extension, and at least some of their substrate-influenced usages are on the path towards grammaticalization.

On the other hand, replica grammaticalization suggests that grammaticalization in the replica language is the result of replicating a process of grammaticalization that has taken place in the model language. This type of grammaticalization has been attested in the so-called "hot-news perfect" in Irish English (see McCawley 1971; Harris 1991; Filppulla

1999; Pietsch 2009). It explains the grammaticalization of *after* from a locative preposition to an aspectual marker in Irish English, as exemplified in (216):

- (216) a. Irish: Tá sí tréis an bád a dhíol.
 be:NON-PAST she after the boat selling
- b. Irish English: She's after selling the boat.
 'She has just sold the boat.' (Heine and Kuteva 2003:540)

In (216)a, the locative preposition *tréis* ('after') is used as an aspectual marker. It shows that *tréis* has been grammaticalized from a locative preposition to an aspectual marker, which is used to indicate that something occurred in the recent or immediate past, relative to the time of speaking or reference. Irish speakers apply the same grammaticalization mechanism to *after* in English, giving *after* the same aspectual meaning of hot-news perfect. By doing so, they replicate the grammaticalization process that has taken place in Irish. The case for Irish English hot-news perfect is strong, as researchers have shown historical trajectories of the hot-news perfect in Irish and subsequently in Irish English (Fulppula 1999; Pietsch 2009).

In order to examine whether CSE markers replicate a grammaticalization process that has taken place in the Chinese substrate, the following section further discusses the grammaticalization process of the CSE expressions in relation to their Chinese counterparts.

9.2.2 Grammaticalization of *already* in relation to Chinese *le*

As discussed earlier, we observe the grammaticalization of *already* from a phasal polarity expression to an aspectual marker in CSE, which is exemplified in (217). In both examples, the verb is *to change*, and both express the aspectual meaning of current relevance, which indicates a link between the present and the past. However, it is the use of *already* that serves as a perfect marker in the CSE example (217)a, as both the explicit perfect marking (*have* in English) and the inflectional morpheme *-ed* are lacking. In the Standard English example (217)b, the present perfect form *has changed* links the present and the past, and the use of *already* involves two reference points situated before and after a phasal change. We can see that the use of *already* in Standard English typically enriches the present perfect with

connotations of unexpectedness and anteriority (see Chapter 5.1.2). A reduction of these connotations and an addition of the grammatical function of perfect marking are visible in the CSE example.

- (217) a. CSE *already*: 1937 and then after that, name *already* change.
 ‘The name has changed.’ [OHI-E000284-CNS]
- b. PhP *already*: And the results of that clash will change me, have *already*
 changed me tremendously.
 [COHA-1937-FIC-EnemyGods]

The grammaticalization of *already* in CSE could be schematically represented using the following chain:

[*already*: PhP expression] > [*already*: ASP marker]

Here we assume that the PhP meaning of *already* is basic, but the grammatical function as an aspectual marker is acquired through the process of replica grammaticalization. A presumed grammaticalization of the Chinese *le* should be identical with the above chain:

[*le*: PhP expression] > [*le*: ASP marker]

However, as discussed in Chapter 6.2, the perfective *le* is directly derived from the verbal usage *liǎo* in the sense of ‘to complete’. It has possibly undergone phonological reduction from a triphthong to a schwa [lə] (see Sun 1996:88), which became today’s *le* as a perfective marker. The grammaticalization process could be condensed in the following schema:

[*liǎo*: verb meaning ‘to complete/finish’] > [*le*: perfective/completive marker]

The nature of this process can be illustrated with two examples from Chinese, one from Middle Chinese, and one from modern Mandarin Chinese (see Chapter 6.2.5 and 6.2.6). Consider the sentences in (218):

- (218) a. 官 事 未 易 了 也
 guān shì wèi yì liǎo yě
 official matter NEG easy complete PTC
 ‘The government business is no easy to finish.’
 (From *Jinshu Fuxian zhuan*, cited in Sun 1996:88)
- b. 我 吃 了 饭 了
 wǒ chī le fàn le
 I eat ASP meal ASP
 ‘I have eaten.’

(218)a is an instance of a verbal use of the Chinese *liǎo* in Middle Chinese, and (218)b is an instance of *le* as an aspectual marker in modern Mandarin Chinese. We notice that *le* occurs twice in (218)b, one directly after the actual verb 吃 *chī* ‘to eat’, and the other sentence-finally. As discussed in Chapter 6.2.1, the antecedent *le* is close to the interpretation of a perfective/completive aspectual marker (Li and Thompson 1981), and the sentence-final *le* belongs to the interpretation of marking a current relevant state (CRS) (Comrie 1976; Li and Thompson 1981).

The question remains whether Chinese also expressed the concept of PhP by means of *liǎo/le* at a point before the verbal *liǎo* grammaticalized into the aspectual *le*. If yes, we could prove that the grammaticalization of *already* from a PhP expression to an aspectual marker does replicate the grammaticalization process in the model language. If not, the grammaticalization of the CSE marker *already* is only a case of ordinary contact-induced grammaticalization. A schematic representation is listed in Figure 9.1. The solid-lined arrow in CSE indicates a real course of development, the dashed arrow represents an assumed chain of grammaticalization of the Chinese *liǎo/le*.

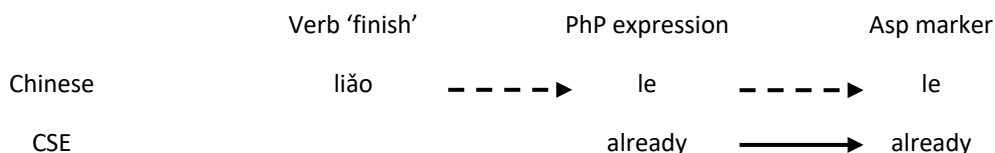


Figure 9.1: Grammaticalization of Chinese *le* and CSE *already*

Previous literature on Chinese *le* does not treat it as a PhP expression but as an aspectual marker (see for example Li and Thompson 1981), though both PhP expressions and aspectual markers are temporally bounded, as they are used to signal temporal conceptualizations, i.e. whether an event/action has been initiated, completed or terminated. The potential of *le* serving as a PhP expression has yet to be established. Inspired by the scenario of van der Auwera (1993:621), as in (219), we explore the possibility of the Chinese *le* serving as a PhP expression.

(219) Suppose you want to marry a certain woman. You propose and you find out that she is *already* married. There is nothing necessarily early about this marriage. You simply come too late to have a chance [...] (van der Auwera 1993:621)

(220) 我 想 娶 她 可是 她 结婚 了 (personal knowledge)
 wǒ xiǎng qǔ tā kěshì tā jiéhūn le
 I want marry her but she marry ASP
 ‘I want to marry her, but she is already married.’

Example (220) in Chinese is instantiated on the scenario in (219). Given the context, the meaning of *le* is ambiguous: It follows the predicate 结婚 *jiéhūn* ‘to marry’, where its grammatical function is to mark the perfective aspect, but the contextual frame suggests that the English *already* offers a more plausible interpretation of the utterance. The temporal interpretation of *le* is exactly the same as in (219) in which the reading of *already* involves two temporal points. The first temporal point is the change of state from “not married” to “married”, and the second one is the contrasting alternative held by the speaker which is placed at a later point than the first one.

Clearly, the context given here can be described in terms of “bridging contexts” (Evan and Wilkins 1998:5, cited in Heine 2002:84), which are crucial in semantic change. Similar terms have been proposed such as “critical context” (Diewald 1999), “inferences”, “implicatures” and “suggestions” (Grice 1967). The most important properties of bridging contexts are described in Heine (2002:84):

- (i) They trigger an inferential mechanism to the effect that, rather than the source meaning, there is another meaning, the target meaning, that offers a more plausible interpretation of the utterance concerned.
- (ii) While the target meaning is the one most likely to be inferred, it is still cancellable (see Grice 1967), that is, an interpretation in terms of the source meaning cannot be ruled out.
- (iii) A given linguistic form may be associated with a number of different bridging contexts.
- (iv) Bridging contexts may, but need not, give rise to conventional grammatical meanings.

The context of the Chinese *le* in (220) obviously meets all the above requirements. However, it seems to suggest that the interpretation of *le* as a PhP expression is developed from its aspectual meaning. Therefore, instead of following the chain of a semantic shift as suggested in the dashed line, a more plausible chain goes from *liǎo* as an actual verb, via *le* as an aspectual marker and finally arrives at the stage where it acquires the function of a PhP expression. Compare the assumed chain of the grammaticalization of *liǎo* with the more plausible one in Figure 9.2:

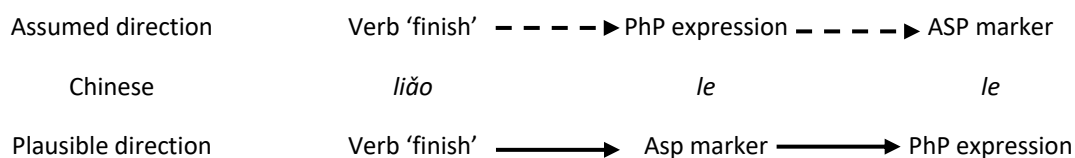


Figure 9.2: Grammaticalization of *liǎo*

According to Heine (2002), bridging contexts do not necessarily lead to grammaticalization. What is required in a semantic shift are “switch contexts” (Heine 2002:85), or “isolating contexts” (Diewald 1999), in which the interpretation of an old meaning is no longer possible, and the new meaning provides the only possible interpretation. I did not find many Chinese examples that give rise to a switch context, but interestingly, an example in Tai resembles a switch context, where its aspectual marker *léew*, borrowed from Chinese *liǎo* (see Bisang 1998:651), seems to have developed into a PhP expression. Consider the following examples:

- (221) a. Kháw khon-cà? paj júu léew.
 he must.have go CONT ASP
 ‘He must have gone already.’ (Noss 1964:165)

- b. Kháw (cà?) māj paj lɛɛw.
 he PROSP NEG go ASP
 ‘He is no longer going.’ (Boonyapatipark 1983:179)

As shown in (221), *lɛɛw* does not need to mark the perfective aspect as the verbal suffix *-cà?* already serves the function of marking the perfective aspect. The Tai *lɛɛw* in this context corresponds to the English *already* that requires two temporal reference points. Another interesting example from Tai shows that by negating *lɛɛw*, the utterance expresses the meaning of ‘no longer’, which is one of the PhP expressions described in Löbner (1989). According to Löbner (1989:72), the semantic concept of *no longer* can be achieved by internally negating *already* (also see Chapter 8.1.6). This provides further evidence for the argumentation that *lɛɛw* has acquired the function of a PhP expression.

Returning to Chinese, a clear-cut example has yet to be identified, which can illustrate a switch context in which *le* can only function as a PhP expression. However, some examples, such as the one in (222), show that *le* does not function as a perfective/completive marker, but is associated with a “contrary to expectation” interpretation, which belongs to one of the properties of the “current relevant state” of the perfect, as described by Li et al. (1982). They see *le* “as an exponent of the perfect aspect,” and *le* carries the basic discourse function of the perfect being, which “relates some state of affairs to the ‘current time’” (Li et al. 1982:22). Soh (2009) also argues that both the “change of state” and the “contrary to expectation” interpretations involve changes across temporal domain. Clearly, there is a certain overlap between the function of the perfect and that of the PhP adverbials, but still the aspectual marker *le* does not equal the PhP expression *already*. The key difference lies in that the perfect relates **a current state** with respect to **one particular reference time**, while PhP *already* involves **two temporal reference points**.

- (222) 这 个 木瓜 很 甜 了
 Zhè ge mùguā hěn tián le
 This CL papaya very sweet le
 ‘This papaya is very sweet.’
 (The speaker or the listener had thought or said it was not sweet)
 (Li et al. 1982:33)

In order to confirm the observation that the PhP meaning of *le* developed from its aspectual usage, a more detailed study is required to compare their occurrences in a historical timeline. However, it is beyond the scope of this current study and I shall not further discuss it here. It is worth noting, though, that it has been observed that what is expressed by a PhP expression in one language (see the examples in (221) in Tai) can be linked to aspect in a different language (Bisang 2002:656). And Classical Chinese seems to be a language that expressed the concept of inchoativity ‘already’, continuativity ‘still’, discontinuity ‘no longer’, and continuative negative ‘not yet’ within the framework of tense-aspect-modality (Bisang 2002).

Modern Chinese also expresses the concept of English *already* by means of the adverbial 已经 *yǐjīng* ‘already’, which marks inchoativity in the same way as English *already*. However, the inchoative phasal adverbial is a relatively new usage which did not exist in Classical Chinese (Bisang 2002:55). In fact, the modern Chinese adverbial *yǐjīng* ‘already’ is a result of a grammaticalization process. It started from 已 *yǐ* functioning as a verb with the meaning of ‘stop, halt; finish’, via the stage when it was used as a preverbal marker of a perfected action, to the final stage when it formed a disyllabic phasal adverbial with another character 经 *jīng* ‘to pass, to go’ (see Bisang 2002:55). Again, the grammaticalization of *yǐjīng* ‘already’ suggests that phasal adverbials developed from aspectual markers in Chinese.

9.2.3 Grammaticalization of *ever* in relation to Chinese *guò*

Like *already*, replica grammaticalization is not applicable to the case of CSE *ever* derived from the experiential marker 过 *guò* (see Chapter 5.3.2 and Chapter 6.3.1). The grammaticalization of *guò* could be schematically represented using the following chain (see Chapter 6.3.2):

[*guò*: verb meaning ‘to pass’] > [*guò*: aspectual marker in V+Asp (*guò*)+O]

Chapter 6.3.2 discussed that the verb *guò*, meaning ‘to pass a locative/temporal object’, first developed into a resultative marker, which forms an “action-result” construction with the preceding verb, e.g. 跳过 *tiàoguò* meaning ‘to jump across the pool’. At the last stage of the grammaticalization, *guò* built a stronger tie with the preceding verb, becoming a verbal suffix meaning ‘activity-experienced’, as exemplified in (123), listed in (223) again:

- (223) 饮 过 酒 今 番 不 忘
 yǐn guò jiǔ jīn fān bú wàng
 drink ASP wine this time NEG forget
 ‘Once we drink wine, we don’t forget this moment.’
 (*Zhu geliang bowing shaodun*, AD 1300)

We can interpret a sense of ‘at least once’ from the above example, i.e. the sentence can be literally translated as ‘Once we drink wine, we don’t forget this moment’. It is worth noting that the aspectual *guò*, as shown in (223), does not distinguish positive and negative contexts.

Experiential *ever* is modeled on aspectual *guò* in Chinese. Unlike Standard English *ever*, which is a negatively orientated polarity item, experiential *ever* does not require a negative context (see Chapter 5.3.2). Analogous to *already*, if we assume that the negative polarity of *ever* is basic, and the grammatical function of *ever* as an aspectual marker is acquired through replica grammaticalization, then the grammaticalization of CSE *ever* and its substrate counterpart *guò* can be schematically represented by the following the chains:

[*ever*: negative polarity] > [*ever*: experiential aspect marker]

*[*guò*: negative polarity] > [*guò*: experiential aspect marker]

Again, the assumed grammaticalization chain of experiential *guò* above does not equate with the actual grammaticalization chain that has taken place in the Chinese experiential marker. In other words, replica grammaticalization does not apply to aspectual *ever*. As such, the case of CSE *ever* only agrees with ordinary contact-induced grammaticalization, with the substrate *guò* providing the grammatical category of the experiential aspect marker. Apart from that, by comparing the ratios of experiential *ever* among OHI (2.11%, early 20th

century), ICE-SG (5.89%, 1997), and Flowerpod (6.67%, 2007–2009, Ziegeler 2015), we observe an increased ratio of CSE *ever* as an aspectual marker. The comparison results suggest an ongoing grammaticalization of CSE *ever* as an aspectual marker (see Chapter 8.3.1).

9.2.4 Grammaticalization of *one* in relation to Chinese *de*

Unlike the aspectual markers *already* and *ever*, which only agree with ordinary grammaticalization, the grammaticalization of *one* from pronominal *one* to emphatic *one* seems to support the scenario of the replica grammaticalization. Consider the following examples in (224):

- (224) a. Pronominal *one*: It was a very closed party. It was a very powerful *one*
[...] [OHI-000166-DB]
b. Nominalizer/relativizer *one*: That means bottom is shop, the top is people
staying *one*. [OHI-001953-LAS]
c. Emphatic *one*: You know, men are always very simple *one*.
[OHI-001953-LAS]

Pronominal *one* in CSE can follow an adjective, just like Standard English. Apart from that, it can follow nominal words (e.g. N-*one* silk *one* ‘a dress made of silk’), and possessive pronouns (e.g. my, your, his/her), which are ungrammatical in Standard English (see Chapter 5.4 and Chapter 8.4). The nominalizer/relativizer function seems to be a reanalysis of pronominal *one* – modeled on the Chinese substrate *de* – towards a wider distributional range of uses. Besides “nominalizing” an adjective, it can also follow a verb phrase and transform it into a noun phrase as exemplified in (224)b. Furthermore, the emphatic function of CSE *one* appears to be a reinterpretation of the nominalizer/relativizer function. As shown in (224)c, *men are always very simple one* can be reinterpreted as [[men are always very simple] *one*]. Here *one* shifts as the head of the NP [men are always [very simple [one]]] to the speaker’s personal assessment of the utterance. Thus, the development of the emphatic marker *one* can be summarized by the following chain:

[*one*: pronominal] > [*one*: nominalizer/relativizer] > [*one*: emphatic]

The functional shift of CSE *one* as a nominalizer/relativizer towards the more pragmatic function is parallel to the final phase of the grammaticalization of the substrate *de* which follows the same pathway (see Chapter 6.5.3):

[底/的 *dǐ/de*: relativizer/nominalizer] > [的 *de*: emphatic marker]

However, the earlier development of *de* is different from that of CSE *one*. Unlike CSE *one*, which extends its pronominal function to a nominalizer/relativizer function, Chinese *de* developed from a locative/spatial noun 底 *dǐ* ‘bottom’ (initially 氏 *dǐ* ‘foundation’ or ‘base’) in Old to Modern Chinese and to a relativizer/nominalizer in Middle to Modern Chinese (see Yap et al. 2017). However, the continuation of the development of *de* as a relativizer/nominalizer into an emphatic marker is identical to that of CSE *one*. As such, CSE speakers seem to have replicated the grammaticalization process that has taken place in Chinese, using an analogical formula, i.e. they develop the pronominal *one* into a new emphatic marker.

In fact, the syncretism between nominalizer/relativizers and sentence-final emphatic markers has been observed as a frequent linguistic phenomenon in other East Asian and Tibeto-Burman languages (see Davidse, Van linden, and Verstraete 2010:5; Yap and Matthews 2008). Yap and Matthews (2008), for example, studied the etymological sources of nominalizers in languages including Classical and Lhasa Tibetan, Chantyal, Gurung, Lahu, Japanese, Okinawan, Korean, as well as three Chinese dialects, i.e. Chaozhou, Cantonese, and Mandarin. They found cross-linguistically robust grammaticalization trajectories (e.g. from lexical source to pronominal to nominalizer or from nominalizer to other functions, such as relativizer, complementizer, or stance marker).

However, unlike the other East Asian and Tibeto-Burman languages, most of which are verb-final languages (e.g. Korean and Japanese), Chinese is an SVO language. This raises the question as to how the word order principles of Chinese could facilitate the reanalysis of nominalizers into sentence-final particles. As mentioned above, Chinese

nominalizers such as *zhe* and *di/de* were developed from locative/spatial nouns, i.e. so-called “light nouns” in (Yap, Choi, and Cheung 2010:71). Therefore, these nominalizers in Chinese are head-final, and when they occur in sentence-final position, they become natural carriers of the pragmatic functions – emphatic marker and/or mood/stance particles (Davidse et al. 2010:6).

Like Chinese, English is among the SVO languages. It suggests that the development of CSE *one* from its pronominal use to other nominalization constructions, e.g. *one* juxtaposed to a verb phrase or a relative clause modeled on Chinese *de*, is a crucial step for the continuation of the development of *one* into an emphatic marker.

9.2.5 Grammaticalization of *also* in relation to Chinese *yě/dōu*

Different from *already*, *ever*, and *one*, which have developed new grammatical functions in CSE, the change of *also* does not involve a change in its syntactic category, i.e. *also* is an adverb in both CSE and Standard English. Nevertheless, *also* fulfills a number of different roles in CSE, when compared to Standard English (see Chapter 5.2). For example, *also* in Standard English is restricted to positive-polarity contexts while CSE *also* can occur in both positive and negative contexts. In addition, CSE *also* can co-occur with universal quantifiers and concessive *even*, which are well-known functions of the Chinese substrate *dōu/yě* (see Chapter 6.4).

Curiously enough, the study has found cases in both OHI and ICE-SG where *also* is used in the sense of ‘even’, which suggest that *also* has developed an inherent scalar meaning. Consider the following examples:

- (225) a. After that they go on guard duty. Guard duty, you cannot talk to them. You got no chance to talk to them *also*, they ask you to get out.
[OHI-000310-SBG]
‘You got even no chance to talk to them.’
- b. Can’t get myself to start it *also* uhm uhm. [ICE-SG:S1A-091#119:1:C]
‘I can’t even get myself to start it.’

In (225)a, the speaker related a scenario where his friends went on guard duty. Trying to illustrate how busy they were and how strict the rules were, he said that they had no time to talk. He added that they even had no chance to talk. The way he presented it, he considered having the chance to talk more fundamental than having the time to talk. Therefore, the context strongly suggests a scalar reading. The same applies to (225)b: The speaker suggested that the task she was faced with was so intimidating that she could not even get herself to start it. Again, the additive marker *also* gives rise to the ‘even’ reading.

