

HANNA PARVIAINEN

# Crossing the Borders

The Influence of Indian English  
in the Southeast Asian Region



HANNA PARVIAINEN

Crossing the Borders

*The Influence of Indian English  
in the Southeast Asian Region*

ACADEMIC DISSERTATION

To be presented, with the permission of  
the Faculty of Information Technology and Communication Sciences  
of Tampere University,  
for public discussion at Tampere University  
on 11 September 2020, at 12 o'clock.

# ACADEMIC DISSERTATION

Tampere University, Faculty of Information Technology and Communication  
Sciences  
Finland

<i>Responsible supervisor and Custos</i>	Professor Juhani Klemola Tampere University Finland	
<i>Supervisor</i>	Professor Päivi Pahta Tampere University Finland	
<i>Pre-examiners</i>	Professor Devyani Sharma Queen Mary University of London UK	Professor Sebastian Hoffmann University of Trier Germany
<i>Opponent</i>	Professor Devyani Sharma Queen Mary University of London UK	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Copyright ©2020 author

Cover design: Roihu Inc.

ISBN 978-952-03-1644-0 (print)  
ISBN 978-952-03-1645-7 (pdf)  
ISSN 2489-9860 (print)  
ISSN 2490-0028 (pdf)  
<http://urn.fi/URN:ISBN:978-952-03-1645-7>

PunaMusta Oy – Yliopistopaino  
Vantaa 2020

Dedicated to the women in my family,  
whose courage and optimism never cease to inspire me.



# ACKNOWLEDGEMENTS

Nobody writes their dissertation on an island (at least not all of it), and thus, there are many people I would like to thank for their help for getting me to this point.

The first and greatest thanks go to my supervisor, Professor Juhani Klemola, without whose encouragement I never would have dared to pursue a doctoral degree, and whose patience and support helped me overcome many obstacles over the years. I would also like to thank my second supervisor Professor Päivi Pahta for her support. In addition, I wish to thank Mark Kaunisto, whose MA seminar was where I first began working on this topic.

I would also like to thank the preliminary examiners of this work, Professor Devyani Sharma and Professor Sebastian Hoffmann, for their insightful and constructive comments. All remaining shortcomings are naturally my own.

During the writing of this dissertation, I have been fortunate to receive financial support from a number of sources. I am grateful to the GlobE and ChangE consortia, both headed by Professors Markku Filppula, Juhani Klemola, and Anna Mauranen, for not only funding my work but also introducing me to many wonderful colleagues from the University of Eastern Finland and the University of Helsinki. Furthermore, I would like to thank Langnet, the Finnish Graduate School in Language Studies, for their financial support and inspiring seminars where I often received invaluable feedback on my writing and – equally importantly – for providing me with a supportive community of peers.

I also want to thank Robert Fuchs for collaborating with me on one of the articles included in this dissertation – without you, this work would not be what it is today.

Great thanks also go to Professor Marianne Hundt for inviting me to the University of Zürich to work on the ICE-Fiji project and Lena Zipp for all the practical advice in matters related to work and leisure during my visit. I am grateful for the opportunity to contribute my small share to the project and for the advice and feedback I received from you and others on my dissertation during its early stages.

My heartfelt thanks go to my colleagues and friends in the Languages Unit at Tampere University. Thank you Paula Rautionaho and Paul Rickman for all the advice, peer support, and laughs. I also wish to thank the teaching staff I worked

with and the students I taught over the years – I learned so much from you all. A special thanks go to Kate Moore for all the lessons in and outside the classroom during our trips to the most unexpected corners of the world. (I should also thank you for letting me stay in your summer cottage – who knew that all that was needed to restart the thesis writing process was a cottage on an actual island!) In addition, I would like to thank all the people I have shared my office with over the years and all the colleagues in the 5<sup>th</sup> floor coffee (/tea) room in Pinni B. Without the daily laughs, finishing this dissertation might have happened slightly sooner, but it would have been infinitely less fun.

There are a number of other people outside of academia who I am grateful to have in my life and who have ensured that I actually have a life outside work. Firstly, I want to thank the Ihalainen family, especially Ilkka and Emmi, for their support over the years. Another great round of thanks goes to my lovely Ruokailta™ crew – Aino, Eeva, and Kaisa – who have for almost two decades now made sure that no matter how stressful or busy life gets, there is always time for food, wine, and merriment. My thanks also go to those I was not able to see as often as I wanted to, but who are no less dear to me – thank you Juuli, Kukka-Maaria, and Mitcho.