If we agree that the *even* + *also* construction is modeled on the *lián* + *yě/dōu* construction, where *also* is used to license the additional scalar component of *even* (see Chapter 8.2.2), then the use of stand-alone *also* with an inherent scalar meaning in (225) seems to suggest a further grammaticalization from the use as an additive adverb via the function of reinforcing a concessive meaning to the more marked scalar reading.

Like CSE *also*, *yě* or *dōu* can give rise to the scalar reading without *lián* ‘even’.

Consider the examples below:

- (226) 连 说话 也 没有 力气
lián shuōhuà yě méiyǒu lìqì
 even talk also NEG strength
 ‘I don’t even have the strength to speak.’
 [CCL-Contemporary-History-Li WenCheng]

- (227) a. Mandarin (personal knowledge)
shuōhuà yě méiyǒu lìqì
 talk also NEG strength
 b. Cantonese
gong2ye5 dou1 mou5 hei3
 talk also NEG strength
 ‘I don’t even have the strength to speak.’

From the above examples, we can see that an inherent scalar reading of *yě/dōu* is possible in both Mandarin and Cantonese. The construction is derived from the *lián* + *yě/dōu* construction with *lián* being omitted. It results in an additional scalar reading of *yě/dōu*, on top of its additive meaning. Clearly, the development of *yě/dōu* matches that of CSE *also*, which supports the analysis of the replica grammaticalization.

9.3 CSE in relation to speakers' social background

After identifying the change patterns of the four CSE expressions over time, other important questions to ask are where these changes start and how the substrate-influenced variables spread. These questions were addressed by integrating the language-external factors (gender, ethnic background, educational level, and year of birth) into the study. The results in this study show that ethnicity and level of education are the strongest social factors that best account for the variations in the CSE markers (see Chapter 8.1.9, Chapter 8.2.3, Chapter 8.3.2, and Chapter 8.4.3). The following sections discuss the wider theoretical implications of these findings.

9.3.1 CSE in different ethnic groups

In contrast to the findings of previous studies on CSE, where its change was assumed to be unidimensional (e.g. Schneider 2007; Bao 2015, Ziegeler 2020), the results of the present study demonstrate that the development of CSE, as manifested in the four CSE expressions, tends to be complex and multidimensional. For example, the variations found in the use of *already*, *also*, *ever* and *one* according to different ethnic groups suggest that CSE should not be treated as a homogeneous entity (see Figure 8.9, Figure 8.17, Figure 8.27, and Figure 8.34 in Chapter 8). The new focus on the internal heterogeneity of CSE spoken amongst speakers from different ethnic backgrounds allows us to see what motivates these cross-ethnic differences.

Chapter 8 mentioned the substrate-influenced CSE expressions differ in frequency between different ethnic groups. The multiple linear regression analyses confirm that the factor ethnicity is a consistent predictor of the frequencies of non-standard variables. The aspectual marker *already*, for example, serves as a good indicator of Chinese influence on CSE, i.e. Chinese speakers use proportionately higher number of *already* as well as aspectual *already* and sentence-final *already*. Chinese speakers also achieved relatively higher ratio of *also* and *one*, though to a lesser extent than *already*. In addition, Peranakan speakers, Malay

speakers and Indian speakers, as well as speakers from other ethnic backgrounds manifest higher frequencies of these substrate-influenced variables in CSE than speakers with a Eurasian or British background.

Besides direct substrate influence from ethnic languages spoken by speakers from other ethnic backgrounds, it is also possible that these Chinese-influenced features, both semantically and syntactically, were first introduced by the Chinese communities into the feature pool of CSE (see Chapter 4.8). In a sequential step, these CSE markers influenced speakers of other ethnic backgrounds. The use of *already* in negative sentences amongst speakers of a Malay background is a case in point. As mentioned in Chapter 5.1.1, Malay *sudah/dah* is rarely used in negative contexts. However, the following example was produced by an interviewee of a Malay background, which supports our hypothesis of an indirect influence from CSE. However, it is important to note that similar examples among the Malay group do not occur very frequently. There are only 4 comparable cases found amongst Malay speakers in OHI.

- (228) Especially when you hear these B29 bombs and so on. Then you *already* don't think of yourself that you are going to work anymore. [OHI-000013-SEA]

British and Eurasian groups

The study also shows that the British and Eurasian groups rarely adopted those Chinese-influenced features. The results are not very surprising, as many of the Eurasians had become anglicized and identified more with the British than their Portuguese ancestors when they came to Singapore, and adopted English as their first language (see Chapter 3.2.8, also see Ho 2013b).

Within the framework of contact-induced change, Thomason and Kaufman (1991:41) predict that the interference features usually enter the target language as spoken by the shifting speakers quite rapidly. However, the effects of a substrate language on native speakers of a target language may take much longer. If the process of language shift is rapid, the shifting speakers may not learn some patterns of the target language. The hypothesis

holds true in this study. Here in this study, the substrate language is Chinese, and the target language is Standard English; the native speakers of the target language are the speakers of the British and Eurasian groups, while the shifting speakers are the Chinese speakers and speakers from the other groups. We can see that the British and Eurasian groups rarely produced substrate-influenced uses of *already*, *also*, *ever*, and *one*, while the Chinese group and the other groups of speakers well accepted the use of these CSE expressions with substrate-influenced meanings and structures.

As noted by Thomason (2001:74), “[M]any cases of group language shift result in perfect acquisition of the target language (TL), in the sense that members of the shifting group speak the same variety of the languages as original TL group members. But in other cases, for various social reasons, the shift results in changes in the TL.” It is important to note that contact-induced changes in the target language are often not the result of the shift group lacking the ability to learn or having insufficient access to the target language. Instead, it is often the decision made by the shifting groups to use non-standard/substrate-influenced features that results in the interference process. As observed in this study, many educated Chinese, Peranakan, Malay, and Indian speakers – those who had received university education – used *already* as an aspectual marker frequently.

Cultural orientation and indexicality vs. ethnicity

The observation of considerable flexibility in the use of CSE elements underlies the culture orientation model developed by Alsagoff (2007, 2010) and the model of social indexing by Leimgruber (2009, 2013) (see Chapter 4.7.3). Alsagoff (2010:340) argues that the variation in the use of English in Singapore arises from a cultural tension between “being/doing global” and “being/doing local”. She adopted the notion of a “macro-culture”, which refers to a collective identity of Singaporeans “forged through common ways of speaking, living and doing, and is thus associated with notions of citizenry, national identity based on a collective disposition and history” (Alsagoff 2010:340). Leimgruber (2009, 2013) also suggests that CSE speakers draw different elements of CSE (e.g. discourse markers such as *lah* and *ah*)

(see Chapter 4.3.2) to demonstrate their community membership in a local context while they use more standard features (e.g. use of inflectional markers such as *-ed* and *-s*) or avoid code-switching and borrowings in a more formal and global context.

However, as the results in this study show, the variation in the use of the CSE markers is tied primarily to ethnicity, which can be associated with the use of the corresponding mother tongue(s) of the ethnic groups. Besides serving as an inter-ethnic *lingua franca* or a collective entity, CSE turned out to be heterogenous, i.e. each of the different ethnic groups, as represented by the speakers in OHI, spoke a slightly different variety of CSE. Apart from that, speakers did not seem to be consciously aware of whether a specific expression represents a local term or a global term. For example, we found *already* appearing twice in a clause, with the antecedent *already* functioning as a PhP expression – the standard code, and the sentence-final *already* as an aspectual marker – the CSE code (see Chapter 8.1.7). It seems to suggest that speakers tend to mix standard and substrate-influenced codes unconsciously, instead of being consciously aware of the cultural norms defined by us as observers. The findings suggest that finer-grained descriptions on speakers of different ethnic communities are needed for more accurate and robust theoretical discussions on contact languages.

9.3.2 CSE according to educational level

Apart from ethnicity, we also see a strong correlation between the frequencies of CSE markers and speakers' level of education. This is particularly manifested in the wide range of variation amongst the Chinese speakers. For example, the ratio of *already* amongst the Chinese speakers ranges from 0 per page to 0.89 per page, with higher frequency of aspectual *already* and sentence-final *already* pointing to the more basilectal speakers.

To a certain degree, the result of a negative correlation between frequencies of substrate-influenced linguistic variables and educational level suggests that there is a lectal stratification (acrolect, mesolect, and basilect) among the Singaporean speakers in OHI, which is part of the post-creole continuum hypothesis (Platt 1980:108–135, see Chapter

4.7.1). The results in OHI show that speakers with a higher educational level produce much lower frequencies of *already*, *also*, and *one* with substrate-derived grammatical meanings while speakers with a lower educational level tend to use substrate-influenced *already*, *also*, *ever* and *one* much more frequently (see Figure 8.10, Figure 8.18, and Figure 8.35 in Chapter 8).

However, the post-creole continuum model did not address the question as to why speakers with a high education level also use CSE expressions in their speech. As shown in Chapter 8.5, speakers in OHI, including those with a high level of education, tend to switch to a more informal register in the second half of their interviews, when compared to the first half. It shows that instead of being a variety that arises from a lack of competence to command the standard variety, CSE remains a variety with a sociolinguistic preference, desired by speakers in a more relaxed/spontaneous speech.

The results in OHI also reveal that speakers with a high educational level tend to use the standard codes more consistently, as there is less heterogeneity among the educated speakers. In other words, there is less variation in the frequencies of substrate-influenced expressions among speakers with a high level of education. In contrast, there is significant variation among speakers with lower educational levels. This finding does not fully agree with the post-creole continuum model, which suggests that speakers at the higher end of the social continuum have a wider range of available lects, while those at the lower end have a more restricted range of linguistic choices (Platt 1975:369, also see Chapter 4.7.1). As suggested by Figure 8.10 (concerning *already*), Figure 8.18 (*also*) and Figure 8.35 (*one*) in relation to educational level in Chapter 8, speakers with a low or medium level of education used substrate-influenced variables in a broader range of frequencies. That is to say that basilectal and mesolectal speakers actually occupy a wider range of the lectal continuum than the acrolectal speakers.

Furthermore, the model of Platt (1975) suggested that Singaporean speakers would move towards the acrolectal end with the widening spread of English education and the increasing number of Singaporeans using English as their dominant language. However, this

study shows that CSE has remained relatively stable in the past 100 years and that there have always been individual variations subject to register, style, as well as ethnic background.

The main problem of the previous models is that they tend to describe CSE either by means of diglossia or generalize CSE in relation to macro-cultural constructions of identity and communicative purpose. As Gupta (2001:365) pointed out: “countries do not speak English – people do”. Similarly, as suggested by Mufwene’s (2001:16) feature pool theory, it is the “intervention of will” of individual speakers that plays a significant role in language evolution (see Chapter 4.8). The conscious decisions of speakers to use language in a similar or different way compared to some other specific speakers for reasons of identity, can have an important impact on a communal language (Mufwene 2001:16). The findings of the individual variation in terms of the frequencies of substrate-influenced *already*, *also*, *ever*, and *one* in CSE in relation to ethnicity and education level suggest that a more comprehensive model is needed to capture CSE and its delicate micro-variations. These micro-variations are presented in the forms of frequency differences of the four CSE expressions in this study according to individual speakers, their ethnic background, and level of education.

9.4 Summary

The preceding sections first discussed implications of the diachronic perspective on CSE. The findings suggest that CSE has remained stable. In addition, stylistic as well as other sociolinguistic variations may have been present from very early on in the development of CSE. This chapter has shown how the aspectual markers *already* and *ever*, the emphatic marker *one* and the additive marker *also* in CSE are related to their Chinese substrates. Both the transfer analysis (Bao 2015; see Chapter 4.7.4) and contact-induced grammaticalization (Heine and Kuteva 2005; see Chapter 4.4) apply to the four CSE expressions to a certain extent.

Bao’s (2005, 2015) systemic transfer and lexifier filter theory provides a robust theoretical basis for the link between the functions of the CSE expressions and their Chinese

counterparts. However, the influence of Bao's (2005, 2015) lexifier filter, which suggests that the lexifier language imposes constraints on the later distribution of the substrate-influenced variables, does not explain the increased ratio of substrate-influenced tokens by comparing the results in OHI with ICE-SG.

In contrast, ordinary contact-grammaticalization is applicable to the grammaticalization of the CSE expressions in this study. The grammaticalization of *already*, *also*, *ever*, and *one* is based on a functional need originating in the substrate language, which was found in their counterparts in the Chinese substrates. This chapter further discussed the possibility of a replica grammaticalization of the four CSE expressions. The cases of the aspectual markers *already* and *ever* seem to have only followed ordinary contact-induced grammaticalization, as their pathways of grammaticalization do not match those of their Chinese counterparts. However, emphatic *one* and additive *also* appear to be two cases of replica grammaticalization, as these two expressions seem to have replicated the grammaticalization models that already existed in Chinese into CSE. It might be a coincidence that the grammaticalization of *one* has undergone the same grammaticalization path of Chinese *de*. The same may apply to the case of *also*, which has developed an additional scalar meaning, just like Chinese *yě/dōu*. The choice between ordinary grammaticalization and replica grammaticalization seems to vary depending on the grammatical item.

Finally, this chapter concludes that a more comprehensive model is needed to embrace the sociolinguistic findings in this study. Further studies on CSE should consider the interaction between individual speakers and their social backgrounds, such as ethnicity and educational level, which have been either generalized away or neglected in previous studies regarding the diachronic development of contact-varieties.

10 Concluding remarks and outlook

The aim of this study was to offer a historical sociolinguistic reconstruction of Colloquial Singapore English (CSE), by using data produced by speakers who were born in the late 19th and early 20th century drawn from the Oral History Interviews (OHI).

The study started with an overview of the historical and linguistic background of CSE, with a summary of the various models presented to date on the emergence of CSE and its relationship to Standard Singapore English as well as its Chinese substrates. The research background and the research questions were described in detail in Chapter 2. Chapter 3 gave an account of the historical background which elucidates the language-related policies and the ethnic dynamics in Singapore from the precolonial era (before 1819) to the modern-day era (1965–present). Chapter 4 presented the theoretical underpinnings of this study, including general language contact theories (e.g. Thomason 2001), scales of borrowability and hierarchies of shift-induced interference (e.g. Field 2002; Matras 2009), “contact-induced grammaticalization” (Heine and Kuteva 2005), and various models on the emergence of CSE and its relationship to Standard Singapore English (Platt 1975; Gupta 1989; Alsagoff 2007, 2010; Leimgruber 2009, 2013) as well as its Chinese substrates (Bao 2005, 2015; Bao and Hong 2006). Apart from that, Chapter 4 introduced the “Dynamic Model” (Schneider 2007), which provides a diachronic outline for the emergence and development of Singapore English.

Since it is impossible to provide a comprehensive account of the diachronic developments, this study focused on four grammatical exponents of CSE, namely the aspectual marker *already*, the additive marker *also*, the experiential marker *ever*, and the emphatic marker *one*, which have parallels in the Chinese substrates. Here, Mandarin Chinese was adopted as a representative of the Chinese substrates due to the morphosyntactic similarities of between Mandarin and the other Sinitic substrates (e.g. Hokkien and Cantonese) (see Bao 2015:20; Hiramoto 2015, also see Chapter 4.7.4 and the introduction in Chapter 5). Chapter 5 discussed the uses of these four expressions in CSE and argued that the functional extensions of these CSE expressions are due to cross-linguistic influences

from the Chinese substrates. Chapter 6 explored the semantic functions and syntactic features of the Chinese counterparts of the four CSE markers. The parallels between them support the assumption that the CSE variants and their semantic extension and/or grammaticalization are largely due to cross-linguistic influence from the Chinese substrates.

Chapter 7 presented the main database of this study – the *Oral History Interviews* drawn from the *National Archives of Singapore* (OHI-NAS 2020), followed by a description of the methods applied in the empirical part. The study calculated the occurrences of *already*, *also*, *ever* and *one* according to their various functions – both standard and substrate-influenced – based on 100 interviews from OHI conducted between 1979 and 2009 amongst speakers born between 1899 and 1983. The speakers were subdivided into different groups with respect to their ethnic background, level of education, age group, and gender. In the next step, the occurrences of the four CSE expressions were compared with those in ICE-SG, which represents a more recent data source among university students starting from 1997 (Nelson 2002:3). The Chinese Corpus held by the Center of Chinese Linguistics (CCL) was used as a complementary database for examples of the Chinese substrates.

Chapter 8 presented the results and data analyses. Chapter 9 discussed the results relating to the previous theories and gave answers to the research questions. Questions were addressed concerning the stability of Colloquial Singapore English, the correlations between the production of CSE markers and social backgrounds of the interviewees as well as the effects of the Chinese substrates on these four CSE markers.

The results of this study suggest that CSE has undergone no substantial changes, at least judging from the use of *already*, *also*, *ever*, and *one*. As the OHI data represents CSE in the early 20th century, it can be assumed to belong to stage 2 and stage 3 in the Dynamic Model. Yet the OHI data did not suggest a strong shift away from CSE towards Standard English. In contrast, there seems to be a slight increase in the substrate-influenced usages, especially among younger speakers in the OHI data. The same observation also holds true when comparing the frequencies of these CSE markers in OHI with those in ICE-SG, i.e. the ratios of substrate-influenced tokens are generally higher in ICE-SG than in OHI. We need to bear in mind, though, that there are substantial differences between individual speakers.

Interestingly enough, OHI speakers show similar behavior in their usage of *already* and *also*: the frequencies of *already* and *also* – both overall frequencies and clause-final frequencies – are positively correlated. The individual differences are strongly contingent on the ethnic background and the level of education of the speakers. Except for the British and Eurasian communities, these novel usages of the four CSE expressions were well-accepted by the other speech communities. Although these substrate-influenced features exist among speakers with a Chinese ethnic background as well as speakers with other ethnic backgrounds, they differ with regard to the frequencies of these variables. The Chinese speaking communities generally produced higher frequencies of these CSE expressions. On the other hand, other ethnic communities such as the Indian and Malay speakers use these expressions relatively frequently as well, though to a lesser degree.

Apart from that, this study has provided evidence that stylistic as well as other sociolinguistic variations may have been present from very early on in the development of CSE. However, these sociolinguistic variations were only expected to appear in the stages of differentiation in Schneider's (2007) Dynamic Model and therefore were either neglected or generalized away in research studying the emergence of CSE. Moreover, the degree of nativization and stabilization crucially depends on individual speakers, as the use of CSE is contingent on their ethnic background and level of education. Although CSE is considered a resource that speakers draw on for purposes of cultural indexing (Leimgruber 2014), especially those in the upper strata of society, there are speakers in the OHI data who use CSE as their native code, in which they draw exponents from their ethnic languages.

The reason why different ethnic groups in Singapore speak a slightly different variety of CSE may be that the intra-ethnic ties are more prevalent than inter-ethnic ties – both historically and synchronically. According to a recent study by Chua, most Singaporean relationships are based on strong intra-ethnic ties (Chua 2015, also see Chapter 3.2). If this holds true for today's Singapore, it must apply in an even stronger way to the historical and institutional contexts of Singapore, supported by, for instance: (i) the ethnic compartmentalization in precolonial period; (ii) the official division of Singapore's population in colonial times under British rules; and (iii) the occupation of the ethnic groups

in different economic sectors (see Chapter 3.3.2). Therefore, it is plausible that ethnicity turned out to be a relatively strong factor in the multiple linear regression analyses, being more important than age, gender, and year of birth. With attitudes towards CSE rising and more people embracing it to express solidarity with Singapore, cultural indexing (Leimgruber 2014) is likely to become the main function of CSE.

The study also explored the use of the CSE expressions in relation to their Chinese counterparts in further detail. To a large extent, Chinese has provided the extended semantic meanings and syntactic features, which were added to the already existing usages in Standard English. These features may be strengthened by other ethnic languages such as Malay and Tamil, which exhibit typological similarities. In a small number of cases, the influence from Chinese in the CSE expressions is so strong that the whole CSE sentence no longer meets the surface structural requirements of English. Furthermore, the study has also explored the possibility of a replica grammaticalization of these CSE expressions modeled on the Chinese substrate. The findings suggest that the paths of grammaticalization of *also* and *one* mirror those of their Chinese equivalents, which supports the model of replica grammaticalization. However, the choice between ordinary grammaticalization and replica grammaticalization seems to vary depending on the grammatical item. In view of the population composition of Singapore, in which the Chinese communities make up approximately 75% of the population (see Chapter 3.2), the resurgence in the prominence of Mandarin due to socio-economic reasons (see Chapter 3.3.4), and the recent immigration of people from all parts of China, I submit that there may continue to be a grammaticalization of these CSE expressions.

Since it is impossible to give a comprehensive account of the emergence and development of CSE, this study chose four salient linguistic variables in CSE, namely the use of *already*, *also*, *ever*, and *one*. However, research has yet to be conducted to take into account a wider range of structural features that define CSE, as outlined in Chapter 1.2, including the lack of consistent morphological marking of the past tenses, the use of bald singular noun phrases, and the use of pragmatic markers such as *ah*, *lah*, *meh*, *leh*, and *lor*. Apart from that, future studies on CSE from a diachronic perspective should also examine

phonological changes (e.g. the segmental, rhythmic, and intonational features in CSE) and their relations to the substrates by using the recordings of OHI.

Although the number of interviews of this study is relatively high (n=100), with the total length of interviews adding up to 16,147 pages, the study only examines the use of four expressions. Apart from that, the interviewees are not evenly distributed across different ethnic groups. The majority of speakers (n=57) have a Chinese ethnic background. Therefore, the numbers of speakers in the other ethnic groups are fairly small, e.g. the Malay group and the Peranakan group each include 8 speakers, while the British and the Eurasian groups consist of only 7, and 6 interviewees. The group of “Other” includes only 2 speakers, with one having a mix of Chinese and Indian ethnic background while the other with an Iraqi background. Apart from that, we have an unbalanced ratio of female vs. male speakers (approximately 4:1) (see Chapter 7.1.2). Therefore, the multiple linear regression analyses only demonstrated moderate predictive power of the use of substrate-influenced *already*, *also*, *ever* and *one* in relation to the social variables of the individual speakers.

Nevertheless, the sociolinguistic findings in this study suggest that future studies of contact varieties should take into account internal variations within the emerging variety, especially when there are different ethnic groups present in the society. Instead of treating a new contact variety as a collective entity at the societal level, subtle differences among speakers of different groups need to be considered to contribute to more accurate explanations and more robust theories of contact phenomena in the local contexts.

References

- Abshire, Jean E. 2011. *The History of Singapore*. Santa Barbara, CA: Greenwood.
- Aljamal, Yousef. 2015. *Dreaming in Arabic: Malaysia Celebrates Arabic Language*.
- Alsagoff, Lubna. 1995. "Colloquial Singapore English: The Relative Clause Construction." Pp. 77–87 in *The English Language in Singapore: Implications for Teaching*, edited by S. C. Teng and M. L. Ho. Singapore: Singapore Association for Applied Linguistics.
- Alsagoff, Lubna. 2007. "Singlish: Negotiating Culture, Capital and Identity." Pp. 25–46 in *Language, capital, culture*, edited by V. Vaish, S. Gopinathan, and Y. Liu. Rotterdam: Sense Publishers.
- Alsagoff, Lubna. 2010. "English in Singapore: Culture, Capital and Identity in Linguistic Variation." *World Englishes* 29:336–48. doi: 10.1111/j.1467-971X.2010.01658.x.
- Alsagoff, Lubna. 2015. "Creative Enactments of Language Teacher Education Policy: A Singapore Case Study." Pp. 125–41 in *International Perspectives on English Language Teacher Education*, edited by T. S. C. Farrell. London: Palgrave Macmillan UK.
- Alsagoff, Lubna, and Chee Lick Ho. 1998. "The Grammar of Singapore English." Pp. 127–51 in *English in new cultural contexts*, edited by J. Foley. Singapore and Oxford: Oxford University Press.
- Ang, Jolene. 2019. "A Friendlier, More Relatable Speak Good English Movement." *The Straits Times*. Retrieved March 19, 2020 (<https://www.straitstimes.com/singapore/a-friendlier-more-relatable-speak-good-english-movement>).
- Anon. 2012. "The Eurasian Diaspora: What It Means for Those Who Emigrated-and Those Who Stayed at Home." *Eurasian Association*. Retrieved (<http://www.eurasians.org.sg/wp-content/uploads/2016/06/15-Jan-Mar-2012.pdf>).
- Ansaldo, Umberto. 1999. "Comparative Constructions in Sinitic: Areal Typology and Patterns of Grammaticalization." PhD Thesis, Stockholm University, Stockholm.
- Ansaldo, Umberto. 2004. "The Evolution of Singapore English: Finding the Matrix." Pp. 129–52 in *Singapore English: A grammatical description*, edited by L. Lim. Amsterdam: John Benjamins.
- Ansaldo, Umberto. 2009. *Contact Languages: Ecology and Evolution in Asia*. Cambridge and New York: Cambridge University Press.

- Ansaldo, Umberto, and Lisa Lim. 2012. "English in Asia." in *Areal Features of the Anglophone World*, edited by R. Hickey. Berlin and Boston: De Gruyter.
- Ansaldo, Umberto, Lisa Lim, and Salikoko S. Mufwene. 2007. "The Sociolinguistic History of the Peranakans: What It Tells Us about 'Creolization.'" Pp. 203–26 in *Deconstructing Creole*, edited by U. Ansaldo, S. Matthews, and L. Lim. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Arumainathan, P. 1971. *Census of Population 1979, Singapore*. Singapore: Department of Statistics.
- Asian Studies Archive. 1994. "Singapore: History." Retrieved February 25, 2020 (<https://web.archive.org/web/200703232095958/http://inic.utexas.edu/asnic/countries/Singapore/Singapore-History.html>).
- van der Auwera, Johan. 1993. "'Already' and 'Still': Beyond Duality." *Linguistics and Philosophy* 16(6):613–53. doi: 10.1007/BF00985436.
- van der Auwera, Johan. 1998. "Phasal Adverbials in the Languages of Europe." in *Adverbial Constructions in the Languages of Europe*, edited by J. van der Auwera. Berlin and New York: De Gruyter.
- van Baar, Theodorus M. 1997. *Phasal Polarity*. Dordrecht: Foris Publications.
- Bailey, Charles J., and Karl Maroldt. 1977. "The French Lineage of English." Pp. 21–53 in *Langues en contact: pidgins, Creoles*. Tübingen: Narr.
- Bakker, Peter, and Robert A. Papen. 1997. "Michif: A Mixed Language Based on Cree and French." Pp. 295–363 in *Contact languages: A wider perspective*, edited by S. G. Thomason. Amsterdam and London: John Benjamins.
- Bamiro, Edmund O. 1995. "Syntactic Variation in West African English." *World Englishes* 14(2):189–204. doi: 10.1111/j.1467-971X.1995.tb00349.x.
- Bao, Zhiming. 1995. "Already in Singapore English." *World Englishes* 14(2):181–88. doi: 10.1111/j.1467-971X.1995.tb00348.x.
- Bao, Zhiming. 1998. "The Sounds of Singapore English." Pp. 154–74 in *English in New Cultural Contexts: Reflections from Singapore*. Singapore: Oxford University Press.
- Bao, Zhiming. 2005. "The Aspectual System of Singapore English and the Systemic Substratist Explanation." *Journal of Linguistics* 41(2):237–67. doi: 10.1017/S0022226705003269.
- Bao, Zhiming. 2009. "One in Singapore English." *Studies in Language* 33(2):338–365. doi: 10.1075/sl.33.2.05bao

- Bao, Zhiming. 2011. "One in Singapore English." Pp. 79–106 in *Benjamins Current Topics*, edited by J. C. Clements and S. Gooden. Amsterdam: John Benjamins.
- Bao, Zhiming. 2015. *The Making of Vernacular Singapore English: System, Transfer and Filter*. Cambridge and New York: Cambridge University Press.
- Bao, Zhiming. 2017. "Transfer Is Transfer; Grammaticalization Is Grammaticalization." in *Changing English*, edited by M. Filppula, J. Klemola, A. Mauranen, and S. Vetchinnikova. Berlin and Boston: De Gruyter.
- Bao, Zhiming. 2021. "The Origins of Singapore English: A Sociohistorical Sketch." Pp. 19–37 in *Multilingual global cities: Singapore, Hong Kong, Dubai.*, edited by P. Siemund and J. R. E. Leimgruber. Singapore: Routledge.
- Bao, Zhiming, and Khin Khin Aye. 2010. "Bazaar Malay Topics." *Journal of Pidgin and Creole Languages* 25(1):155–71. doi: 10.1075/jpcl.25.1.06bao.
- Bao, Zhiming, and Huaqing Hong. 2006. "Diglossia and Register Variation in Singapore English." *World Englishes* 25(1):105–14. doi: 10.1111/j.0083-2919.2006.00449.x.
- Bao, Zhiming, and Hui Min Lye. 2005. "Systemic Transfer, Topic Prominence, and the Bare Conditional in Singapore English." *Journal of Pidgin and Creole Languages* 20(2):269–91. doi: 10.1075/jpcl.20.2.03zhi.
- Barr, Michael D. 2000. "Lee Kuan Yew and the 'Asian Values' Debate." *Asian Studies Review* 24(3):309–34. doi: 10.1080/10357820008713278.
- Barrett, E. C. G. 2004. "Winstedt, Sir Richard Olaf (1878–1966), Civil Servant and Malay Scholar."
- Baskaran, Lohanayahi. 1987. "Aspects of Malaysian English Syntax." PhD Thesis, University of London.
- Bergs, Alexander, and Laurel J. Brinton, eds. 2017. *The History of English: Varieties of English*. Berlin and Boston: De Gruyter.
- Biber, Douglas. 1992. "The Multi-Dimensional Approach to Linguistic Analyses of Genre Variation: An Overview of Methodology and Findings." *Computers and the Humanities* 26(5/6):331–45.
- Biber, Douglas E., and Bethany Gray. 2011. "Grammatical Change in the Noun Phrase: The Influence of Written Language Use." *English Language and Linguistics* 15(2):223–50. doi: 10.1017/S1360674311000025.

- Biewer, Carolin. 2015. *South Pacific Englishes: A Sociolinguistic and Morphosyntactic Profile of Fiji English, Samoan English and Cook Islands English*. Amsterdam: John Benjamins.
- Blommaert, Jan. 2010. *The Sociolinguistics of Globalization*. Cambridge and New York: Cambridge University Press.
- Bloom, David. 1986. "The English Language and Singapore: Accritical Survey." Pp. 337–452 in *Singapore Studies: Critical Surveys of the Humanities and Social Sciences*, edited by B. K. Kapur. Singapore: Singapore University Press.
- Bokhorst-heng, Wendy. 1998. "Language Planning and Management in Singapore." Pp. 287–309 in *English in new cultural contexts*, edited by J. Foley. Singapore and Oxford: Oxford University Press.
- Bokhorst-Heng, Wendy. 1999. "Singapore's Speak Mandarin Campaign: Language Ideological Debates in the Imagining of the Nation." in *Language Ideological Debates*, edited by J. Blommaert. Berlin and New York: De Gruyter.
- Bokhorst-Heng, Wendy D. 2005. "Debating Singlish." *Multilingua - Journal of Cross-Cultural and Interlanguage Communication* 24(3):185–209. doi: 10.1515/mult.2005.24.3.185.
- Bolton, Kingsley. 2000a. "Researching Hong Kong English: Bibliographical Resources." *World Englishes* 19(3):445–52. doi: 10.1111/1467-971X.00192.
- Bolton, Kingsley. 2000b. "The Sociolinguistics of Hong Kong and the Space for Hong Kong English." *World Englishes* 19(3):265–85. doi: 10.1111/1467-971X.00179.
- Bolton, Kingsley, and David Graddol. 2012. "English in China Today." *English Today* 28(3):3–9. doi: 10.1017/S0266078412000223.
- Bolton, Kingsley, and Bee Chin Ng. 2014. "The Dynamics of Multilingualism in Contemporary Singapore: The Dynamics of Multilingualism in Contemporary Singapore." *World Englishes* 33(3):307–18. doi: 10.1111/weng.12092.
- Braga-Blake, Myrna, and Ann Ebert-Oehlers. 1992. "Where the Twain Met." Pp. 13–30 in *Singapore Eurasians*, edited by M. Braga-Blake and A. Ebert-Oehlers. Singapore: Eurasian Association.
- Brown, Charles Cuthbert. 1983. *Sejarah Melayu or Malay Annals: An Annotated Translation*. Kuala Lumpur: Oxford University Press.
- Bruthiaux, Paul. 2003. "Squaring the Circles: Issues in Modeling English Worldwide." *International Journal of Applied Linguistics* 13(2):159–78. doi: 10.1111/1473-4192.00042.

- Buckley, Charles Burton. 1984. *An Anecdotal History of Old Times in Singapore: 1819-1867*. Singapore and New York: Oxford University Press.
- Burridge, Kathryn. 2006. "Language Contact and Convergence in Pennsylvania German." Pp. 179–200 in *Grammars in contact: A cross-linguistic typology*. Oxford University Press.
- Buschfeld, Sarah. 2013. *English in Cyprus or Cyprus English: An Empirical Investigation of Variety Status*. Amsterdam and Philadelphia: John Benjamins.
- Buschfeld, Sarah. 2021. "Multilingual Language Acquisition in Singapore." Pp. 205–28 in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*, edited by P. Siemund and J. R. E. Leimgruber. New York: Routledge.
- Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Oxford University Press.
- Bybee, Joan L., and Östen Dahl. 1989. "The Creation of Tense and Aspect Systems in the Languages of the World." *Studies in Language* 13(1):51–103. doi: 10.1075/sl.13.1.03byb.
- Bybee, Joan L., Revere D. Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Carlier, Anne, Michèle Goyens, and Béatrice Lamiroy. 2013. "De: A Genitive Marker in French?: Its Grammaticalization Path from Latin to French." Pp. 141–216 in *The Genitive*, edited by A. Carlier and J.-C. Verstraete. Amsterdam: John Benjamins.
- Cavallaro, Francesco, and Bee Chin Ng. 2009. "Between Status and Solidarity in Singapore." *World Englishes* 28(2):143–59. doi: 10.1111/j.1467-971X.2009.01580.x.
- Cavallaro, Francesco, and Bee Chin Ng. 2014. "Language in Singapore: From Multilingualism to English Plus." Pp. 33–48 in *Challenging the Monolingual Mindset*, edited by J. Hajek and Y. Slaughter. Bristol: Multilingual Matters.
- Cavallaro, Francesco, and Bee Chin Ng. 2021. "Multilingualism and Multiculturalism in Singapore." Pp. 133–59 in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*, edited by P. Siemund and J. R. E. Leimgruber. New York: Routledge.
- Cavallaro, Francesco, Bee Chin Ng, and Mark Fifer Seilhamer. 2014. "Singapore Colloquial English: Issues of Prestige and Identity: Singapore Colloquial English: Issues of Prestige and Identity." *World Englishes* 33(3):378–97. doi: 10.1111/weng.12096.

- Cavallaro, Francesco, and Stefan K. Serwe. 2010. "Language Use and Language Shift among the Malays in Singapore." Pp. 120–70 in *Applied Linguistics Review*. Vol. 1, edited by L. Wei. Berlin and New York: De Gruyter.
- Chambers, J. K. 2003. *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*. 2nd ed. Oxford and Malden: Blackwell.
- Chan, Jim Y. H. 2013. "Contextual Variation and Hong Kong English: Contextual Variation and Hong Kong English." *World Englishes* 32(1):54–74. doi: 10.1111/weng.12004.
- Chao, Yuen-ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Cheeseman, Harold. 1935. "A Century of Education in Singapore."
- Chew, Phyllis Ghim-Lian. 1999. "Linguistic Imperialism, Globalism and the English Language." Pp. 37–47 in *English in a Changing World*, edited by D. Graddol and U. H. Meinhof. Guildford: Biddles Ltd.
- Chew, Phyllis Ghim-Lian. 2012. "Language Teaching Methods in Singapore Primary Schools: An Historical Overview." Pp. 21–23 in *Primary School English-Language Education in Asia: From Policy to Practice*, edited by B. Spolsky and Y. Moon. New York: Routledge.
- Chiew, Seen-Kong. 1980. "Bilingualism and National Identity: A Singapore Case Study." in *Language and Society in Singapore*. Singapore: Singapore University Press.
- Chua, Vincent. 2015. "How Do Singaporeans Connect? Ties Among Chinese, Malays, and Indians." *American Behavioral Scientist* 59(9):1115–28. doi: 10.1177/0002764215580617.
- Clammer, John. 1988. "Minorities and Minority Policy in Singapore." *Southeast Asian Journal of Social Science* 16(2):96–110.
- Clayton, Cathryn H. 2009. *Sovereignty at the Edge: Macau & the Question of Chineseness*. Cambridge: Harvard University Press.
- Clements, J. Clancy. 2003. "Reviewed Work: The Ecology of Language Evolution by Salikoko Mufwene." *Language in Society* 32(4):587–92.
- Clements, J. Clancy. 2006. "The Lexicalization — Grammaticalization Continuum." Pp. 77–101 in *Creole Language Library*, edited by J. C. Clements, T. A. Klingler, D. Piston-Hatlen, and K. J. Rottet. Amsterdam: John Benjamins.
- Coblin, W. South. 2000. "A Brief History of Mandarin." *Journal of the American Oriental Society* 120(4):537–52. doi: 10.2307/606615.

- van Coetsem, Frans. 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht and Holland: Foris Publications.
- van Coetsem, Frans. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Heidelberg: Winter.
- Cohen, Myron L. 1968. "The Hakka or Guest People: Dialect as a Sociocultural Variable in Southeast China." *Ethnohistory* 15(3):237–92.
- Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Constable, Nicole, ed. 1996. *Guest People: Hakka Identity in China and Abroad*. Washington: University of Washington Press.
- Crawford, John. 2000. *Journal of an Embassy from the Governor-General of India to the Courts of Siam and Cochin China Exhibiting a View of the Actual State of Those Kingdoms*. New Delhi: Asian Educational Services.
- Crewe, William James, and Linden Vargish. 1977. *Singapore English and Standard English: Exercises in Awareness*. Singapore: Eastern Universities Press.
- Crystal, David. 2003. *English as a Global Language*. 2nd edition. Cambridge and New York: Cambridge University Press.
- Dahl, Östen, and Bernhard Wälchli. 2016. "Perfects and Imperatives: Two Gram Types in One Grammatical Space." *Letras de Hoje* 51(3):325–48.
- Davidse, Kristin, An Van linden, and Jean-Christophe Verstraete. 2010. "Introduction." Pp. 1–16 in *Typological Studies in Language*, edited by A. Van linden, J.-C. Verstraete, and K. Davidse. Amsterdam: John Benjamins.
- Davies, Mark, and Robert Fuchs. 2015. "Expanding Horizons in the Study of World Englishes with the 1.9 Billion Word Global Web-Based English Corpus (GloWbE)." *English World-Wide* 36(1):1–28.
- Deterding, David. 2005. "Listening to Estuary English in Singapore." *TESOL Quarterly* 39(3):425. doi: 10.2307/3588488.
- Deterding, David. 2007. *Singapore English*. Edinburgh: Edinburgh University Press.
- Dixon, Quentin L. 2005. "The Bilingual Education Policy in Singapore: Implications for Second Language Acquisition." Pp. 625–35 in *Proceedings of the 4th International Symposium on Bilingualism*, edited by J. Cohen. Somerville: Cascadilla Press.

- Edb. 2015. "The New Yorker Series: Transforming Singapore Innovation." Retrieved January 2, 2017 (<https://www.edb.gov.sg/content/edb/en/resources/Publications/the-new-yorker-series-singapore-transforming-innovation.html>).
- Ellis, Nick C. 2002. "Frequency Effects in Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition." *Studies in Second Language Acquisition* 24(2):143–88.
- Erbaugh, Mary S. 1992. "The Secret History of the Hakkas: The Chinese Revolution as a Hakka Enterprise." *The China Quarterly* 132:937–68.
- Evans, Stephen. 2000. "Hong Kong's New English Language Policy in Education." *World Englishes* 19(2):185–204. doi: 10.1111/1467-971X.00168.
- Evans, Stephen. 2009. "The Evolution of the English-Language Speech Community in Hong Kong." *English World-Wide. A Journal of Varieties of English* 30(3):278–301. doi: 10.1075/eww.30.3.03eva.
- Faraci, Biagio. 2014. "The Contribution of Malay Language to the Enrichment and Development of Singlish Lexicon in 13 Words and 9 Typical Dishes." Retrieved (<http://www.singlish.it/malay-words-in-singlish/>).
- Ferguson, Gibson. 2006. *Language Planning and Education*. Edinburgh University Press.
- Field, Andy. 2017. *Discovering Statistics Using IBM SPSS Statistics*. 5th edition. Thousand Oaks, CA: SAGE Publications.
- Field, Fredric. 2002. *Linguistic Borrowing in Bilingual Contexts*. Amsterdam: John Benjamins.
- Firth, Alan. 1996. "The Discursive Accomplishment of Normality: On 'Lingua Franca' English and Conversation Analysis." *Journal of Pragmatics* 26(2):237–59. doi: 10.1016/0378-2166(96)00014-8.
- Fishman, Joshua A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon and Philadelphia: Multilingual matters.
- Foley, Joseph, ed. 1998. *English in New Cultural Contexts: Reflections from Singapore*. Singapore: Oxford University Press.
- Fong, Vivienne. 2005. "Unmarked Already." Pp. 251–67 in *Perspectives on Aspect*, edited by H. J. Verkuyl, H. de Swart, and A. van Hout. Dordrecht: Springer Netherlands.
- Fong, Vivienne. 2004. "The Verbal Cluster." Pp. 75–104 in *Singapore English: A grammatical description*, edited by L. Lim. Amsterdam: John Benjamins.

- Forker, Diana. 2016. "Toward a Typology for Additive Markers." *Lingua* 180:69–100. doi: 10.1016/j.lingua.2016.03.008.
- Fuchs, Robert. 2012. "Focus Marking and Semantic Transfer in Indian English: The Case of *Also*." *English World-Wide* 33(1):27–53. doi: 10.1075/eww.33.1.02fuc.
- Fuchs, Robert, and Ulrike Gut. 2012. "Do Women Use More Intensifiers than Men? Investigating Gender and Age-Specific Language Use with the International Corpus of English." in *ICAME 33*. Leuven.
- Fuchs, Robert, and Ulrike Gut. 2015. "An Apparent Time Study of the Progressive in Nigerian English." Pp. 373–88 in *Studies in Corpus Linguistics*. Vol. 67, edited by P. Collins. Amsterdam: John Benjamins Publishing Company.
- Gardner, Robert. 1985. *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*. London: Edward Arnold.
- Gil, David. 2003. "English Goes Asian: Number and (in)Definiteness in the Singlish Noun Phrase." Pp. 20–27 in *Noun phrase structure in the languages of Europe, Empirical approaches to language typology*, edited by F. Plank. Berlin: De Gruyter.
- Gogolin, Ingrid. 2000. "Minderheiten, Migration Und Forschung. Ergebnisse Des DFG-Schwerpunktprogramms FABER." Pp. 15–35 in, edited by B. Nauck and I. Gogolin. Wiesbaden: Springer.
- Goh, Chok Tong. 1998. "Speech by Prime Minister Goh Chok Tong at the Marine Parade National Day Dinner."
- Goh, Siew Huang Melissa, Ying Wei Lim, and Seah Yiling Fiona Cherie. 2016. "My Father Tongue: A Campaign to Revitalise the Use of Chinese Dialects among Youth in Singapore." Retrieved March 8, 2020 (<https://blogs.ntu.edu.sg/wkwsci-fyp-collection/2016/05/23/%EF%BB%BFmy-father-tongue-a-campaign-to-revitalise-the-use-of-chinese-dialects-among-youth-in-singapore/>).
- Gong, Yuanyuan, Irene Hau-siu Chow, and David Ahlstrom. 2011. "Cultural Diversity in China: Dialect, Job Embeddedness, and Turnover." *Asia Pacific Journal of Management* 28(2):221–38. doi: 10.1007/s10490-010-9232-6.
- Gopinathan, Saravanan, ed. 1998. *Language, Society, and Education in Singapore: Issues and Trends*. 2nd ed. Singapore: Times Academic Press.
- Greenbaum, Sidney, and Gerald Nelson. 1996. "The International Corpus of English (ICE) Project." *World Englishes* 15(1):3–15. doi: 10.1111/j.1467-971X.1996.tb00088.x.
- Gumperz, John J. 1982. *Discourse Strategies*. Cambridge University Press.