I wish to thank my flamenco group for ensuring that at least once a week, I had to put all other things aside and focus on something completely different. I also owe a great thanks to all the tea farmers around the world for supplying the hundreds upon hundreds of litres of (mostly) green tea that were consumed during the writing of this dissertation. Without your contribution, this work would never have been finished.

Lastly, I would like to thank my family. My heartfelt thanks go to my grandparents, especially Mummu and Taata, for their support throughout my years in academia. I am sad that Taata, who was always enthusiastic about my work, passed away before I could complete it – you are with me in spirit. The final and greatest thanks go to Iskä, Barbro, and Siri, and Minna, Ida, and Äiti.

Tampere, August 2020

Hanna Parviainen



# ABSTRACT

Asia is home to over 2,000 languages and while the majority of these are of local origin, there are also languages such as English that entered the linguistic landscape as a result of colonialism. After English was planted in numerous (often multilingual) communities around Asia, the language continued to evolve locally and developed various innovative features as a response to the communicative needs of its new users. Over the years, many of these Englishes have become nativized, that is, they have developed local feature profiles that are unique for each variety. Today, many of these New Englishes have the status of an official language in the countries where they developed, and they are often spoken as a second language by the majority of the population. One of the largest speaker populations of these New Englishes is located in India, where English is one of the official languages. Previous studies show that this local, nativized variety of English, Indian English (IndE), has already become prominent in the region to the extent that it is able to influence the development of other English varieties spoken in South Asia – in other words, it has become a linguistic epicentre. However, no research has been conducted to see whether this influence could have spread further, to the varieties spoken in Southeast Asia. This is the aim of the present study.

The Asian varieties included in this article-based dissertation are the Englishes spoken in India, Singapore (SinE), Hong Kong (HKE) and the Philippines (PhiE). The first three countries are former British colonies while the fourth is a former American colony and therefore, British (BrE) and American (AmE) varieties are also included in the study for points of reference. The choice of the three Southeast Asian countries is based on their differing connections with India. Singapore is a country that has a sizeable Indian ethnic minority, it has maintained close connections with India since colonial times, and it is geographically closest to India when compared with the other two Southeast Asian countries examined here. While Hong Kong also shares India's past as a former British colony, it lacks a significant Indian minority, strong historical connections, and geographical proximity with India. The Philippines in turn was never connected to India through the British colonial empire, the size of its Indian minority is negligible, and it is geographically furthest away from India. Because of these factors, the present study hypothesises that if there is any

sign of IndE extending its epicentric influence to Southeast Asia, it would most likely be detected in SinE and possibly in HKE, but not in PhiE.

The potential epicentric role of IndE in Southeast Asia is studied through the use of three syntactic features, all of which previous studies have shown to be local innovations in IndE and which have also been noted to exist in some of the Southeast Asian varieties included in this study. The features examined here include the use of clause-final *also* and *only*, the use of the invariant tag *isn't it* and the tendency to omit direct objects, and each feature is examined in a separate article included in this article-based dissertation. The study focuses mainly on the use of these features in spoken language where syntactic innovations often occur most frequently. While the majority of the data comes from the *International Corpus of English* (ICE), a family of corpora that has comparable data on a large number of varieties of English, the *Freiburg-Brown Corpus of American English* and *Santa Barbara Corpus of Spoken American English* were also used in some occasions when comparable data were not available in ICE.

The results of the study show that all three syntactic features are used most frequently in IndE, which for all three features is followed by SinE and, depending on the feature, by HKE or PhiE. An investigation into the major substrates of each variety shows that while for IndE the substrate effect is the most plausible explanation for the existence of the feature in the variety, for the three Southeast Asian varieties, the situation is more complex. Interestingly, for SinE and possibly HKE the influence of IndE is a factor that could help explain the results of the two varieties more comprehensively, whereas for PhiE, no indication of IndE's influence could be detected. Therefore, the results of the study lend cautious support to the argument that IndE could have extended its influence on some Southeast Asian varieties, namely SinE.