- Gupta, Anthea Fraser. 1989. "Singapore Colloquial English and Standard English." *Singapore Journal of Education* 10(2):33–39. doi: 10.1080/02188798908547659.
- Gupta, Anthea Fraser. 1992a. "Contact Features of Singapore Colloquial English." *Sociolinguistics Today: International Perspectives* 323–45.
- Gupta, Anthea Fraser. 1992b. "The Pragmatic Particles of Singapore Colloquial English." *Journal of Pragmatics* 18(1):31–57. doi: 10.1016/0378-2166(92)90106-L.
- Gupta, Anthea Fraser. 1994. *The Step-Tongue: Children's English in Singapore*. Clevedon: Multilingual Matters.
- Gupta, Anthea Fraser. 2001. "English in the Linguistic Ecology of Singapore." *The Cultural Politics of English as a World Language*. Freiburg.
- Gupta, Anthea Fraser. 2010. "Singapore Standard English Revisited." Pp. 57–89 in *English in Singapore: Modernity and Management*, edited by L. Lim, A. Pakir, and L. Wee. Hong Kong and London: Hong Kong University Press.
- Gwee, Yee Hean, and Theodore R. Doraisamy. 1969. *150 Years of Education in Singapore*. Singapore: Teachers Training College.
- Hansen, Beke. 2018. *Corpus Linguistics and Sociolinguistics: A Study of Variation and Change in the Modal Systems of World Englishes*. Leiden and Boston: Brill Rodopi.
- Harris, Alice C., and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge and New York: Cambridge University Press.
- Hashim, Azirah, and Gerhard Leitner. 2011. "Contact Expressions in Contemporary Malaysian English: Contact Expressions in Contemporary Malaysian English." *World Englishes* 30(4):551–68. doi: 10.1111/j.1467-971X.2011.01729.x.
- Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the World's Languages*. München: Lincom Europa.
- Haspelmath, Martin. 2006. "Against Markedness (And What to Replace It With)." *Journal of Linguistics* 42(1):25–70.
- Haspelmath, Martin. 2008. "Loanword Typology: Steps toward a Systematic Crosslinguistic Study of Lexical Borrowability." Pp. 43–62 in *Aspects of language contact: new theoretical, methodological and empirical findings with special focus on romancisation processes*, edited by T. Stolz, D. Bakker, and R. Salas Palomo. Berlin: De Gruyter.
- Haugen, Einar Ingvald. 1950. "The Analysis of Linguistic Borrowing." *Language* 26:210–31.

- Haugen, Einar Ingvald. 1953. *The Norwegian Language in America: A Study in Bilingual Behavior*. Philadelphia: University of Pennsylvania Press.
- Heckmann, Friedrich. 2003. "From Ethnic Nation to Universalistic Immigrant Integration: Germany." Pp. 45–78 in *The Integration of Immigrants in European Societies*, edited by F. Heckmann and D. Schnapper. Oldenbourg: De Gruyter.
- Heine, Bernd. 2008. "Contact-Induced Word Order Change without Word Order Change." Pp. 33–60 in *Hamburg Studies on Multilingualism*, edited by P. Siemund and N. Kintana. Amsterdam: John Benjamins Publishing Company.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Heine, Bernd, and Tania Kuteva. 2003. "On Contact-Induced Grammaticalization." *Studies in Language* 27(3):529–72. doi: 10.1075/sl.27.3.04hei.
- Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge and New York: Cambridge University Press.
- Heng, Derek Thiam Soon, and Syed Muhd Khairudin Aljunied, eds. 2009. *Reframing Singapore: Memory, Identity, Trans-Regionalism*. Amsterdam: Amsterdam University Press.
- Hengeveld, Kees. 1992. *Non-Verbal Predication: Theory, Typology, Diachrony*. Berlin and New York: De Gruyter.
- Herring, Susan. 1991. "The Grammaticalization of Rhetorical Questions in Tamil." Pp. 253–84 in *Approaches to Grammaticalization: Theoretical and methodological issues*, edited by E. C. Traugott and B. Heine. Amsterdam: John Benjamins.
- Hickey, Raymond. 1990. "Suprasegmental Transfer: On Prosodic Traces of Irish in Irish English." Pp. 219–29 in *Further Insights into Contrastive Analysis*, edited by J. Fisiak. Amsterdam: John Benjamins.
- Hickey, Raymond. 2010. "Contact and Language Shift." Pp. 149–69 in *The Handbook of Language Contact*. Oxford: Wiley-Blackwell.
- Hickey, Raymond. 2012. "English in Ireland." Pp. 79–108 in *Areal Features of the Anglophone World*, edited by R. Hickey. Berlin and Boston: De Gruyter.
- Hiramoto, Mie. 2015. "Sentence-Final Adverbs in Singapore English and Hong Kong English: Sentence-Final Adverbs in Singapore English and Hong Kong English." *World Englishes* 34(4):636–53. doi: 10.1111/weng.12157.

- Ho, Mei Kei. 2017. "Current Population Estimates, Malaysia, 2016-2017." *Department of Statistics, Malaysia*. Retrieved March 24, 2020 (https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=155&bul_id=a1d1UTFZazd5ajJiRWFHNDduOXFFQT09&menu_id=L0pheU43NWJwRWV SZklWdzQ4TlhUUT09).
- Ho, Mian Lian, and John Talbot Platt. 1993. *Dynamics of a Contact Continuum: Singaporean English*. Oxford: Clarendon Press.
- Ho, Mian Lian, and Irene F. H. Wong. 2001. "The Use of Ever in Singaporean English." *World Englishes* 20(1):79–87. doi: 10.1111/1467-971X.00197.
- Ho, Stephanie. 2013a. "Arab Community." *National Library Board*. Retrieved December 18, 2019 (http://eresources.nlb.gov.sg/infopedia/articles/SIP_2013-08-19_183028.html).
- Ho, Stephanie. 2013b. "Eurasian Community." *National Library Board*. Retrieved December 28, 2019 (http://eresources.nlb.gov.sg/infopedia/articles/SIP_2013-08-06_115106.html).
- Hock, Hans Henrich, and Brian D. Joseph. 1996. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin and New York: De Gruyter.
- Hole, Daniel. 2011. "The Deconstruction of Chinese Shì...de Clefts Revisited." *Lingua* 121(11):1707–33. doi: 10.1016/j.lingua.2011.07.004.
- Hole, Daniel, and Malte Zimmermann. 2013. "Cleft Partitionings in Japanese, Burmese and Chinese." Pp. 285–318 in *Linguistics Today*, edited by K. Hartmann and T. Veenstra. Amsterdam: John Benjamins.
- Holmes, Janet, and Nick Wilson. 2017. *An Introduction to Sociolinguistics*. 5th Edition. London and New York: Routledge.
- Hönekopp, Elmar. 2002. "Arbeitsmärkte Und Arbeitnehmerfreizügigkeit." Pp. 125–36 in *Die Einheit Europas. Zur politischen, wirtschaftlichen und sozialen Dimension der EU-Erweiterung*, edited by R. Langewiesche. Münster: Verlag Westfälisches Dampfboot.
- Hoon, Chng Huang. 2003. "'You See Me No up': Is Singlish a Problem?" *Language Problems and Language Planning* 27(1):45–62. doi: 10.1075/lplp.27.1.04hoo.
- Hopper, Paul J. 1991. "On Some Principles of Grammaticization." P. 17 in *Approaches to Grammaticalization: Theoretical and methodological issues*, edited by E. C. Traugott and B. Heine. Amsterdam: John Benjamins.

- Hopper, Paul J., and Elizabeth Closs Traugott. 2006. *Grammaticalization*. 2. ed., 3. print. Cambridge: Cambridge Univ. Press.
- Hornberger, Nancy H. 1988. *Bilingual Education and Language Maintenance: A Southern Peruvian Quechua Case*. Berlin and New York: De Gruyter.
- van Hout, Roeland, and Pieter Muysken. 1994. "Modeling Lexical Borrowability." *Language Variation and Change* 6(1):39–62. doi: 10.1017/S0954394500001575.
- Huber, Magnus. 2014. "Stylistic and Sociolinguistic Variation in Schneider's Nativization Phase: T -Affrication and Relativization in Ghanaian English." Pp. 86–106 in *The Evolution of Englishes: The Dynamic Model and beyond*, edited by S. Buschfeld, T. Hoffmann, M. Huber, and A. Kautzsch. Amsterdam: John Benjamins.
- Huber, Magnus, and Viveka Velupillai, eds. 2007. *Synchronic and Diachronic Perspectives on Contact Languages*. Amsterdam: John Benjamins.
- Huddleston, Rodney D., and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge and New York: Cambridge University Press.
- Hundt, Marianne, and Christian Mair. 1999. "Agile and Uptight Genres: The Corpus-Based Approach to Language Change in Progress." *International Journal of Corpus Linguistics* 4(2):221–42.
- Hunger, Uwe, and Dietrich Thränhardt. 2004. "Migration Und Bildungserfolg: Wo Stehen Wir?" *IMIS-BEITRÄGE* 23:179–97.
- IBP. 2015. *Singapore Criminal Laws, Regulations and Procedures Handbook*. Washington and Singapore: International Business Publications.
- ICA, Singapore. 2011. "ICA - Greater Flexibility with Implementation of Double-Barrelled Race Option from 1 January 2011." Retrieved March 8, 2020 (http://www.ica.gov.sg/news_details.aspx?nid=12443).
- IMB Corp. 2020. *IBM SPSS Statistics for Macintosh, Version 27.0*. Armonk, NY: IBM Corp.
- Israel, Michael. 1995. "Negative Polarity and Phantom Reference." Pp. 162–73 in *Annual Meeting of the Berkeley Linguistics Society*. Vol. 21. Berkeley: The Linguistic Society of America.
- Israel, Michael. 1998. "Ever: Polysemy and Polarity Sensitivity." *Linguistic Notes from La Jolla* 19:29–45.
- Jenkins, Jennifer. 2009. *World Englishes: A Resource Book for Students*. 2nd ed. London ; New York: Routledge.

- Jenkins, Jennifer. 2014. *English as a Lingua Franca in the International University: The Politics of Academic English Language Policy*. London: Routledge.
- Jenkins, Jennifer, Alessia Cogo, and Martin Dewey. 2011. "Review of Developments in Research into English as a Lingua Franca." *Language Teaching* 44(3):281–315. doi: 10.1017/S0261444811000115.
- Jespersen, Otto. 2003. *Essentials of English Grammar: 25th Impression, 1987*. 25 impression 1987. London: Routledge.
- Johanson, Lars. 1992. *Strukturelle Faktoren in Türkischen Sprachkontakten*. Stuttgart: Steiner.
- Johanson, Lars. 2000. "Linguistic Convergence in the Volga Area." *Studies in Slavic and General Linguistics* 28:165–78.
- Johanson, Lars. 2002. *Structural Factors in Turkic Language Contacts*. London: Routledge.
- Johanson, Lars. 2008. "Remodeling Grammar: Copying, Conventionalization, Grammaticalization." Pp. 61–79 in *Language Contact and Contact Languages*, edited by P. Siemund and N. Kintana. Amsterdam: John Benjamins.
- Kachru, Braj B. 1985. "Standards, Codification, and Sociolinguistic Realism: The English Language in the Outer Circle." Pp. 11–30 in *English in the world: Teaching and learning the language and literatures*. Cambridge: Cambridge University Press.
- Kachru, Braj B. 1986. *The Alchemy of English: The Spread, Functions, and Models of Non-Native Englishes*. Urbana: University of Illinois Press.
- Kachru, Braj B., ed. 1992. *The Other Tongue: English across Cultures*. 2nd ed. Urbana: University of Illinois Press.
- Kachru, Braj B. 2005. *Asian Englishes: Beyond the Canon*. Hong Kong: Hong Kong University Press.
- Kaske, Elisabeth. 2008. *The Politics of Language in Chinese Education, 1895-1919*. Leiden and Boston: Brill.
- Khong, Cho-oon. 1995. "Singapore: Political Legitimacy through Managing Conformity." Pp. 108–35 in *Political legitimacy in Southeast Asia: the quest for moral authority*. Stanford University Press.
- Khoo, Chian Kim, ed. 1983. *Economic and Social Statistics, Singapore, 1960-1982*. Singapore: Dept. of Statistics.
- Kiang, Clyde. 1991. *The Hakka Search for a Homeland*. Elgin: Allegheny Press.

- Klein, Wolfgang. 1994. *Time in Language*. London and New York: Routledge.
- Kohlmeier, Manfred, and Peter Schimany. 2005. *Der Einfluss von Zuwanderung Auf Die Deutsche Gesellschaft: Deutscher Beitrag Zur Pilotforschungsstudie "The Impact of Immigration on Europe's Societies" Im Rahmen Des Europäischen Migrationsnetzwerk*. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- König, Ekkehard. 1991. *The Meaning of Focus Particles: A Comparative Perspective*. London and New York: Routledge.
- König, Ekkehard. 2017. "Syntax and Semantics of Additive Focus Markers from a Cross-Linguistic Perspective: A Tentative Assessment of the State of the Art." Pp. 23–44 in *Pragmatics and Beyond New Series*, edited by A.-M. De Cesare and C. Andorno. Amsterdam: John Benjamins.
- Kramer, Rajia. 2018. "Position Paper on Phasal Polarity Expressions." Hamburg: University of Hamburg.
- Krashen, Stephen D., Lucy Tse, and Jeff McQuillan. 1998. *Heritage Language Development*. Culver: Language Education Associates.
- Kroeger, Paul. 2018. *Analyzing Meaning: An Introduction to Semantics and Pragmatics*. Berlin: Language Science Press.
- Kulick, Don. 2004. *Language Shift and Cultural Reproduction: Socialization, Self, and Syncretism in a Papua New Guinean Village*. Cambridge: Cambridge University Press.
- Kuo, Eddie C. Y. 1983. "Literacy in Singapore, 1970 -1980." *RELC Journal* 14(1):1–17. doi: 10.1177/003368828301400101.
- Kuteva, Tania, Seongha Rhee, Debra Ziegeler, and Jessica Sabban. 2018. "On Sentence-Final 'What' in Singlish: Are You the Queen of England, or What?" *Journal of Language Contact* 11(1):32–70. doi: 10.1163/19552629-01101002.
- Kwan-Terry, A. 1978. "The Meaning and the Source of the 'La' and the 'What' Particles in Singapore English." *RELC Journal* 9(2):22–36. doi: 10.1177/003368827800900202.
- Kwan-Terry, Anna. 1989. "The Specification of Stage by a Child Learning English and Cantonese Simultaneously: A Study of Acquisition Processes." Pp. 33–48 in *Interlingual processes*, edited by H. W. Dechert and M. Raupach. Tübingen: Gunter Narr.
- Kyto, Merja, and Paivi Pahta, eds. 2016. *The Cambridge Handbook of English Historical Linguistics*. Cambridge: Cambridge University Press.

- Labov, William. 1963. "The Social Motivation of a Sound Change." *WORD* 19(3):273–309. doi: 10.1080/00437956.1963.11659799.
- Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Digital print. Malden, Mass.: Blackwell.
- Ladusaw, William A. 1980. *Polarity Sensitivity as Inherent Scope Relations*. New York: Garland Pub.
- Ladusaw, William A. 2008. "On the Notion Affective in the Analysis of Negative-Polarity Items." Pp. 457–70 in *Formal Semantics*, edited by P. Portner and B. H. Partee. Oxford: Blackwell.
- Laitin, David D. 1993. "Migration and Language Shift in Urban India." *International Journal of the Sociology of Language* 103(1). doi: 10.1515/ijsl.1993.103.57.
- Lang, Jürgen. 2012. "Dangers of the Linguistic Feature Pool." Pp. 1–16 in *Associação Brasileira de Estudos Crioulos e Similares*. São Paulo: Associação de Crioulos de Base Lexical Portuguesa e Espanhola.
- Lange, Claudia. 2012. *The Syntax of Spoken Indian English*. Amsterdam: John Benjamins.
- Lee, Kuan Yew. 2000. *From Third World to First: The Singapore Story: 1965-2000*. New York: HarperCollins.
- Lehmann, Christian. 2015. *Thoughts on Grammaticalization*. 3rd edition. Berlin: Language Science Press.
- Leimgruber, Jakob R. E. 2009. "Modelling Variation in Singapore English." PhD Thesis, Pembroke College, Oxford.
- Leimgruber, Jakob R. E. 2013. *Singapore English: Structure, Variation, and Usage*. Cambridge and New York: Cambridge University Press.
- Leimgruber, Jakob R. E. 2015. "Bah in Singapore English." *World Englishes* 35(1):78–97. doi: 10.1111/weng.12175.
- Leimgruber, Jakob R. E., Peter Siemund, and Laura Terassa. 2018. "Singaporean Students' Language Repertoires and Attitudes Revisited." *World Englishes* 37(2):282–306. doi: 10.1111/weng.12292.
- Lenk, Uta. 1998. *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*. Tübingen: G. Narr.
- LePoer, Barbara Leitch, ed. 1991. *Singapore: A Country Study*. 2nd ed. Washington, D.C.: Federal Research Division, Library of Congress.

- Ler, Soon Lay Vivien. 2006. "A Relevance-Theoretic Approach to Discourse Particles in Singapore English." Pp. 149–66 in *Approaches to discourse particles, Studies in pragmatics*, edited by K. Fischer. Amsterdam and London: Elsevier.
- Levshina, Natalia. 2015. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam; Philadelphia: John Benjamins.
- Lewis, Diana M. 2014. "Discourse Patterns in the Development of Discourse Markers in English." *Functions of Language* 21(1):95–118. doi: 10.1075/fol.21.1.06lew.
- Lewis, Diana M. 2018. "Grammaticalizing Connectives in English and Discourse Information Structure." Pp. 135–57 in *New Trends in Grammaticalization and Language Change*, edited by S. Hancil, T. Breban, and J. V. Lozano. Amsterdam: John Benjamins.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Li, Charles N., Sandra A. Thompson, and R. McMillan Thompson. 1982. "The Discourse Motivation for the Perfect Aspect." P. 19 in *Tense-Aspect*, edited by P. J. Hopper. Amsterdam: John Benjamins.
- Li, Lijun. 2014. "The Use of 'already' in Singapore English in Relation to the Chinese Particle 'Le.'" Master Thesis, University of Hamburg, Hamburg.
- Li, Lijun, Eliane Lorenz, and Peter Siemund. 2021. "Ages of Pragmatic Particles in Colloquial Singapore English – A Corpus-Based Study." in *Conference: New Ways of Analyzing Variation – Asia Pacific 6*. Singapore: National University of Singapore.
- Li, Lijun, and Peter Siemund. 2021. "From Phasal Polarity Expression to Aspectual Marker: Grammaticalization of 'already' in Asian and African Varieties of English." Pp. 515–44 in *The expression of phasal polarity in African languages*, edited by K. Raija. Berlin: De Gruyter.
- Li, Yuanjin. 1991. *Lin Wenqing Di Si Xiang: Zhong Xi Wen Hua Di Hui Liu Yu Mao Dun* "The Thought of Lim Boon Kong : Convergence and Contradiction between Chinese and Western Culture." Singapore: Xinjiapo Yazhou yan jiu xue hui.
- Lim, Lisa, ed. 2004. *Singapore English: A Grammatical Description*. Amsterdam and Philadelphia: John Benjamins.
- Lim, Lisa. 2007. "Mergers and Acquisitions: On the Ages and Origins of Singapore English Particles." *World Englishes* 26(4):446–73. doi: 10.1111/j.1467-971X.2007.00522.x.
- Lim, Lisa. 2009. "Not Just an 'Outer Circle', 'Asian' English: Singapore English and the Significance of Ecology." Pp. 179–206 in *World Englishes - Problems, Properties*

- and Prospects: Selected paper from the 13th IAWC conference, Varieties of English Around the World*, edited by Hoffmann, Thomas; Siebers, Lucia. Amsterdam: John Benjamins.
- Lim, Lisa. 2010. "Migrants and 'Mother Tongues': Extralinguistic Forces in the Ecology of English in Singapore." Pp. 19–54 in *English in Singapore, Asian Englishes today*, edited by L. Lim, A. Pakir, and L. Wee. Hong Kong and London: Hong Kong University Press.
- Lim, Lisa, Anne Pakir, and Lionel Wee. 2010. *English in Singapore: Modernity and Management*. edited by L. Lim, A. Pakir, and L. Wee. Hong Kong; London: Hong Kong University Press; Eurospan [distributor].
- Liu, Jane, and Marilyn Carpenter. 2005. "Trends and Issues of Women's Education in China." *The Clearing House* 78(6):277–81.
- Löbner, Sebastian. 1989. "German Schon - Erst - Noch: An Integrated Analysis." *Linguistics and Philosophy* 12(2):167–212. doi: 10.1007/BF00627659.
- Lock, Graham. 1982. "The Chinese Language in Singapore: Status and Features." *Language Learning and Communication* 1(3):233–344.
- Low, Ee Ling. 2010. "English in Singapore and Malaysia: Similarities and Differences." Pp. 224–46 in *The Routledge handbook of world Englishes*, edited by A. Kirkpatrick. London and New York: Routledge.
- Low, Ee Ling, and Adam Brown. 2005. *English in Singapore: An Introduction*. Singapore and New York: McGraw Hill.
- Lyons, John. 1981. *Language and Linguistics: An Introduction*. Cambridge and New York: Cambridge University Press.
- Mackey, William F. 1962. "The Description of Bilingualism." *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 7(2):51–85. doi: 10.1017/S0008413100019393.
- Mair, Christian. 2009. "Corpus Linguistics Meets Sociolinguistics: The Role of Corpus Evidence in the Study of Sociolinguistic Variation and Change" edited by A. Renouf and A. Kehoe. *Language and Computer* 69(1):7–32. doi: 10.1163/9789042025981_003.
- Mair, Christian. 2013. "The World System of Englishes: Accounting for the Transnational Importance of Mobile and Mediated Vernaculars." *English World-Wide* 34(3):253–78. doi: 10.1075/eww.34.3.01mai.

- Mair, Christian. 2017. "Crisis of the 'Outer Circle'? – Globalisation, the Weak Nation State, and the Need for New Taxonomies in World Englishes Research." in *Changing English*, edited by M. Filppula, J. Klemola, A. Mauranen, and S. Vetchinnikova. Berlin and Boston: De Gruyter.
- Malay Language Council. 2019. "Bulan Bahasa." Retrieved March 20, 2020 (<https://www.languagecouncils.sg/mbms/en/bulan-bahasa/bulan-bahasa>).
- Martin, Mayo. 2015. "Mr Lee Kuan Yew among 50 Influential Peranakans Featured in Showcase." *Today Weekend*. Retrieved December 27, 2019 (<http://www.todayonline.com/entertainment/arts/mr-lee-kuan-yew-among-50-influential-peranakans-featured-showcase>).
- Matisoff, James A. 1973. "Tonogenesis in Southeast Asia." *Consonant Types and Tone* 1.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- Matthews, Stephen, and Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.
- Matthews, Stephen, and Virginia Yip. 2009. "Contact-Induced Grammaticalization: Evidence from Bilingual Acquisition." *Studies in Language* 33(2):366–95. doi: 10.1075/sl.33.2.06mat.
- McCawley, James D. 1971. "Tense and Time Reference in English." Pp. 97–113 in *Studies in Linguistic Semantics*, edited by C. J. Fillmore and D. T. Langendoen. New York: Holt, Rinehart, and Winston.
- McCoard, Robert W. 1978. *The English Perfect: Tense-Choice and Pragmatic Inferences*. Amsterdam and New York: Elsevier.
- McWhorter, John H. 2012. *What Language Is: And What It Isn't and What It Could Be*. New York: Gotham Books.
- Miksic, John N. 2013. *Singapore & the Silk Road of the Sea, 1300-1800*. Singapore: NUS Press, National University of Singapore : National Museum of Singapore.
- Milroy, James. 1997. "Internal vs External Motivations for Linguistic Change." *Multilingua - Journal of Cross-Cultural and Interlanguage Communication* 16(4):311–24. doi: 10.1515/mult.1997.16.4.311.
- Mirhosseini, Seyyed-Abdolhamid, and Parisa Abazari. 2016. "'My Language Is Like My Mother': Aspects of Language Attitudes in a Bilingual Farsi-Azerbaijani Context in Iran." *Open Linguistics* 2(1). doi: 10.1515/opli-2016-0018.

- Mobini-Kesheh, Natalie. 1999. *The Hadrami Awakening: Community and Identity in the Netherlands East Indies, 1900-1942*. Ithaca: Southeast Asia Program.
- Moravcsik, Edith A. 1975. "Verb Borrowing." *Wiener Linguistische Gazette* 8:3–30.
- Moravcsik, Edith A. 1978. "Universals of Language Contact." Pp. 93–122 in *Universals of human language*.
- MRG. 2017. "World Directory of Minorities and Indigenous Peoples: Singapore-Indians." Retrieved December 27, 2019 (<http://minorityrights.org/minorities/indians-3/>).
- Mufwene, Salikoko. 2008. *Language Evolution: Contact, Competition and Change*. London and New York: Continuum International Publishing Group.
- Mufwene, Salikoko S. 1990. "Transfer and the Substrate Hypothesis in Creolistics." *Studies in Second Language Acquisition* 12(1):1–23. doi: 10.1017/S0272263100008718.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
- Narrog, Heiko, Seongha Rhee, and John Whitman. 2018. "Grammaticalization in Japanese and Korean." Pp. 166–88 in *Grammaticalization from a typological perspective*.
- Nathan, J. E. 1922. *The Census of British Malaya 1921*. London: Waterlow and Sons.
- National Library Board. 2014. "Singapore Free School Is Opened - Singapore History." *History SG*. Retrieved December 8, 2020 (<https://eresources.nlb.gov.sg/history/events/34e52586-72fa-4d0c-9551-a9fd86cb89ab>).
- National Library Board. 2017. *HISTORYSG: An Online Resource Guide*. Singapore: NLB Singapore.
- National Museum of Singapore. 2019. *Investigating History: Singapore under the Japanese Occupation 1942-1945*. Singapore: National Museum of Singapore.
- Nelson, Gerald. 2002. "International Corpus of English: The Singapore Corpus." Retrieved February 15, 2020 (http://courses.washington.edu/hypertext/ice-cwb/Manuals/ICE-SIN_Manual.PDF).
- Newbold, Thomas John. 1839. *Political and Statistical Account of the British Settlements in the Straits of Malacca*. London: John Murray.

- Ng, Chin Bee, and Francesco Cavallaro. 2019. "Multilingualism in Southeast Asia: The Post-Colonial Language Stories of Hong Kong, Malaysia and Singapore." Pp. 27–50 in *Multidisciplinary Perspectives on Multilingualism*, edited by S. Montanari and S. Quay. Berlin and Boston: De Gruyter.
- Ng, Patrick Chin Leong. 2011. "Language Planning in Action: Singapore's Multilingual and Bilinugal Policy." *Risumeikan Asia Pacific Journal* 30:1–12.
- Ng, Patrick Chin Leong. 2017. *A Study of Attitudes of Dialect Speakers Towards the Speak Mandarin Campaign in Singapore*. Singapore: Springer Singapore.
- Nga, Nguyen Thuy. 2017. "English Borrowings and Scale of Borrowability in Vietnamese Magazines." *VNU Journal of Foreign Studies* 33(3). doi: 10.25073/2525-2445/vnufs.4154.
- OHI-NAS. 2020. "Oral History Interviews." *National Archives of Singapore*. Retrieved February 13, 2020 (https://www.nas.gov.sg/archivesonline/oral_history_interviews).
- Olsson, Bruno. 2013. "Iamitives: Perfects in Southeast Asia and Beyond." Stockholm University, Stockholm.
- Omar, Asmah Haji. 1977. *The Phonological Diversity of the Malay Dialects. Agian Pebinaan Dan Pengembangan Bahasa*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Omar, Asmah Haji. 1983. *The Malay Peoples of Malaysia and Their Languages*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Omoniyi, Tope, and Mukul Saxena. 2010. "Chapter 1. Introduction." Pp. 1–22 in *Contending with Globalization in World Englishes*, edited by M. Saxena and T. Omoniyi. Bristol, Blue Ridge Summit: Multilingual Matters.
- Onysko, Alexander. 2007. *Anglicisms in German: Borrowing, Lexical Productivity, and Written Codeswitching*. Berlin and New York: Walter de Gruyter.
- Pakir, Anne. 1994. "Education and Invisible Language Planning: The Case of English in Singapore." Pp. 158–79 in *English and Language Planning: A Southeast Asian Contribution*.
- Pakir, Anne. 2004. "English-Knowing Bilingualism in Singapore." Pp. 254–78 in *Imagining Singapore*, edited by B. K. Choon, A. Pakir, and T. C. Kiong. Singapore: Eastern Universities Press by Marshall Cavendish.
- Panagiotidis, Phoivos. 2003. "One, Empty Nouns, and Theta-Assignment." *Linguistic Inquiry* 34(2):281–92.

- Paul, Waltraud, and John Whitman. 2008. "Shi ... de Focus Clefts in Mandarin Chinese." *The Linguistic Review* 25(3–4):413–51. doi: 10.1515/TLIR.2008.012.
- Payne, Thomas Edward. 2006. *Exploring Language Structure: A Student's Guide*. Cambridge, UK ; New York: Cambridge University Press.
- Peng, Chia Oai. 2012. "Chinese Education in Southeast Asia." Pp. 446–58 in *Routledge Handbook of the Chinese Diaspora*, edited by C.-B. Tan. London: Routledge.
- Pietsch, Lukas. 2009a. *English in Ireland: Grammar in Language Contact*. Hamburg: University of Hamburg.
- Pietsch, Lukas. 2009b. "Hiberno-English Medial-Object Perfects Reconsidered: A Case of Contact-Induced Grammaticalisation." *Studies in Language* 33(3):528–68. doi: 10.1075/sl.33.3.02pie.
- Pillai, Anitha Devi. 2018. "Malayalee Community and Culture in Singapore." Pp. 265–302 in *The Singapore Ethnic Mosaic*, edited by M. Mathew. Singapore: World Scientific.
- Platt, John T. 1975. "The Singapore English Speech Continuum and Its Basilect 'Singlish' as a 'Creoloid.'" *Anthropological Linguistics* 17(7):363–74.
- Platt, John T. 1978. "Sociolects and Their Pedagogical Implications." *RELC Journal* 9(1):28–38. doi: 10.1177/003368827800900104.
- Platt, John T., and Mian Lian Ho. 1989. "Discourse Particles in Singaporean English: Substratum Influences and Universals." *World Englishes* 8(2):215–21. doi: 10.1111/j.1467-971X.1989.tb00656.x.
- Platt, John T., and Heidi Weber. 1980. *English in Singapore and Malaysia: Status, Features, Functions*. Kuala Lumpur and New York: Oxford University Press.
- Platt, John Talbot, Heidi Weber, and Mian Lian Ho. 1984. *The New Englishes*. London and Boston: Routledge and Kegan Paul.
- Portes, Alejandro, and Lingxin Hao. 1998. "E Pluribus Unum: Bilingualism and Loss of Language in the Second Generation." *Sociology of Education* 269–94.
- Portes, Alejandro, and Rubén G. Rumbaut. 2001. *Legacies: The Story of the Immigrant Second Generation*. Berkeley: University of California Press.
- Poussa, Patricia. 1982. "The Evolution of Early Standard English: The Creolization Hypothesis." *Studia Anglica Posnaniensia* 14:69–85.
- Prasad, Vina Jie-Min, and Jaime Koh. 2014. "Raffles Institution." *Singapore Infopedia*. Retrieved December 8, 2020 (https://eresources.nlb.gov.sg/infopedia/articles/SIP_17_2004-12-21.html).

- Queen, Robin M. 2001. "Bilingual Intonation Patterns: Evidence of Language Change from Turkish-German Bilingual Children." *Language in Society* 30(1):55–80.
- Rajadurai, Joanne. 2005. "Revisiting the Concentric Circles: Conceptual and Sociolinguistic Considerations." *Asian EFL Journal* 7(4):111–30.
- Ramsey, Samuel Robert. 1989. *The Languages of China*. 2nd print. Princeton: Princeton University Press.
- Rhee, Seongha. 2004. "Grammaticalization and Lexicalization of Rhetorical Questions in Korean." *Studies in Modern Grammar* 35:111–39.
- Richards, Jack C., and Mary Tay. 1977. "The La Particle in Singapore English." Pp. 141–55 in *The English language in Singapore*, edited by W. J. Crewe. Singapore: Eastern Universities Press.
- Rochemont, Michael S. 1986. *Focus in Generative Grammar*. Amsterdam and Philadelphia: John Benjamins.
- Romaine, Suzanne. 1995. *Bilingualism*. 2nd ed. Oxford: Blackwell.
- Ross, Malcolm D. 1991. "Refining Guy's Sociolinguistic Types of Language Change." *Diachronica* 8(1):119–29. doi: 10.1075/dia.8.1.14ros.
- Rubdy, Rani. 2007. "Singlish in the School: An Impediment or a Resource?" *Journal of Multilingual and Multicultural Development* 28(4):308–24. doi: 10.2167/jmmd459.0.
- Rudolph, Jürgen. 1993. "The 'Speak Mandarin Campaign' of the Babas." *Commentary* 11(1):93–102.
- Sadock, Jerrold M. 1991. *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago: University of Chicago Press.
- Salazar, Danica. 2014. "Towards Improved Coverage of Southeast Asian Englishes in the Oxford English Dictionary." *Lexicography* 1(1):95–108. doi: 10.1007/s40607-014-0003-2.
- Sandhu, Kernial Singh, and A. Mani, eds. 1993. *Indian Communities in Southeast Asia*. Singapore: Times Academic Press.
- Sankoff, David. 1988. "Sociolinguistics and Syntactic Variation." Pp. 140–61 in *Linguistics: The Cambridge Survey*, edited by F. J. Newmeyer. Cambridge: Cambridge University Press.
- Sankoff, David, and Suzanne Laberge. 1978. "The Linguistic Market and the Statistical Explanation of Variability." Pp. 239–50 in *Linguistic variation: models and methods*. New York: Academic Press.

- Sapir, Edward. 2008. *Language: An Introduction to the Study of Speech*. Milton Keynes: Lightning Source.
- Saw, Swee-Hock. 1969. "Population Trends In Singapore, 1819-1967." *Journal of Southeast Asian History* 10(1):36–49.
- Saw, Swee-Hock. 2012. *The Population of Singapore*. 3rd ed. Singapore: Institute of Southeast Asian Studies.
- Schneider, Edgar W. 2003. "The Dynamics of New Englishes: From Identity Construction to Dialect Birth." *Language* 79(2):233–81. doi: 10.1353/lan.2003.0136.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the World*. Cambridge and New York: Cambridge University Press.
- Schneider, Edgar W. 2014. "New Reflections on the Evolutionary Dynamics of World Englishes: New Reflections on the Evolutionary Dynamics of World Englishes." *World Englishes* 33(1):9–32. doi: 10.1111/weng.12069.
- Schröter, Verena. 2018. "Verena. 'Where Got Such Thing One?' A Multivariate Analysis of Null Subjects in Asian Englishes." Albert-Ludwigs-Universität Freiburg, Freiburg.
- Schuchardt, Hugo. 1884. *Slawo-Deutsches Und Slawo-Italienisches: Dem Herrn Franz von Miklosich Zum 20. November 1883*. Graz: Leuschner u. Lubensky.
- Schweinberger, Martin. 2012. "The Discourse Marker LIKE in Irish English." Pp. 179–202 in *Varieties of English Around the World*, edited by B. Migge and M. Ní Chiosáin. Amsterdam: John Benjamins.
- Schwenter, Scott A., and Rena Torres Cacoullous. 2008. "Defaults and Indeterminacy in Temporal Grammaticalization: The 'Perfect' Road to Perfective." *Language Variation and Change* 20(1):1–39. doi: 10.1017/S0954394508000057.
- Sharma, Devyani. 2009. "Typological Diversity in New Englishes." *English World-Wide* 30(2):170–95. doi: 10.1075/eww.30.2.04sha.
- Sharma, Devyani. 2012. "Shared Features in New Englishes." Pp. 211–32 in *Areal Features of the Anglophone World*, edited by R. Hickey. Berlin, Boston: DE GRUYTER.
- Shi, Yuzhi. 2002. *The Establishment of Modern Chinese Grammar: The Formation of the Resultative Construction and Its Effects*. Amsterdam: John Benjamins.
- Siemund, Peter. 2008. "Language Contact: Constraints and Common Paths of Contact-Induced Language Change." Pp. 3–11 in *Language contact and contact languages*, edited by N. Kintana and P. Siemund. Amsterdam and Philadelphia: John Benjamins.

- Siemund, Peter. 2013. *Varieties of English: A Typological Approach*. Cambridge: Cambridge University Press.
- Siemund, Peter. 2018. "Modeling World Englishes from a Cross-Linguistic Perspective." Pp. 133–62 in *Varieties of English Around the World*. Vol. G61, edited by S. C. Deshors. Amsterdam: John Benjamins Publishing Company.
- Siemund, Peter, and Lijun Li. 2017. "Towards a Diachronic Reconstruction of Colloquial Singapore English." Pp. 11–32 in *Negation and Contact*, edited by D. Ziegeler and Z. Bao. Amsterdam and Philadelphia: John Benjamins.
- Siemund, Peter, and Lijun Li. 2020. "Multilingualism and Language Policy in Singapore." Pp. 205–28 in *Language Diversity in the Sinophone World: Historical Trajectories, Language Planning, and Multilingual Practices*, edited by H. Klöter and M. S. Saarela. Berlin: Routledge.
- Siemund, Peter, Monika Edith Schulz, and Martin Schweinberger. 2014. "Studying the Linguistic Ecology of Singapore: A Comparison of College and University Students." *World Englishes* 33(3):340–62. doi: 10.1111/weng.12094.
- Sim, Cheryl. 2016. "Bilingual Policy Infopedia Singapore." *Singapore Infopedia*. Retrieved March 13, 2020 (https://eresources.nlb.gov.sg/infopedia/articles/SIP_2016-09-01_093402.html?s=The%20bilingual%20education.).
- Smith, Carlota S. 1997. *The Parameter of Aspect*. Dordrecht: Springer Netherlands.
- Smith, Ian. 1985. "Multilingualism and Diffusion: A Case Study from Singapore English." *Indian Journal of Applied Linguistics* 11(2):105–28.
- 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 & Linguistic Theory* 27(3):623–57. doi: 10.1007/s11049-009-9074-4.
- Soh, Hooi Ling, and Meijia Gao. 2006. "Perfective Aspect and Transition in Mandarin Chinese: An Analysis of Double -Le Sentences." Pp. 107–22 in *Proceedings of the 2004 Texas Linguistics Society Conference*, edited by P. Denis, E. McCready, A. Palmer, and B. Reese. Somerville, Mass.: Cascadilla Proceedings Project.
- Song, Ong Siang. 1984. *One Hundred Years' History of the Chinese in Singapore*. Singapore and New York: Oxford University Press.
- Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge and New York: Cambridge University Press.

- Sun, Chaofen. 1996. *Word-Order Change and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.
- Suntum, Ulrich van, and Dirk Schlotböller. 2002. *Arbeitsmarktintegration von Zuwanderern: Einflussfaktoren, internationale Erfahrungen und Handlungsempfehlungen*. Gütersloh: Bertelsmann-Stiftung.
- Swietochowski, Tadeusz. 1995. *Russia and Azerbaijan: A Borderland in Transition*. New York: Columbia University Press.
- Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Malden, MA: Wiley-Blackwell.
- Tamil Language Council. 2020. "About the Tamil Language Council." Retrieved March 20, 2020 (<https://www.languagecouncils.sg/tamil/en/about/about-the-tamil-language-council>).
- Tan, Charlene. 2012. "'OUR SHARED VALUES' IN SINGAPORE: A CONFUCIAN PERSPECTIVE." *Educational Theory* 62(4):449–63. doi: 10.1111/j.1741-5446.2012.00456.x.
- Tan, Cheryl Lu-Lien. 2016. "The Singlish Language Reflects the Power of My People." *Time*. Retrieved March 20, 2020 (<https://time.com/4441078/singlish-language/>).
- Tan, Chwee Haut. 1989. "Confucianism and Nation Building in Singapore." *International Journal of Social Economics* 16(8):5–16. doi: 10.1108/03068298910133106.
- Tan, Eugene K. B. 2007. "The Multilingual State in Search of the Nation: The Language Policy and Discourse in Singapore's Nation-Building." Pp. 74–117 in *Language, Nation and Development in Southeast Asia*, edited by L. Suryadinata and H. G. Lee. Singapore: ISEAS–Yusof Ishak Institute.
- Tan, Guan Heng. 2008. *100 Inspiring Rafflesians, 1823-2003*. New Jersey: World Scientific.
- Tan, Siew Imm. 2016. "Charting the Endomormative Stabilization of Singapore English." Pp. 69–84 in *Communicating with Asia: The future of English as a global language*, edited by G. Leitner, A. Hashim, and H.-G. Wolf. Cambridge: Cambridge University Press.
- Tan, Soo Jiuan, and Chin Tiong Tan. 2012. "Singapore." Pp. 493–536 in *International business handbook*, edited by V. H. Kirpalani. London: Routledge.
- Tan, Thomas T. W., ed. 1990. *Chinese Dialect Groups: Traits and Trades*. Singapore: Opinion Books.