In order to explain what could have caused this, the study also presents a detailed investigation of the historical, cultural, economic, linguistic and social connections between India and the three Southeast Asian countries; special attention is paid to the Indian minorities in each country, including their numbers, L1s, occupational profiles, social standings, and their connections with India. Interestingly, the results of this investigation mirror the pattern that emerges from the corpus studies presented in the three articles: the closer the connections a Southeast Asian country has with India, the closer its use of the studied features are in relation to those of IndE. Therefore, there is socio-historical evidence that supports the hypothesis of IndE functioning as an emerging epicentre for some of the English varieties spoken in Southeast Asia.

The fourth and final article included in this study uses the apparent-time method to examine how established one of the features, the use of clause-final *also* and *only*, is in IndE, HKE and PhiE (for this article, SinE had to be excluded due to the lack of metadata in the ICE-corpus). The results of this investigation indicate that the feature is more established in IndE, where it is used more by older and male speakers, whereas in HKE and PhiE the feature is used more by younger and female speakers, which in turn suggests that the feature has been taken up more recently in the two varieties. Therefore, it seems that, at least in the case of clause-final *also* and *only*, the use of the innovative feature in IndE predates those of HKE and PhiE and hence, the results lend support to the argument that IndE could have contributed to the growing use of the feature in some Southeast Asian varieties.

As the results of the present study show, there is some evidence that supports the idea that IndE has extended its influence to some of the English varieties spoken in Southeast Asia. For the varieties included here, the evidence appears to be strongest for SinE, while for HKE the results seem inconclusive, though the possibility of IndE's influence could not be excluded completely. For PhiE, there seems to be no indication of IndE's influence and thus, the results align with the original hypothesis of this study.



# TIIVISTELMÄ

Aasiassa puhutaan yli 2000 kieltä, ja vaikka näistä suurin osa on paikallista alkuperää, joukossa on myös kieliä kuten englanti, jotka saapuivat paikalliseen kielimaisemaan kolonialismin seurauksena. Sen jälkeen kun englanti oli istutettu moniin, usein monikielisiin yhteisöihin ympäri Aasiaa, kieli jatkoi kehittymistään paikallisesti vastatakseen uusien käyttäjiensä kommunikatiivisiin tarpeisiin. Vuosien kuluessa monet näistä englanneista ovat nativisoituneet, eli ne ovat kehittäneet omia yksilöllisiä paikallisia kielellisiä profilejaan, jotka erottavat ne muista englannin varieteeteista. Nykyään monet näistä uusista englanneista ovat virallisia kieliä alkuperämaissaan, ja valtaosa väestöstä puhuu niitä usein toisena kielenään. Yksi suurimmista tällaisista maista on Intia, jossa englanti toimii yhtenä virallisista kielistä. Aikaisempien tutkimusten mukaan tämä paikallinen nativisoitunut englannin varieteetti, intianenglanti, on jo saavuttanut niin merkittävän aseman, että se on alkanut vaikuttaa muiden Etelä-Aasiassa puhuttujen englannin varieteettien kehitykseen – toisin sanoen, siitä on tullut kielellinen episentrumi. Sitä, olisiko intianenglannin vaikutus voinut levitä myös joihinkin Kakkois-Aasiassa puhuttuihin varieteetteihin ei kuitenkaan ole vielä toistaiseksi tutkittu lähemmin, ja niinpä tämä artikkeliväitöskirja pyrkii selvittämään, olisiko tällaisesta kehityksestä mahdollisesti havaittavissa joitain merkkejä osassa Kaakkois-Aasian varieteetteja.

Tutkimuskohteiksi valittiin Intiassa, sekä Singaporessa, Hong Kongissa ja Filippiineillä puhutut englannin varieteetit. Koska kolme ensimmäistä maata ovat entisiä Britannian siirtomaita neljännen ollessa entinen Yhdysvaltojen siirtomaa, myös britti- ja amerikanenglannit lisättiin tutkimukseen vertailukohdiksi muille varieteeteille. Kyseiset Kaakkois-Aasian maat valittiin tutkimuskohteiksi koska niiden kahdenväliset suhteet Intiaan ovat olleet hyvin erilaisia kautta historian. Singaporessa on huomattava etninen intialainen vähemmistö, maa on ylläpitänyt läheisiä suhteita Intiaan aina kolonialismin ajoista saakka, ja se on näistä kolmesta Kaakkois-Aasian maasta maantieteellisesti lähimpänä Intiaa. Vaikka Hong Kong on Intian tavoin entinen Britannian siirtomaa, sillä ei ole merkittävää intialaista vähemmistöä, läheisiä historiallisia suhteita Intiaan, eivätkä nämä kaksi maata ole myöskään maantieteellisesti lähellä toisiaan. Filippiinit puolestaan ei ollut koskaan