- Tan, Yingying. 2014. "English as a 'mother Tongue' in Singapore." *World Englishes* 33(3):319–39. doi: 10.1111/weng.12093.
- Tang, Leslie. 2017. "Population By-Census 2016 Main Results." Retrieved March 24, 2020 (<https://www.bycensus2016.gov.hk/data/16bc-main-results.pdf>).
- Tarling, Nicholas, ed. 2012. *Studying Singapore's Past: C.M. Turnbull and the History of Modern Singapore*. Singapore: NUS Press.
- Tay, Kelly, and Andrea Soh. 2015. "Beyond 50: Singapore's Growth Strategy Shifts." *Business Times*. Retrieved (<http://www.businesstimes.com.sg/government-economy/beyond-50-singapores-growth-strategy-shifts>).
- Tay, Mary W. J. 1982. "The Phonology of Educated Singapore English." *English World-Wide* 3(2):135–45. doi: 10.1075/eww.3.2.02tay.
- Teo, Ming Chew. 2014. "Transferability of Semantic Functions in Contact-Induced Grammaticalization: The Case of Singapore Colloquial English *One*." *Studies in Language* 38(4):836–63. doi: 10.1075/sl.38.4.07teo.
- Teo, Ming Chew. 2019. "The Role of Parallel Constructions in Imposition: A Synchronic Study of *Already* in Colloquial Singapore English." *Journal of Pidgin and Creole Languages* 34(2):346–76. doi: 10.1075/jpcl.00042.teo.
- Teo, Ming Chew. 2020. *Crosslinguistic Influence in Singapore English: Linguistic and Social Aspects*. New York: Routledge.
- Terassa, Laura Anna Maria. 2017. "Morphological Simplification in Asian Englishes: Frequency, Substratum Transfer, and Institutionalization." PhD Thesis, Freiburg, University of Freiburg.
- Tessensohn, Denyse. 2001. *Elvis Lived in Katong: Personal Singapore Eurasiana*. Singapore: Dagmar Books.
- The Straits Times. 2004. "Lim Boon Keng Bicultural Broker." Retrieved March 10, 2020 (<eresources.nlb.gov.sg/newspapers/Digitised/Article/straitstimes20040626-1.2.35.1>).
- Thomason, Sarah G., ed. 1997. *Contact Languages: A Wider Perspective*. Vol. 17. Amsterdam: John Benjamins.
- Thomason, Sarah G. 2001. *Language Contact*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G., and Terrence Kaufman. 1991. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

- Tongue, Ray K. 1974. *The English of Singapore and Malaysia*. Singapore: Eastern Universities Press.
- Treffers-Daller. 1999. "Borrowing and Shift Induced-Interference: Contrasting Patterns in French-Germanic Contact in Brussels and Strasbourg." *Bilingualism: Language and Cognition* 2(1):1–22.
- Treffers-Daller, Jeanine. 1991. "Towards a Uniform Approach to Code-Switching and Borrowing." Pp. 259–79 in *Papers for the workshop on constraints, conditions and models*. Strasbourg: European Science Foundation.
- Treffers-Daller, Jeanine. 2009. "Code-Switching and Transfer: An Exploration of Similarities and Differences." Pp. 58–74 in *The Cambridge Handbook of Linguistic Code-switching*, edited by B. E. Bullock and A. J. Toribio. Cambridge: Cambridge University Press.
- Tregonning, Kennedy G. 1972. *A History of Modern Singapore and Malaysia*. Singapore: Eastern Universities Press.
- Trocki, Carl. 1979. *Prince of Pirates: The Temenggongs and the Development of Johor and Singapore 1784-1885*. Singapore: Oxford University Press.
- Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. New York: Oxford University Press.
- Tsui, Amy B. M., and James W. Tollefson. 2017. *Language Policy, Culture, and Identity in Asian Contexts*. London: Routledge.
- Turnbull, Constance Mary. 1989. *A History of Malaysia, Singapore, and Brunei*. Sydney and Boston: Allen & Unwin.
- Turnbull, Constance Mary. 1996. *A History of Singapore: 1819-1988*. 3rd ed. Singapore: Oxford University Press.
- Turnbull, Constance Mary. 2009. *A History of Modern Singapore, 1819-2005*. Singapore: NUS Press.
- Van der Wouden, Ton. 2002. *Negative Contexts: Collocation, Polarity and Multiple Negation*. London: Routledge.
- Van Rooy, Bertus. 2010. "Social and Linguistic Perspectives on Variability in World Englishes." *World Englishes* 29(1):3–20. doi: 10.1111/j.1467-971X.2009.01621.x.
- Wang, Gungwu. 2003. *China and the Chinese Overseas*. Singapore: Eastern Universities Press.

- Wardhaugh, Ronald, and Janet M. Fuller. 2015. *An Introduction to Sociolinguistics*. Seventh edition. Chichester, West Sussex, UK ; Malden, MA: Wiley Blackwell.
- Warner, Anthony. 1982. *Complementation in Middle English and the Methodology of Historical Syntax: A Study of the Wyclifite Sermons*. State College: Pennsylvania State University Press.
- Wee, Lionel. 2004a. "Reduplication and Discourse Particles." Pp. 105–28 in *Singapore English: A grammatical description*, edited by L. Lim. Amsterdam: John Benjamins.
- Wee, Lionel. 2004b. "Singapore English: Morphology and Syntax." Pp. 1058–72 in *A Handbook of Varieties of English: Morphology and Syntax*, edited by B. Kortmann and E. Schneider. Berlin and New York: Mouton de Gruyter.
- Wee, Lionel. 2014. "The Evolution of Singlish: Beyond Phase 5?" Pp. 126–41 in *The Evolution of Englishes: The Dynamic Model and beyond*, edited by S. Buschfeld, T. Hoffmann, M. Huber, and A. Kautzsch. Amsterdam: John Benjamins.
- Wee, Lionel. 2018. *The Singlish Controversy: Language, Culture and Identity in a Globalizing World*. Cambridge University Press.
- Wee, Lionel, and Umberto Ansaldo. 2004. "Nouns and Noun Phrases." Pp. 57–74 in *Singapore English: A grammatical description*, edited by L. Lim. Amsterdam: John Benjamins.
- Wei, Li. 1994. *Three Generations, Two Languages, One Family: Language Choice and Language Shift in a Chinese Community in Britain*. Clevedon and Philadelphia: Multilingual Matters.
- Weinreich, Uriel. 1953. *Languages in Contact, Findings and Problems*. Vol. no. 1. New York: Linguistic Circle of New York.
- Wendel, John, and Patrick Heinrich. 2012. "A Framework for Language Endangerment Dynamics: The Effects of Contact and Social Change on Language Ecologies and Language Diversity." *International Journal of the Sociology of Language* 218:145–66. doi: 10.1515/ijsl-2012-0062.
- Whitney, William D. 1881. "On Mixture in Language." *Transactions of the American Philological Association (1869-1896)* 12:5–26. doi: 10.2307/2935666.
- Wichmann, Søren, and Jan Wohlgemuth. 2008. "Loan Verbs in a Typological Perspective." Pp. 89–121 in *Aspects of language contact: new theoretical, methodological and empirical findings with special focus on romancisation processes*. Berlin and New York: De Gruyter.

- Widodo, Johannes. 2009. "Morphogenesis and Hybridity of Southeast Asian Coastal Cities." Pp. 79–92 in *Southeast Asian Culture and Heritage in a Globalising World: Diverging Identities in a Dynamic Region*, edited by R. Ismail, B. Shaw, and G. L. Ooi. London: Surrey.
- Wilkinson, Richard James. 1912. "The Capture of Malacca." *Journal of the Straits Branch of the Royal Asiatic Society* 61:71–76.
- Winford, Donald. 2005. "Contact-Induced Changes: Classification and Processes." *Diachronica* 22(2):373–427.
- Wong, Jock. 2005. "'Why You so Singlish One?' A Semantic and Cultural Interpretation of the Singapore English Particle One." *Language in Society* 34(02). doi: 10.1017/S0047404505050104.
- Wong, Tessa. 2015. "The Rise of Singlish." *BBC News*. Retrieved March 20, 2020 (<https://www.bbc.com/news/magazine-33809914>).
- Wong, Wee Kim. 2011. *Census of Population 2010. Statistical Release 1, Demographic Characteristics, Education, Language and Religion*. Singapore: Dept. of Statistics.
- Wong, Wee Kim. 2015. "Chinese Resident Population by Age Group, Dialect Group and Sex, 2015." *Data.Gov.Sg*. Retrieved March 5, 2020 (<https://data.gov.sg/dataset/chinese-resident-population-by-age-group-dialect-group-and-sex-2015>).
- Wong, Wee Kim. 2016. "Population Trends 2016." Retrieved February 3, 2019 (http://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/population_and_population_structure/population2016.pdf).
- Wong, Wee Kim. 2019. "Population Trends 2019." Retrieved February 5, 2020 (<https://www.singstat.gov.sg/-/media/files/publications/population/population2019.pdf>).
- Xiang, Ming. 2008. "Plurality, Maximality and Scalar Inferences: A Case Study of Mandarin Dou." *Journal of East Asian Linguistics* 17(3):227–45. doi: 10.1007/s10831-008-9025-9.
- Yang, Xiaolong, and Yicheng Wu. 2019. "A Dynamic Account of *lian...dou* in Chinese Verb Doubling Cleft Construction." *Lingua* 217:24–44. doi: 10.1016/j.lingua.2018.10.007.
- Yap, Foong Ha, Fanny Pik-ling Choi, and Kam-siu Cheung. 2010. "Delexicalizing Di: How a Chinese Noun Has Evolved into an Attitudinal Nominalizer." Pp. 63–92 in *Formal*

- Evidence in Grammaticalization Research*, edited by A. Van linden, J.-C. Verstraete, and K. Davidse. Amsterdam: John Benjamins.
- Yap, Foong Ha, Yi Deng, and Marco Caboara. 2017. "Attitudinal Nominalizer(s) in Chinese: Evidence of Recursive Grammaticalization and Pragmaticization." *Lingua* 200:1–21. doi: 10.1016/j.lingua.2017.07.011.
- Yap, Foong Ha, and Stephen Matthews. 2008. "The Development of Nominalizers in East Asian and Tibeto-Burman Languages." Pp. 309–41 in *Rethinking Grammaticalization: New perspectives*, edited by M. J. López-Couso and E. Seoane. Amsterdam: John Benjamins.
- Yeoh, Brenda, and Weiqiang Lin. 2012. "Rapid Growth in Singapore's Immigrant Population Brings Policy Challenges." *Migration Policy Institute*. Retrieved (<http://www.migrationpolicy.org/article/rapid-growth-singapores-immigrant-population-brings-policy-challenges/>).
- Ziegeler, Debra. 2014. "Replica Grammaticalisation as Recapitulation: The Other Side of Contact." *Diachronica* 31(1):106–41.
- Ziegeler, Debra. 2015. *Converging Grammars: Constructions in Singapore English*. Boston: De Gruyter.
- Ziegeler, Debra. 2016. "Historical Replication in Contact Grammaticalization." Pp. 311–52 in *Aspects of Grammaticalization*, edited by D. Olmen, H. Cuyckens, and L. Ghesquière. Berlin and Boston: De Gruyter.
- Ziegeler, Debra. 2017. "Quantification under Negative Scope in Singapore English." Pp. 171–206 in *Negation and Contact: With special focus on Singapore English*, edited by D. Ziegeler and Z. Bao. Amsterdam: John Benjamins.
- Ziegeler, Debra. 2020. "Changes in the Functions of *Already* in Singapore English: A Grammaticalization Approach." *Journal of Pidgin and Creole Languages* 35(2):293–331. doi: 10.1075/jpcl.00062.zie.
- Ziegeler, Debra. 2021. "Grammatical Change and Diversity in Singapore English." Pp. 267–84 in *Multilingual global cities: Singapore, Hong Kong, Dubai.*, edited by P. Siemund and J. R. E. Leimgruber. Singapore: Routledge.

Appendix

I. Background information of interviewees

ACC ID	Name	Topics	Initials	Born	PloBirth	Gender	Ethnicity	Language	Education	Interview	Length (page)
000001	TOH Man Keong (杜马恭)	Special project	TMK	1921	Malaya	M	Chinese	Hokkien Chinese	medium	1979	130
000002	LOW, Peter Por Tuck (刘坡德)	Political History of Singapore 1945–1965	PL	1929	Singapore	M	Chinese	Chinese	medium	1980	172
000004	Snodgrass, John Anthony	Education in Singapore (Part 1: English)	JAS	1906	Singapore	M	British	English	medium	2002	68
000009	LEE Gek Seng (李玉成)	Political History of Singapore 1945–1965	LGS	1927	Singapore	M	Chinese	Chinese	medium	1980	125
000013	Tan Sri Syed Esa Almenoar	Political History of Singapore 1945–1965	SEA	1918	Singapore	M	Malay	Malay	medium	1980	143
000014	LINGAM, S V	Political History of Singapore 1945–1965	SVL	1909	Singapore	M	Chinese	Cantonese Chinese	medium	1980	78
000016	Lee Kip Lin (李急麟)	Japanese Occupation of Singapore	LKL	1925	Singapore	M	Chinese	Chinese	high	1984	186
000021	Mr Chan Chee Seng (陈志成)	Political History of Singapore 1945–1965	CCS	1932	Singapore	M	Chinese	Cantonese	high	1980	94
000024	FONG Sip Chee (冯摄治)	Political History of Singapore 1945–1965	FSC	1938	Singapore	M	Chinese	Chinese	medium	1980	207
000025	Menon, Kanichat Raghava (Dr)	Japanese Occupation of Singapore	KRM	1907	India	M	Indian	Tamil	high	1982	99
000057	LIEN Ying Chow (连赢洲)	Pioneers of Singapore	LYC	1906	China	M	Chinese	Teochew Chinese	medium	1981	244
000071	CHONG Kim Meng	Education in Singapore (Part 2: Chinese)	CKM	1922	Malaya	M	Peranakan	Chinese	high	1981	182
000081	Kannusamy s/o Pakirisamy	Communities of Singapore (Part 2)	K	1914	Singapore	M	Indian	Tamil	medium	1983	466
000095	BAKER, Maurice (Dr)	Political History of Singapore 1945–1965	MB	1920	Malaya	M	Eurasian	English	high	1993	66
000107	Neoh Teik Hong	Japanese Occupation of Singapore	NTH	1927	Singapore	M	Chinese	Chinese	medium	1983	40

A historical sociolinguistic reconstruction of CSE

000114	NEILL, James Desmond Howard	The Public Service	JDHN	1923	Fiji	M	British	English	high	1981	137
000123	BANA, Nazeem	Pioneers of Singapore	NB	1933	India	M	Indian	Tamil	low	1981	59
000133	OTHMAN Wok	Political History of Singapore 1945–1965	OW	1924	Singapore	M	Malay	Malay	high	1982	187
000166	BLOODWORTH, Dennis	Political History of Singapore 1945–1965	DB	1919	UK	M	British	English	medium	1982	251
000205	WARREN, Stanley	Japanese Occupation of Singapore	SW	1917	UK	M	British	English	medium	1982	87
000211	Lee Kim Tah (李金塔)	Pioneers of Singapore	LKT	1902	Singapore	M	Chinese	Hokkien Chinese	medium	1982	125
000213	QUAH, Elsie	Women Through The Years: Economic & Family Lives	EQ	1899	Singapore	F	Chinese	Hokkien	low	1982	40
000237	Hwang Chung Yun	Education in Singapore (Part 1: English)	HCY	1905	China	F	Chinese	Cantonese Chinese	medium	1982	256
000242	ISA, Ibrahim	The Public Service	II	1915	Malaya	M	Malay	Malay	medium	1983	337
000259	Josey, Alex	Political History of Singapore 1945–1965	AJ	1915	UK	M	British	English	low	1983	276
000263	SHANMUGASIVANATHAN	Japanese Occupation of Singapore	S	1927	India	M	Indian	Tamil	low	1983	71
000265	LEE Tian Soo	Japanese Occupation of Singapore	LTS	1925	Singapore	M	Chinese	Chinese	low	1983	98
000284	AZIZ bin Rahim Khan Surattee	Japanese Occupation of Singapore	ARKS	1926	Singapore	M	Malay	Malay	low	1983	67
000296	LELAH, Albert	Communities of Singapore (Part 1)	AL	1913	Iraq	M	Other	Arabic	low	1983	120
000310	SOH Guan Bee	Japanese Occupation of Singapore	SGB	1928	Singapore	M	Chinese	Chinese	medium	1983	162
000316	YAP Siong Eu (叶尚友)	Economic Development of Singapore	YSE	1920	Singapore	M	Chinese	Hokkien	high	1983	107
000345	THEVATHASAN, Gnanasundram	Communities of Singapore (Part 2)	AST	1920	Ceylon	F	Indian	Tamil	low	1983	651
000350	Benjamin Ponnuthurai Alfreds	Communities of Singapore (Part 1)	BPA	1917	India	M	Indian	Tamil	high	1984	381
000362	Ng Kim Boon, Jack	Japanese Occupation of Singapore	JNKB	1926	Malaya	M	Chinese	Chinese	medium	1983	107
000371	Tan, Beng Neo	Women Through The Years: Economic & Family Lives	TBN	1914	Singapore	F	Chinese	English	high	1983	384
000374	GAY Wan Guay (倪员外)	Japanese Occupation of Singapore	GWG	1915	Singapore	M	Chinese	Chinese	high	1984	267
000404	White, Roger	Special project	RW	1948	UK	M	British	English	medium	1984	41

000416	TAN Cheng Hwee	Japanese Occupation of Singapore	TCH	1919	Singapore	M	Chinese	Chinese	medium	1984	170
000454	SOH Teow Seng (苏朝成)	Chinese Dialect Groups	STS	1923	Singapore	M	Chinese	Hokkien	medium	1984	94
000462	Chu Shuen Choo (朱纯藻)	Japanese Occupation of Singapore	CSC	1921	Malaysia	F	Chinese	Chinese	medium	1985	163
000483	TAN Teck Chye Victor	Japanese Occupation of Singapore	TTC	1936	Singapore	M	Chinese	Chinese	medium	1984	62
000485	Ling Lee Hua (林理化)	Japanese Occupation of Singapore	LLH	1922	Malaysia	M	Chinese	Chinese	medium	1985	32
000526	LIM Kim San (林金山)	Economic Development of Singapore	LKS	1916	Singapore	M	Chinese	Chinese	high	1985	279
000838	SEAH Mui Kok (余美国)	History of the Labour Movement	SMK	1923	Singapore	M	Chinese	Chinese	medium	1988	89
000852	Ng Seng Mun	Chinese Dialect Groups	NSM	1927	China	M	Chinese	Cantonese Chinese	medium	1987	133
000920	HENSON, Alexander	Prisoners-of-War (POWs)	AH	1913	Singapore	M	Eurasian	English and Cantonese	medium	1988	86
001109	Minnie Nunes	Education in Singapore (Part 1: English)	MN	1907	Singapore	F	Peranakan	Hokkien Chinese and Malay	medium	1994	111
001255	Mohamed Sidek bin Siraj	Communities of Singapore (Part 3)	MSS	1912	Malaysian	M	Malay	Malay	medium	1991	65
001309	Michael Gorrie	The Public Service	MG	1923	Pakistan	M	Eurasian	Urdu, Pushtu, Tamil	high	1991	93
001423	Vernon Palmer	The Public Service	VP	1925	Singapore	M	Eurasian	Scottish, Siamese, Burmese, German	high	1993	104
001478	Sathyeti Ram Swami Naidu	The Public Service	SRSD	1923	Singapore	M	Indian	Tamil	medium	1994	57
001536	Leela Raj	Education in Singapore (Part 1: English)	LR	1934	India	F	Indian	Tamil	medium	1994	265
001599	Lu Yaw	Education in Singapore (Part 1: English)	LY	1918	China	M	Chinese	Chinese	medium	1994	233
001600	Reverend Neivelle Tan	Special project	NT	1940	Singapore	M	Peranakan	Hokkien	high	1995	549
001613	Comber, Leon	Special Project	LC	1921	UK	M	British	English	medium	1995	91

A historical sociolinguistic reconstruction of CSE

001631	Oh Choo Neo, Patricia	Education in Singapore (Part 1: English)	OCN	1924	Singapore	F	Chinese	Hokkien, Mandarin	medium	1995	316
001632	Chua Hui Neo	Women Through The Years: Economic & Family Lives	CHN	1939	Singapore	F	Chinese	Baba Malay, Hokkien, English	low	1995	79
001663	Mohamed Siraj @ Khatijun Nissa Siraj	Women Through The Years: Economic & Family Lives	RZ	1925	Singapore	F	Other	Tamil, Hindi, Urdu, Malay	medium	1995	548
001880	Lim Kwee Neo	Women Through The Years: Economic & Family Lives	LKN	1921	Singapore	F	Peranakan	Hokkien, Malay	medium	1997	80
001915	Sumitera Bte Mohd Letak	Medical Services in Singapore	SML	1942	Singapore	F	Malay	Malay	low	1997	112
001917	Lee Miu Ling	Women Through The Years: Economic & Family Lives	LML	1992	Singapore	F	Chinese	Cantonese, English, Hokkien, Malay	medium	1997	142
001953	LEO Ah Sin	Special Project	LAS	1940	Singapore	M	Chinese	Chinese	low	1997	170
001970	Ng Joo Kee	Special Project	NJK	1959	Singapore	M	Chinese	Chinese	high	1997	42
002017	Foo Kee Seng	Story of Joo Chiat Changing Landscapes & Community	FKS	1916	Singapore	M	Peranakan	Hainanese, Hokkien, Teochew, Cantonese	medium	1998	179
002027	Dr. Wong Yip Chong	Medical Services in Singapore	WYC	1929	Singapore	M	Chinese	Hokkien, Cantonese, Teochew	medium	1998	147
002038	Dr. Tan Nghoh Chuan	Medical Services in Singapore	TNC	1931	Malaya	M	Peranakan	Hokkien, English	high	1999	99

002039	Shuki Ram Chandra nee Davies	Story of Joo Chiat Changing Landscapes & Community	SRC	1929	Singapore	F	Indian	English, Tamil, Hokkien, Malay	medium	1998	380
002044	Mrs Rita Fernando [& Mr Rennie Fernando]	Story of Joo Chiat Changing Landscapes & Community	RF	1932	Singapore	F	Malay	Malay	high	1998	141
002057	DE CONCEICAO, Aloysius Leo	Story of Joo Chiat Changing Landscapes & Community	AC	1928	Singapore	M	Eurasian	Hokkien, Baba-Malay Bahasa Melayu, Malay	medium	1999	168
002068	Lim Kee Chan	Story of Joo Chiat Changing Landscapes & Community	LKC	1932	Singapore	M	Chinese	Hokkien Chinese	medium	1998	129
002107	Philip Chew Peng Kia (周炳锐)	The Public Service	CPK	1935	Singapore	M	Chinese	Hokkien Chinese, Cantonese Chinese	medium	1999	83
002108	Richard Tan Swee Guan (陈瑞源)	Story of Joo Chiat Changing Landscapes & Community	TSG	1955	Singapore	M	Peranakan	Hokkien Chinese	medium	1999	126
002130	LAU, Charlie Kieng Hiong (Dr) (刘景熺)	Medical Services in Singapore	LKH	1920	Singapore	M	Chinese	Chinese	high	1999	165
002165	Lim Kheng Ann Dr.	Medical Services in Singapore	LKA	1924	Singapore	M	Chinese	Chinese	high	1999	126
002179	S Rajaratnam	Medical Services in Singapore	SR	1921	Singapore	M	Indian	English	high	1999	107
002186	Lee Liang Hye	Communities of Singapore (Part 1)	LLH	1924	Singapore	M	Peranakan	Hokkien Chinese	high	1986	99

A historical sociolinguistic reconstruction of CSE

002198	Tan I Tong (陈以桐)	Economic Development of Singapore	TIT	1921	Indonesia	M	Chinese	Hokkien Chinese	medium	1999	109
002204	TAN, Leo Wee Hin (Prof)	Special project	TWH	1944	Singapore	M	Chinese	Chinese	high	2004	98
002206	HEE, Mary Sook Yin	Medical Services in Singapore	MH	1937	Singapore	F	Chinese	Hakka, Cantonese	low	1999	184
002275	IAU, Robert Kuo Kwong (姚国光)	The Public Service	IKK	1935	Singapore	M	Chinese	Chinese	high	2001	261
002325	CHEN, Kenneth Koon Lap (陈冠立)	Political History of Singapore 1985–2005	KC	1940	Burma	M	Chinese	Chinese	high	2001	135
002597	Hwang Peng Yuan	The Public Service	HPY	1936	Singapore	M	Chinese	Chinese	high	2002	272
002598	Neo, Jack	Singapore Film Industry	JN	1960	Singapore	M	Chinese	Hokkien Chinese	medium	2002	23
002715	CHAN Chin Bock (曾振木)	The Public Service	CCB	1933	Singapore	M	Chinese	Chinese	high	2002	103
002749	TAN, Dixie Dr.	Medical Services in Singapore	DT	1935	Malaysia	F	Chinese	Mandarin, Cantonese, Hokkien, Malay	high	2003	288
002818	Maarof Salleh	Communities of Singapore (Part 3)	MS	1948	Singapore	M	Chinese	Malay	high	2004	204
002827	CHOY, Elizabeth (teacher)	Special project	EC	1910	Malaysia	F	Chinese	Chinese (Hokkien and Cantonese), Malay	medium	2004	438
002847	KOH, James Cher Siang (许慈祥)	The Public Service	JK	1946	Singapore	M	Chinese	Chinese	high	2008	146
002951	Ms Jean Yip & Ms Dawn Yip	Special project	JY	1958	Singapore	F	Chinese	Cantonese Chinese	medium	2005	92
002986	Wong Kee Hung	Singapore Film Industry	WKH	1911	Malaya	M	Chinese	Cantonese	high	2005	143

002995	Lee Hock Seng	Japanese Occupation of Singapore	LHS	1922	Malacca	M	Chinese	Hokkien Chinese, Malay	medium	2005	40
003155	Brian Richmond	Special project	BR	1947	Singapore	M	Eurasian	English, Malay	medium	2007	70
003182	Arthur Daniel Matthew	The Public Service	ADM	1922	Singapore	M	Indian	Tamil, English	medium	2007	154
003196	Lim-Yeo Siok Peng (杨淑萍)	The Public Service	LSP	1954	Singapore	F	Chinese	Chinese	high	2007	50
003206	Mahmud Awang	Political History of Singapore 1945–1965	MA	1928	Malaya	M	Malay	Malay	medium	2007	77
003223	Chan Soo Sen	Political History of Singapore 1985–2005	CSS	1956	Singapore	M	Chinese	Hokkien, Cantonese	high	2005	559
003409	LEE, Samuel Jian Wie	Medical Services in Singapore	LSC	1983	Singapore	M	Chinese	Mandarin	high	2009	15
E000005	Arun Mahiznan	Interviews for "Ong Teng Cheong Planner Politician President"	A	1946	India	M	Indian	Tamil	high	2001	9
E000284	Chen Nee Sian	WESLEY ARCHIVES & HERITAGE COMMITTEE	CNS	1921	Singapore	M	Chinese	Chinese	high	2001	27
E000285	Chen Swee Soo	Wesley Methodist Church	CSS	1930	Singapore	F	Chinese	Cantonese Chinese	medium	2002	25

II. *Already* summary

a) Individual speakers in OHI (semantic coding of *already*)

ACC ID	Initials	A	B	C	D	E	F	G			H	I	J					P	Q
								C	D	E			K	L	M	N	O		
000001	TMK	45	0	22	23	0	7	8	6	0	0	0	4	1	0	0	1	0	0
000002	PL	23	0	8	15	0	3	3	1	0	0	0	7	0	0	0	0	0	0
000004	JAS	5	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000009	LGS	31	0	14	16	0	5	3	0	0	0	0	2	0	0	0	0	0	1
000013	SEA	15	1	12	2	0	1	1	1	0	0	0	1	0	1	0	0	0	0
000014	SVL	18	1	9	8	0	1	1	1	0	0	0	4	1	0	0	0	0	0
000016	LKL	12	0	6	6	0	0	2	2	0	0	0	3	0	0	1	0	0	0
000021	CCS	4	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000024	FSC	33	0	22	10	1	5	3	3	1	0	0	0	1	0	0	0	0	0
000025	KRM	15	0	5	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000057	LYC	102	3	44	54	1	27	17	4	1	0	3	14	1	3	0	0	0	1
000071	CKM	47	0	26	21	0	1	6	4	0	0	2	3	0	0	1	0	0	0
000081	K	46	1	17	28	0	22	9	3	0	0	0	15	0	8	0	0	0	0
000095	MB	7	0	3	4	0	0	1	0	0	0	0	0	0	0	0	0	0	0
000107	NTH	5	0	3	2	0	2	1	0	0	0	0	1	0	0	0	0	0	0
000114	JDHN	6	0	2	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0
000123	NB	11	0	4	7	0	5	2	0	0	2	0	0	0	0	1	0	0	0
000133	OW	29	0	17	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000166	DB	41	1	27	13	0	0	1	0	0	0	0	0	0	0	0	0	0	0
000205	SW	10	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000211	LKT	26	0	10	16	0	5	1	1	0	0	0	4	0	1	0	0	0	0
000213	EQ	14	0	8	6	0	3	6	1	0	1	1	6	0	5	0	0	0	1
000237	HCY	35	0	18	17	0	19	3	1	0	0	1	2	0	2	0	0	1	0
000242	II	67	1	35	30	1	24	34	0	1	0	0	1	2	1	0	1	2	0
000259	AJ	4	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000263	S	16	0	3	13	0	0	1	0	0	0	0	1	0	0	0	0	0	0
000265	LTS	51	2	25	24	0	16	15	4	0	0	2	5	1	0	0	0	1	2
000284	ARKS	20	0	5	15	0	10	1	2	0	0	1	2	0	0	1	0	1	0
000296	AL	15	0	4	11	0	9	2	1	0	1	0	2	0	1	1	0	0	0
000310	SGB	37	0	11	24	2	18	9	1	2	0	2	5	1	1	0	0	1	0
000316	YSE	17	0	7	10	0	3	1	3	0	0	0	1	0	1	0	0	1	1
000345	AST	82	2	37	43	0	51	10	0	0	1	0	12	0	0	0	0	0	1
000350	BPA	24	0	4	20	0	5	0	1	0	1	0	3	0	0	0	0	1	1
000362	JNKB	4	0	1	3	0	1	0	0	0	0	0	1	0	0	0	0	0	0
000371	TBN	13	0	4	9	0	1	0	0	0	0	0	0	0	0	0	0	0	0

000374	GWG	98	0	23	75	0	12	2	5	0	0	0	6	0	1	0	0	1	0
000404	RW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
000416	TCH	20	0	9	11	0	9	1	0	0	0	1	0	0	0	0	0	0	0
000454	STS	11	0	3	8	0	4	2	0	0	0	2	0	1	0	0	0	0	0
000462	CSC	87	0	24	61	2	27	12	10	2	3	4	2	0	0	0	0	0	0
000483	TTC	3	0	1	2	0	3	0	0	0	0	0	3	0	0	0	0	0	0
000485	LLH	6	1	3	2	0	2	1	0	0	0	0	1	0	0	0	0	0	0
000526	LKS	41	2	16	23	0	8	6	4	0	2	0	5	1	1	2	0	1	0
000838	SMK	31	0	17	14	0	4	4	5	0	1	2	3	2	2	0	0	0	0
000852	NSM	79	0	27	50	2	34	18	4	0	2	9	18	1	12	0	0	1	0
000920	AH	10	0	0	9	1	0	0	0	1	0	0	0	0	0	0	0	0	0
001109	MN	23	0	8	15	0	10	5	3	0	0	0	8	0	3	2	1	0	0
001255	MSS	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
001309	MG	8	0	1	7	0	0	1	0	0	0	0	0	0	0	0	0	0	0
001423	VP	13	0	0	12	1	0	0	0	1	0	0	0	0	0	0	0	0	0
001478	SRSD	11	0	1	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
001536	LR	93	8	45	39	1	15	12	2	1	4	1	9	0	1	0	0	0	0
001599	LY	108	0	41	66	1	29	13	4	1	5	2	2	2	1	0	0	0	0
001600	NT	111	1	44	66	0	15	8	5	0	1	1	2	0	1	0	0	0	2
001613	LC	11	0	4	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
001631	OCN	203	7	102	84	10	69	66	10	10	6	17	52	7	30	3	1	1	0
001632	CHN	47	0	33	12	2	13	27	0	2	1	10	11	0	7	0	0	0	0
001663	RZ	139	5	92	41	1	53	34	3	1	2	1	43	0	4	0	0	0	0
001880	LKN	30	1	14	15	0	11	10	2	0	0	0	7	1	0	0	0	0	0
001915	SML	40	1	16	23	0	21	9	0	1	0	2	21	0	3	0	0	0	0
001917	LML	43	1	24	18	0	13	10	9	0	0	0	3	0	2	0	0	0	0
001953	LAS	150	2	53	88	7	67	43	9	7	8	18	90	0	23	0	0	0	0
001970	NJK	11	0	3	7	1	5	1	1	1	0	1	4	0	0	0	0	0	0
002017	FKS	44	0	20	23	1	22	8	0	1	1	2	19	2	1	0	0	0	0
002027	WYC	24	1	10	13	0	3	2	0	0	0	2	1	0	1	0	0	0	1
002038	TNC	21	0	9	12	0	3	0	0	0	0	0	1	0	0	0	0	0	2
002039	SRC	3	0	1	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0
002044	RF	19	1	13	5	0	1	1	0	0	0	0	3	0	1	0	0	0	1
002057	AC	9	0	5	4	0	6	0	0	0	0	0	4	0	0	0	0	0	0
002068	LKC	42	2	20	18	2	23	4	2	1	0	1	20	0	3	1	0	0	2
002107	CPK	24	1	12	10	1	10	6	1	1	0	3	10	1	2	0	0	0	0
002108	TSG	29	0	16	12	1	10	6	0	1	2	0	14	0	4	0	0	0	0
002130	LKH	42	1	24	14	3	19	6	1	3	1	3	18	1	2	0	0	0	0
002165	LKA	28	0	17	11	0	6	2	2	0	0	0	6	0	0	0	0	0	0
002179	SR	32	0	12	19	1	17	1	1	0	4	0	7	1	0	0	0	0	0
002186	LLH	26	0	8	18	0	8	5	0	0	0	0	1	0	0	0	0	0	0
002198	TIT	72	3	45	23	1	35	12	4	0	1	3	31	1	7	1	0	0	0

A historical sociolinguistic reconstruction of CSE

002204	TWH	12	0	4	7	1	5	2	0	1	1	0	4	0	0	1	0	0	0
002206	MH	86	1	37	44	4	36	19	6	4	9	8	44	0	8	0	0	0	1
002275	IKK	86	0	34	50	1	18	6	1	0	0	0	16	0	1	0	0	0	3
002325	KC	51	1	19	29	1	23	5	2	1	1	0	24	2	4	0	0	0	0
002597	HPY	57	0	18	39	0	15	7	4	0	0	0	9	0	2	0	0	0	0
002598	JN	11	0	1	9	1	8	1	0	1	0	0	7	0	3	0	0	0	0
002715	CCB	31	2	17	12	0	7	6	1	0	0	2	4	1	1	0	0	0	0
002749	DT	41	1	16	23	1	5	1	0	1	0	1	4	0	0	0	0	0	0
002818	MS	40	0	18	22	0	7	2	0	0	0	0	5	0	1	0	0	0	0
002827	EC	62	7	26	27	2	10	2	0	2	2	0	5	0	1	0	0	0	0
002847	JK	44	1	21	21	0	8	4	0	0	1	0	8	0	2	0	0	0	0
002951	JY	57	2	25	20	10	19	18	1	7	5	4	31	0	3	0	0	0	0
002986	WKH	53	1	23	27	2	24	14	1	3	3	6	25	1	7	0	0	0	0
002995	LHS	8	0	2	6	0	8	0	0	0	0	0	7	0	0	0	0	0	0
003155	BR	18	0	8	10	0	0	8	0	0	0	0	1	0	0	0	0	0	0
003182	ADM	33	1	12	20	0	15	9	1	0	0	0	8	0	2	0	0	0	0
003196	LSP	9	0	2	7	0	1	0	0	0	0	0	1	0	0	0	0	0	0
003206	MA	11	0	9	2	0	5	4	0	0	0	1	6	0	1	0	0	0	0
003223	CSS	151	2	65	79	5	56	18	12	4	4	8	62	1	16	0	0	0	2
003409	LSC	9	0	4	5	0	1	1	0	0	0	0	1	0	0	0	0	0	0
E000005	MA	8	0	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E000284	CNS	16	0	8	5	3	5	4	0	2	1	2	5	0	2	0	0	0	0
E000285	CSS	9	1	4	2	2	2	0	0	1	1	0	2	0	0	0	0	0	0
SUM		3676	70	1613	1912	77	1110	610	157	67	79	129	803	34	190	15	4	12	22

b) Different text types in ICE-SG (semantic coding of *already*)

File	A	B	C	D	E	F	G			H	I	J					P	Q
							C	D	E			K	L	M	N	O		
S1A-001-100	292	8	105	164	11	130	52	20	7	15	22	102	1	31	4	0	0	0
S1B-001-080	106	5	35	61	1	23	4	2	1	2	1	16	0	1	0	0	0	7
S2A-001-070	50	1	19	28	2	16	0	5	2	3	1	17	2	2	1	0	0	0
S2B-001-050	41	2	11	26	1	9	2	2	1	0	0	6	0	2	0	0	0	1
SUM	489	16	170	279	15	178	58	29	11	20	24	141	3	36	5	0	0	8

A total

B unclear

C with stative predicate

D with non-stative predicate

E with modal verb

F completive

G inchoative

H prospective

I in negative sentence

J other

K lack of inflectional marker/perfect markers

L double

M predicate ellipsis

N word order change

O parallel with still

P discourse functions

Q pre-adjective

III. *Also summary*

a) Individual speakers in OHI (semantic coding of *also*)*

ACC ID	Initials	NEG	even	universal quantifiers				double
				all	every*	always	any*	
000001	TMK	1	0	0	1	0	0	0
000002	PL	2	0	0	0	0	0	0
000013	SEA	1	0	1	1	0	0	1
000014	SVL	2	0	2	0	0	0	0
000016	LKL	1	1	0	0	0	0	0
000024	FSC	1	0	0	0	0	0	0
000025	KRM	1	0	1	1	1	0	0
000057	LYC	0	1	2	0	0	0	0
000071	CKM	1	0	0	0	0	0	0
000081	K	12	5	10	4	0	0	6
000095	MB	0	0	1	0	1	0	0
000107	NTH	0	0	1	0	0	0	0
000114	JDHN	1	0	0	0	0	0	0
000123	NB	4	0	0	1	2	1	0
000166	DB	0	0	1	1	0	0	0
000205	SW	1	0	0	0	0	0	0
000211	LKT	1	0	0	0	0	0	1
000213	EQ	0	0	1	0	0	0	0
000237	HCY	3	0	3	0	0	0	0
000242	II	9	13	12	6	2	1	1
000263	S	1	5	0	3	0	1	0
000265	LTS	4	0	1	1	0	0	0
000284	ARKS	2	2	0	3	0	0	1
000296	AL	3	0	1	0	0	2	0
000310	SGB	4	1	15	2	1	0	0
000345	AST	5	1	1	0	1	0	0
000350	BPA	3	1	1	0	0	0	1
000374	GWG	1	0	0	0	0	1	0
000416	TCH	0	0	2	0	0	0	0
000454	STS	1	0	2	1	0	0	0
000462	CSC	4	0	1	1	0	0	0
000483	TTC	0	0	1	3	0	0	0
000838	SMK	2	0	1	1	0	0	0
000852	NSM	3	0	0	2	0	0	0
001109	MN	0	1	3	3	2	1	0

A historical sociolinguistic reconstruction of CSE

001255	MSS	0	0	1	0	0	0	0
001309	MG	1	0	1	0	0	0	0
001423	VP	1	0	0	0	0	0	0
001536	LR	5	1	2	4	0	1	0
001599	LY	5	1	0	1	0	0	0
001600	NT	2	1	0	1	3	0	0
001631	OCN	13	4	12	1	5	2	2
001632	CHN	1	0	0	4	0	0	0
001663	RZ	22	9	19	6	8	3	2
001880	LKN	1	0	0	1	0	0	0
001917	LML	2	0	0	4	1	1	0
001953	LAS	4	3	6	6	4	1	1
002017	FKS	1	0	0	0	0	0	0
002027	WYC	1	0	1	0	0	0	0
002038	TNC	0	1	0	0	1	0	0
002039	SRC	0	0	0	1	0	0	0
002044	RF	0	0	0	0	0	1	0
002057	AC	1	2	1	0	0	0	3
002068	LKC	2	1	4	1	0	1	0
002107	CPK	0	0	0	0	0	0	1
002108	TSG	0	0	2	0	1	0	0
002130	LKH	1	2	1	0	0	0	1
002165	LKA	1	0	0	0	0	0	0
002179	SR	0	0	1	0	0	0	0
002198	TIT	1	0	4	1	0	0	1
002204	TWH	1	0	1	0	1	0	0
002206	MH	2	2	4	0	0	0	0
002325	KC	1	0	0	0	1	0	0
002597	HPY	0	0	0	1	0	0	0
002598	JN	0	0	0	1	0	0	0
002715	CCB	1	0	0	0	0	0	0
002749	DT	1	0	0	1	0	0	0
002818	MS	1	0	0	0	1	0	0
002827	EC	6	1	9	6	2	0	2
002847	JK	1	0	2	1	0	0	0
002951	JY	0	0	1	4	0	0	0
002986	WKH	2	1	2	0	0	1	1
003206	MA	1	4	1	0	0	0	0
003223	CSS	2	0	2	0	0	0	0
003409	LSC	0	1	0	0	0	0	0
SUM		158	65	141	80	38	18	25

*Only speakers who produce at least one substrate-influenced token of *also* are listed here.

b) Different text types in ICE-SG (semantic coding of *also*)

File	NEG	even	universal quantifiers				double
			all	every*	always	any*	
S1A-001-100	11	2	1	1	0	3	2
S1B-001-080	0	0	0	0	0	0	1
S2A-001-070	1	0	1	0	0	0	0
S2B-001-050	0	0	0	0	0	0	0
SUM	12	2	2	1	0	3	3

IV. *Ever* summary

a) Individual speakers in OHI (semantic coding of *ever*)*

ACC ID	Initials	Standard English			CSE		
		existential	universal	emphatic	aspectual	affirmative response	Other
000001	TMK	1	0	0	0	0	0
000002	PL	2	0	0	0	0	0
000004	JAS	29	0	0	0	0	0
000009	LGS	6	0	0	0	0	0
000013	SEA	8	1	1	0	0	0
000014	SVL	1	0	0	0	0	0
000016	LKL	19	1	0	0	0	0
000021	CCS	2	1	0	0	0	0
000024	FSC	3	1	0	0	0	0
000025	KRM	1	0	0	0	0	0
000057	LYC	3	0	0	0	0	0
000071	CKM	6	4	1	0	0	0
000081	K	14	1	0	1	0	0
000095	MB	7	0	0	0	0	0
000107	NTH	1	0	0	0	0	0
000114	JDHN	16	2	0	0	0	0
000123	NB	1	0	0	0	0	0
000133	OW	15	0	0	1	0	0
000166	DB	24	2	1	1	0	0
000205	SW	5	1	0	0	0	0
000211	LKT	5	0	0	0	0	0
000237	HCY	23	1	0	0	0	0
000242	II	4	1	0	0	0	0
000259	AJ	1	0	0	0	0	0
000263	S	2	0	0	0	0	0
000265	LTS	5	0	0	0	0	0
000284	ARKS	8	0	0	0	0	0
000310	SGB	11	0	0	0	0	0
000316	YSE	4	1	0	0	0	0
000345	AST	52	11	6	0	0	0
000350	BPA	12	0	0	0	0	0
000362	JNKB	6	1	0	0	0	0
000371	TBN	14	0	0	0	0	0
000374	GWG	16	0	0	0	0	0
000404	RW	2	0	0	0	0	0