osa brittiläistä imperiumia, jonka vuoksi sen yhteydet Intiaan pysyivät heikkoina kolonialismin aikaan. Maan intialainen vähemmistö on myös ollut aina erittäin pieni ja lisäksi se on maantieteellisesti kauimpana Intiasta. Näiden seikkojen vuoksi tämän tutkimuksen hypoteesina on, että mikäli intianenglanti olisi ulottanut vaikutuksensa Kaakkois-Aasiaan, merkkejä tästä löytyisi todennäköisimmin singaporenenglannista ja mahdollisesti hongkonginenglannista, mutta ei filippiinienenglannista.

Intianenglannin potentiaalista episentrumin roolia tarkastellaan tässä tutkimuksessa kolmen syntaktisen piirteen kautta, joiden on aikaisempien tutkimusten myötä osoitettu olevan paikallisia innovaatioita intianenglannissa, ja joita on todettu käytettävän myös joissain tässä tutkimuksessa tarkastelluissa Kaakkois-Aasian englanneissa. Tutkittaviksi piirteiksi valikoituivat lauseenloppuiset fokuspartikkelit *also* ja *only*, invariantti liitekysymys *isn't it*, sekä tendenssi suorien objektien poisjättöön lauseista, ja jokaisen piirteen käyttöä tarkastellaan omassa artikkelissaan, jotka on kaikki liitetty osaksi tätä väitöskirjaa. Tutkimuksen pääasiallinen fokus on puhutussa kielessä, jossa syntaktisten innovaatioiden esiintyvyys on usein suurinta. Tutkimuksessa käytetty data tulee suurimmaksi osaksi *International Corpus of English* -korpusperheestä, joka tarjoaa vertailukelpoista dataa monista englannin varieteeteista, mutta joissain tapauksissa myös *Freiburg-Brown Corpus of American English* ja *Santa Barbara Corpus of Spoken American English* -korpuksia käytettiin tilanteissa, joissa vertailukelpoista ICE-dataa ei ollut saatavilla.

Tutkimuksen tulokset osoittavat, että kaikkien kolmen syntaktisen innovaation esiintyvyys on suurinta intianenglannissa ja toiseksi korkeinta singaporenenglannissa, jota puolestaan seuraa, piirteestä riippuen, joko hongkongin- tai filippiinienenglanti. Varieteettien tärkeimpiä substraattikieliä tarkasteltaessa käy ilmi, että intianenglannin kohdalla substraativäikutus on todennäköisin selitys piirteiden esiintymiselle varieteetissa, mutta Kaakkois-Aasian varieteettien kohdalla tilanne on monimutkaisempi. Singaporen- ja mahdollisesti hongkonginenglannin kohdalla intianenglannin vaikutus voisikin osaltaan tarjota kattavamman selityksen tutkimuksessa raportoiduille tuloksille, kun taas filippiinienenglannin kohdalla ei havaittu merkkejä intianenglannin vaikutuksesta. Tämän vuoksi voidaankin siis todeta, että tässä tutkimuksessa esitetyt tulokset osaltaan tukevat väitettä, että intianenglannin vaikutus olisi voinut levitä myös Kaakkois-Aasian englanteihin, erityisesti singaporenenglantiin.