000416	TCH	7	0	0	0	0	0
000454	STS	7	0	0	0	0	0
000462	CSC	3	0	0	0	0	0
000483	TTC	2	0	0	0	0	0
000526	LKS	14	0	0	1	0	0
000838	SMK	1	0	0	0	0	0
000852	NSM	0	3	0	0	0	0
000920	AH	0	2	0	0	0	0
001109	MN	2	0	0	0	0	0
001255	MSS	1	0	0	0	0	0
001309	MG	23	1	0	0	0	0
001423	VP	12	1	0	0	0	0
001478	SRSD	1	0	0	0	0	0
001536	LR	9	0	0	0	0	0
001599	LY	5	1	0	0	0	0
001600	NT	23	7	0	0	0	0
001613	LC	4	0	0	0	0	0
001631	OCN	6	0	0	1	0	0
001632	CHN	3	0	0	0	0	0
001663	RZ	12	0	0	1	0	0
001880	LKN	1	0	0	0	0	0
001915	SML	2	1	1	2	0	0
001917	LML	21	1	0	0	0	0
001953	LAS	2	1	0	1	0	0
001970	NJK	2	0	0	0	0	0
002017	FKS	2	0	0	0	0	0
002027	WYC	2	0	0	0	0	0
002038	TNC	3	2	0	0	0	0
002039	SRC	0	3	0	0	0	0
002044	RF	10	1	1	0	0	0
002057	AC	4	0	0	0	0	0
002107	CPK	3	0	0	0	0	0
002108	TSG	7	3	0	2	0	0
002130	LKH	1	0	0	0	0	0
002165	LKA	1	0	0	0	0	0
002179	SR	1	0	0	0	0	0
002186	LLH	3	4	1	0	0	0
002198	TIT	3	0	0	0	0	0
002204	TWH	6	0	0	0	0	0
002206	MH	1	1	0	3	0	0
002275	IKK	14	3	0	1	0	0
002325	KC	5	0	0	0	0	0

A historical sociolinguistic reconstruction of CSE

002597	HPY	11	0	0	0	0	0
002715	CCB	1	0	0	0	0	0
002749	DT	10	0	0	0	0	0
002818	MS	1	1	0	0	0	0
002827	EC	12	2	0	0	0	0
002847	JK	4	0	0	0	0	0
002951	JY	14	0	0	0	0	0
003155	BR	6	0	0	0	0	0
003182	ADM	19	1	0	0	0	0
003196	LSP	3	0	0	0	0	0
003206	MA	1	0	0	0	0	0
003223	CSS	16	4	0	1	0	0
E000284	CNS	1	0	0	0	0	0
E000285	CSS	0	1	0	0	0	0
SUM		657	73	12	16	0	0

*Only speakers who produce at least one token of *ever* are listed in the table.

b) Different text types in ICE-SG (semantic coding of *ever*)

File	Standard English			CSE		
	esistential	universal	emphatic	aspectual	affirmative response	other
S1A-001-100	24	5	1	4	0	0
S1B-001-080	5	5	0	0	0	0
S2A-001-070	16	8	1	0	0	0
S2B-001-050	8	7	0	1	0	0
SUM	53	25	2	5	0	0

V. *One* summary

a) Individual speakers in OHI (semantic coding of *one*)

ACC ID	Initials	CSE				Standard English	
		N-one	Possessive-one	Relativizer	Emphatic	Numeral	Pronominal
000001	TMK	0	0	0	0	0	4
000002	PL	1	0	1	2	2	10
000004	JAS	0	0	0	0	10	7
000009	LGS	1	0	0	1	6	6
000013	SEA	0	0	0	0	8	5
000014	SVL	0	0	1	0	9	1
000016	LKL	1	0	0	0	9	7
000021	CCS	0	0	0	1	4	5
000024	FSC	0	0	0	0	15	12
000025	KRM	0	0	0	0	2	2
000057	LYC	0	0	0	1	10	11
000071	CKM	0	0	0	1	6	7
000081	K	5	0	3	0	13	24
000095	MB	0	0	0	0	4	0
000107	NTH	1	0	3	1	2	3
000114	JDHN	1	0	0	0	2	8
000123	NB	0	0	0	0	0	5
000133	OW	0	0	0	0	4	9
000166	DB	0	0	1	3	5	22
000205	SW	0	0	0	0	3	5
000211	LKT	0	0	0	0	0	6
000213	EQ	0	0	0	2	3	5
000237	HCY	3	0	4	4	24	60
000242	II	4	0	3	1	24	34
000259	AJ	0	0	0	0	1	1
000263	S	0	0	0	0	0	9
000265	LTS	0	0	0	0	0	6
000284	ARKS	0	0	0	0	2	6
000296	AL	1	0	1	1	9	10
000310	SGB	5	0	1	1	18	48
000316	YSE	0	0	0	1	0	4
000345	AST	1	1	1	1	32	50
000350	BPA	1	0	0	0	11	6
000362	JNKB	0	0	1	0	1	7
000371	TBN	5	0	0	2	13	30

A historical sociolinguistic reconstruction of CSE

000374	GWG	1	0	0	0	7	16
000404	RW	0	0	1	0	1	1
000416	TCH	1	0	0	0	6	3
000454	STS	0	0	0	0	4	9
000462	CSC	3	0	0	0	8	14
000483	TTC	0	0	0	0	1	6
000485	LLH	0	0	0	0	1	0
000526	LKS	0	0	0	1	11	14
000838	SMK	2	0	0	0	5	7
000852	NSM	0	0	1	1	3	16
000920	AH	0	0	1	0	3	3
001109	MN	1	0	1	1	14	16
001255	MSS	0	0	1	0	1	4
001309	MG	0	0	0	0	1	3
001423	VP	0	0	0	0	5	1
001478	SRSD	0	0	0	0	4	3
001536	LR	1	0	0	3	25	35
001599	LY	0	0	0	0	10	16
001600	NT	1	0	0	2	32	34
001613	LC	0	0	0	0	0	1
001631	OCN	3	0	7	6	22	92
001632	CHN	0	0	0	1	4	0
001663	RZ	2	0	0	2	37	67
001880	LKN	0	0	0	0	2	13
001915	SML	0	0	1	0	2	12
001917	LML	2	0	0	2	9	16
001953	LAS	24	1	28	33	35	84
001970	NJK	0	0	0	0	1	3
002017	FKS	1	0	4	0	20	38
002027	WYC	0	0	1	1	5	9
002038	TNC	0	0	0	0	4	9
002039	SRC	0	0	0	0	3	0
002044	RF	1	0	0	0	0	11
002057	AC	0	0	0	0	15	8
002068	LKC	0	0	1	0	10	24
002107	CPK	3	0	1	0	7	10
002108	TSG	0	0	1	0	8	7
002130	LKH	0	0	0	1	8	12
002165	LKA	0	0	0	0	6	2
002179	SR	0	0	0	0	4	14
002186	LLH	1	0	2	0	4	7
002198	TIT	0	0	1	2	14	16

002204	TWH	0	0	0	0	6	2
002206	MH	5	0	2	7	9	16
002275	IKK	3	0	0	0	9	12
002325	KC	1	0	0	1	6	7
002597	HPY	2	0	3	1	30	37
002598	JN	0	0	0	0	1	3
002715	CCB	1	0	1	0	0	3
002749	DT	0	0	0	0	4	14
002818	MS	2	0	0	0	18	12
002827	EC	2	0	0	1	12	21
002847	JK	1	0	0	0	11	11
002951	JY	2	0	0	3	6	15
002986	WKH	0	0	1	1	7	3
002995	LHS	1	0	1	0	4	9
003155	BR	0	0	0	0	2	2
003182	ADM	2	0	4	0	10	9
003196	LSP	0	0	0	0	3	3
003206	MA	3	0	1	1	5	23
003223	CSS	10	0	2	0	53	48
003409	LSC	0	0	0	0	0	1
E000005	MA	0	0	0	0	2	0
E000284	CNS	0	0	0	0	1	2
E000285	CSS	0	0	0	0	3	4
SUM		112	2	87	94	816	1348

b) Different text types in ICE-SG (semantic coding of *one*)

File	CSE				Standard English	
	N-one	Possessive-one	Relativizer	Emphatic	Numeral	Pronominal
S1A-001-100	21	4	19	52	69	140
S1B-001-080	1	0	0	2	42	57
S2A-001-070	2	1	2	1	41	43
S2B-001-050	0	0	3	1	14	7
SUM	24	5	24	56	166	247

Abstract

This study seeks to offer a historical sociolinguistic reconstruction of Colloquial Singapore English (CSE), a widespread contact variety of English spoken in Singapore, which shows traces of all other ethnic languages spoken in the region such as Chinese, Malay, and Tamil. So far, scholarly work has been dedicated to the phonological and grammatical features of CSE (Deterding 2007; Lim 2004; Wee 2004b), as well as the social conditions that determine their occurrences (Leimgruber, Siemund und Terassa 2018; Lim, Pakir und Wee 2010). Various models have been postulated to capture the emergence of CSE and its relationship to Standard Singapore English (Platt 1975; Gupta 1989; Alsagoff 2007, 2010; Leimgruber 2009, 2013) as well as its Chinese substrates (Bao 2005, 2015; Bao and Hong 2006). However, there is relatively little research that probes into its history, mainly due to a scarcity of historical data.

The main purpose of this study is to work towards a diachronic reconstruction of CSE by exploring a novel historical data source, namely the *Oral History Interviews* held by the *National Archives of Singapore* (OHI-NAS 2020). The database allows a significant step back in time, as the majority of the speakers sampled were born between the 1890s and 1950s. The study first provides a sociohistorical account of Singapore with personal recollections chosen from OHI, which elucidates the language-related policy and the ethnic dynamics in Singapore. Since it is impossible to provide a comprehensive picture of the development of CSE, the present study focuses on four salient grammatical markers of CSE, namely the aspectual marker *already*, the additive marker *also*, the experiential marker *ever*, and the emphatic marker *one*. These CSE expressions differ significantly from native Englishes in terms of their semantic functions and syntactic positions, but they mirror – to a large extent – the usages of the Chinese substrates. The study asks to what extent the usages of *already*, *also*, *ever*, and *one* can be related to their Chinese substrates and which social factors can explain the variation among individual speakers in using CSE markers. Furthermore, this study explores whether CSE speakers replicate a grammaticalization model into CSE that

exists in the Chinese substrates based on the use of *already*, *also*, *ever*, and *one* (Heine and Kuteva 2003, 2005 on replica grammaticalization).

The study is based on 100 interviews from OHI conducted between 1979 and 2009 amongst speakers born between 1899 and 1983. These speakers were divided into seven groups depending on their ethnic background, i.e. Chinese, Indian, Peranakan, Malay, British, Eurasian, and other. Furthermore, speakers were further differentiated based on gender and level of education received (low = primary education; medium = secondary education; high = university education). Multiple linear regression analyses were applied to examine the four CSE expressions – *already*, *also*, *ever*, and *one* – in relation to the social background of the speaker, including gender, year of birth, level of education, and ethnic background. For the establishment of a diachronic reconstruction of these four CSE expressions, the results obtained from OHI were further compared with the Singaporean component of the International Corpus of English (ICE-SG), a well-explored corpus of Singapore English, which represents a more recent sample of CSE among university students starting from 1997 (Nelson 2002:3). Furthermore, the study compares the functional extension and/or grammaticalization of the four CSE expressions with that of their equivalents in Chinese by using corpus data drawn from the *Center for Chinese Linguistics* (CCL) at Peking University as well as secondary data from the scholarly circle on the grammaticalization of 了 *le*, 过 *guò*, 也/都 *yě/dōu*, and 的 *de*.

This study builds on the Dynamic Model of Schneider (2007), which proposes that Singapore English in general has reached stage 4 (endonormative stabilization), in which linguistic innovations become increasingly accepted. It is possible that CSE is moving towards stage 5, in which “internal differentiation” emerges (Schneider 2007:54). Internal differentiation refers to “differences within a society and between individuals with respect to their economic status, social categories and personal predilections” (Schneider 2007:53). The findings of this study suggest that CSE remains relatively stable, apparently undergoing no substantial changes, at least judging from the use of *already*, *also*, *ever* and *one*. Furthermore, among the social variables under investigation, the most important factors in

determining the frequencies of the substrate-influenced tokens are ethnicity and educational level.

To a large extent, the Sinitic languages have provided the extended semantic functions and syntactic features for the CSE expressions. In some cases, the influence from Chinese in the CSE expressions is so strong that the entire CSE sentence no longer meets the surface structural requirements of English (see Bao 2005, 2015 on systemic transfer and lexifier filters). With regards to replica grammaticalization (Heine and Kuteva 2003, 2005), the findings suggest that the pathways of grammaticalization of the aspectual markers *already* and *ever* do not match those of their Chinese counterparts. However, emphatic *one* and additive *also* seem to have replicated the grammaticalization models into CSE that already existed in Chinese.

The sociolinguistic variation found in this study based on the historical data of CSE reveals that social factors such as ethnic background and level of education as well as individual speaker preference may have been present from very early on. It suggests that further studies on CSE and other new varieties of English – both synchronic and diachronic – should continue to investigate these social variables of the individual speakers in the diverse local contexts.

Deutsche Kurzfassung der Ergebnisse

Diese Studie stellt eine historische soziolinguistische Rekonstruktion des „Colloquial Singapore English“ (CSE) dar. CSE ist eine Kontaktvarietät des in Singapur gesprochenen Englisch, die Spuren aller anderen in der Region gesprochenen ethnischen Sprachen wie Chinesisch, Malaiisch und Tamilisch aufweist. Bisher widmete sich die Forschung überwiegend den phonologischen und grammatikalischen Charakteristiken von CSE (Deterding 2007; Lim 2004; Wee 2004b) sowie den sozioökonomischen Rahmenbedingungen, die ihr Auftreten determinieren (Leimgruber, Siemund und Terassa 2018; Lim, Pakir und Wee 2010). Es wurden verschiedene Modelle postuliert, um die Entstehung von CSE und seine Beziehung zu „Standard Singapore English“ (Platt 1975; Gupta 1989; Alsagoff 2007, 2010; Leimgruber 2009, 2013) sowie seinen chinesischen Substraten zu erfassen (Bao 2005, 2015; Bao und Hong 2006). Es wurden bis heute jedoch wenige Studien publiziert, die sich mit der historischen Entwicklung befassen, was hauptsächlich auf den Mangel an historischen Daten zurückzuführen ist.

Das Ziel dieser Studie besteht darin, eine diachrone Rekonstruktion des CSE zu etablieren, hinzuarbeiten, die auf der Erforschung der historische Datenquelle „Oral History Interviews des National Archives of Singapore“ (OHI-NAS 2020) basiert. Diese Datenbank ermöglicht die Erforschung des CSE in einem historischen Kontext, da die Mehrheit der befragten Sprecher und Sprecherinnen zwischen den 1890er und 1950er Jahren geboren wurde. Die Studie liefert zunächst einen soziohistorischen Bericht über Singapur mit persönlichen Erinnerungen, die aus den OHIs ausgewählt wurden, um die sprachbezogene Politik und die ethnische Dynamik in Singapur zu erläutern. Da es unmöglich ist, ein umfassendes Bild der Entwicklung des CSE zu rekonstruieren, fokussiert sich die vorliegende Studie auf vier herausragende grammatikalische Marker von CSE, nämlich den Aspektmarker *already* ‚bereits‘ oder ‚schon‘, den additiven Marker *also* ‚auch‘, ‚ebenfalls‘ oder ‚außerdem‘, den experientiellen Aspektmarker *ever* ‚jemals‘, und den emphatischen Marker *one* ‚eins‘. Diese CSE-Ausdrücke unterscheiden sich in ihren semantischen Funktionen und syntaktischen Positionen erheblich von den

muttersprachlichen Varietäten, spiegeln jedoch weitgehend die Verwendung der chinesischen Substrate wider. Die Studie stellt ferner die Frage, inwieweit die Verwendung von *already*, *also*, *ever* und *one* mit ihren chinesischen Substraten zusammenhängt und welche sozialen Faktoren die Unterschiede zwischen einzelnen Sprechern bei der Verwendung von CSE-Markern erklären können.

Die Studie basiert auf 100 Interviews des OHI-NAS, die zwischen 1979 und 2009 mit Sprecherinnen und Sprechern durchgeführt wurden, die zwischen 1899 und 1983 geboren wurden. Diese Sprecher und Sprecherinnen wurden entsprechend ihrem ethnischen Hintergrund in sieben Gruppen eingeteilt, nämlich Chinesisch, Indisch, Peranakanisch, Malaiisch, Britisch, Eurasisch, und andere. Darüber hinaus wurde zwischen weiblichen Sprecherinnen und männlichen Sprechern sowie zwischen niedrigem (Grundschulbildung), mittlerem (Sekundarschulbildung) und hohem Bildungsniveau (Universitätsausbildung) unterschieden. Multilineare Regressionsmodelle wurden angewendet, um die oben genannten vier CSE-Ausdrücke in Bezug auf den sozialen Hintergrund der Sprecherinnen und Sprecher zu untersuchen, einschließlich Geschlecht, Geburtsjahr, Bildungsniveau und ethnischen Hintergrund. Für die Etablierung einer diachronen Rekonstruktion der oben genannten CSE-Ausdrücke wurden die Ergebnisse basierend auf den OHI mit der singapurischen Komponente des International Corpus of English (ICE-SG) verglichen. ICE-SG stellt ein gut erforschtes Korpus singapurischen Englischs dar, welches gesprochene Sprache von Universitätsstudierenden ab 1997 enthält und somit neuere Daten von CSE liefert (Nelson 2002:3). Darüber hinaus vergleicht die Studie die funktionale Erweiterung und / oder Grammatikalisierung der vier CSE-Ausdrücke mit denen ihrer Äquivalente im Chinesischen unter Verwendung von Korpusdaten aus dem *Center for Chinese Linguistics* (CCL) (‘Zentrum für chinesische Linguistik’) der Universität Peking sowie Sekundärdaten aus dem wissenschaftlichen Kreis zur Grammatikalisierung von 了 *le*, 过 *guò*, 也/都 *yě/dōu* und 的 *de*. Die Studie untersucht weiter, ob CSE-Sprecher und Sprecherinnen ein Grammatikalisierungsmodell in CSE replizieren, das auf dem chinesischen Substrat aufbaut, basierend auf der Verwendung von *already*, *also*, *ever* und *one* (Heine und Kuteva 2003, 2005 zur Replika-Grammatikalisierung).

Diese Studie baut auf dem „Dynamic Model“ von Schneider (2007) auf, der postuliert, dass „Singapore English“ generell Phase 4 (endonormative Stabilisierung) erreicht hat, in dem sprachliche Innovationen zunehmend akzeptiert werden. Möglicherweise bewegt sich CSE in Richtung Phase 5, in der eine „interne Differenzierung“ auftritt (Schneider 2007: 54, eigene Übersetzung). Interne Differenzierung bezieht sich dort auf „Unterschiede innerhalb einer Gesellschaft und zwischen Individuen hinsichtlich ihres wirtschaftlichen Status, ihrer sozialen Kategorien und ihrer persönlichen Vorlieben“ (Schneider 2007: 53, eigene Übersetzung). Die Ergebnisse dieser Studie legen nahe, dass CSE relativ stabil ist und anscheinend keine wesentlichen Veränderungen erfährt, zumindest basierend auf der Verwendung von *already*, *also*, *ever* und *one*. Abgesehen davon scheint es schon sehr früh zu soziolinguistischen Variationen gekommen sein. Unter den untersuchten sozialen Variablen sind die wichtigsten Faktoren, die die bei Häufigkeit der vom Substrat beeinflussten Tokens determinieren, die ethnische Zugehörigkeit und das Bildungsniveau.

Die sinitischen Sprachen haben weitgehend die erweiterten semantischen Funktionen und syntaktischen Merkmale für die CSE-Ausdrücke bereitgestellt. In einigen Fällen ist der Einfluss des Chinesischen auf die CSE-Ausdrücke so stark, dass der ganze CSE-Satz nicht mehr den Anforderungen der englischen Oberflächenstruktur entspricht (vgl. „systemic transfer and lexifier filter“ Bao 2005, 2015). In Bezug auf die Replika-Grammatikalisierung (Heine und Kuteva 2003, 2005) legen die Ergebnisse nahe, dass die Wege der Grammatikalisierung der Aspektmarker *already* und *ever* nicht mit denen ihres jeweiligen chinesischen Äquivalents übereinstimmen. Bei emphatischem *one* und additivem *also* scheinen jedoch die Grammatikalisierungsmodelle in CSE repliziert worden zu sein, die bereits im Chinesisch existierten.

Die soziolinguistische Variation, die in dieser Studie auf der Grundlage der historischen Daten von CSE gefunden wurde, zeigt, dass soziale Faktoren wie ethnischer Hintergrund und Bildungsniveau sowie bestimmte Präferenzen einzelner Sprecher und Sprecherinnen möglicherweise schon sehr früh existierten. Diese Erkenntnisse lassen darauf schließen, dass weitere Studien zu CSE und anderen neuen Varietäten des Englischen –

sowohl synchron als auch diachron – erforderlich sind, um diese sozialen Variablen der einzelnen Sprecher und Sprecherinnen in den verschiedenen lokalen Kontexten weiter zu untersuchen.

Veröffentlichungen die aus dieser Dissertation hervorgegangen sind

- Li, Lijun, and Peter Siemund. 2021. "From Phasal Polarity Expression to Aspectual Marker: Grammaticalization of 'already' in Asian and African Varieties of English." Pp. 515–44 in *The expression of phasal polarity in African languages*, edited by K. Raija. Berlin: De Gruyter.
- Siemund, Peter, and Lijun Li. 2020. "Multilingualism and Language Policy in Singapore." Pp. 205–28 in *Language Diversity in the Sinophone World: Historical Trajectories, Language Planning, and Multilingual Practices*, edited by H. Klöter and M. S. Saarela. Berlin: Routledge.
- Siemund, Peter, and Lijun Li. 2017. "Towards a Diachronic Reconstruction of Colloquial Singapore English." Pp. 11–32 in *Negation and Contact*, edited by D. Ziegeler and Z. Bao. Amsterdam and Philadelphia: John Benjamins.

Eidesstattliche Erklärung über das selbstständige Verfassen der vorliegenden Arbeit

Hierdurch versichere ich an Eides Statt, dass ich die vorliegende Dissertation selbständig angefertigt, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den herangezogenen Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Dissertation wurde in keinem früheren Promotionsverfahren eingereicht.

Lijun Li

Hamburg, 01.03.2021