Selvittääkseen mistä nämä tulokset voisivat johtua, tämä tutkimus sisältää myös yksityiskohtaisen selvityksen Intian ja näiden kolmen Kaakkois-Aasian maan historiallisista, kulttuurisista, kielellisistä ja sosiaalisista yhteyksistä; erityisen huomion kohteena ovat jokaisen maan intialaisten vähemmistöjen koot, äidinkielet,

ammattilliset profiilit, sosiaaliset asemat sekä yhteydet Intiaan. Tämän selvityksen loppuhavaintona voidaan todeta, että mitä tiiviimmät yhteydet maalla on Intiaan, sitä lähempänä kolmen kielellisen piirteen käyttö on intianenglannin tarjoamaa mallia. Tämän vuoksi voidaankin sanoa, että on olemassa myös epäsuoraa sosiohistoriallista näyttöä siitä, että intianenglannin kielellinen vaikutus olisi levinnyt myös joihinkin Kaakkois-Aasiassa puhuttuihin englannin varieteetteihin.

Neljäs ja viimeinen väitöskirjaan liitetty artikkeli hyödyntää näennäisaikametodia selvittääkseen kuinka vakiintunutta yhden piirteen, eli lauseenloppuisten *also*- ja *only*-fokuspartikkelien käyttö on intian-, hongkongin- ja filippiinienenglannissa (singaporenenglantia ei voitu sisällyttää tähän artikkeliin ICE-korpuksesta puuttuvan metadatan vuoksi). Tutkimustulokset osoittavat, että piirre on vakiintuneempi intianenglannissa, jossa sitä käyttävät enimmäkseen miehet ja ikääntyneet henkilöt. Hongkongin- ja filippiinienenglannissa kyseistä piirrettä puolestaan käyttävät enemmän naispuoliset ja nuoret, joka viittaa siihen, että piirrettä on alettu käyttää näissä kahdessa varieteetissa vasta hiljattain. Tämän vuoksi voidaan sanoa, että tulokset ainakin lauseenloppuisten *also*- ja *only*-fokuspartikkelien osalta tukevat käsitystä siitä, että innovatiivisen piirteen käyttö on alkanut intianenglannissa muita aikaisemmin, ja se on siten voinut osaltaan vaikuttaa vastaavien piirteiden käyttöönottoon muissa Kaakkois-Aasian varieteeteissa.

Kuten tämän tutkimuksen tulokset osoittavat, on olemassa näyttöä siitä, että intianenglannin vaikutus on voinut yltää joihinkin Kaakkois-Aasiassa puhuttuihin englannin varieteetteihin. Tämä vaikutus näyttää olevan voimakkainta singaporenenglannissa, kun taas hongkonginenglannin kohdalla todistusaineistoa vaikutuksesta ei löytynyt, vaikkei intianenglannin vaikutusta pystyttykään kokonaan sulkemaan pois. Filippiinienenglannin kohdalla ei voitu havaita mitään merkkejä intianenglannin vaikutuksesta, ja siten lopulliset tutkimustulokset ovat yhteneväiset alkuperäisen tutkimushypoteesin kanssa.





# CONTENTS

1	Introduction.....	23
1.1	Background and research environment .....	24
1.2	Objectives and scope of the study.....	25
1.3	Structure of the dissertation.....	27
2	Theoretical foundations.....	28
2.1	Modelling World Englishes .....	28
2.1.1	Kachru's Three Circles Model .....	29
2.1.2	Mair's World System of Standard and Non-Standard Englishes .....	33
2.1.3	Schneider's Dynamic Model .....	35
2.1.4	Summary of models and their relevance to the present study.....	39
2.2	Contact environments .....	41
2.2.1	Hong Kong.....	41
2.2.1.1	History of Hong Kong.....	41
2.2.1.2	Linguistic ecology of Hong Kong .....	42
2.2.2	India .....	46
2.2.2.1	History of India .....	46
2.2.2.2	Linguistic ecology of India.....	48
2.2.3	The Philippines .....	52
2.2.3.1	History of the Philippines .....	52
2.2.3.2	Linguistic ecology of the Philippines .....	54
2.2.4	Singapore.....	57
2.2.4.1	History of Singapore.....	57
2.2.4.2	Linguistic ecology of Singapore .....	59
2.3	Defining epicentres .....	63
2.4	Previous research on IndE influence in Asia.....	66
3	Data and Methodology .....	72
3.1	Data .....	72
3.1.1	The International Corpus of English.....	72
3.1.2	The Santa Barbara Corpus of Spoken American English .....	77
3.1.3	The Freiburg-Brown Corpus of American English .....	77
3.1.4	Summary of the data .....	78
3.2	Methods .....	78

4	Results.....	82
4.1	Research question 1.....	82
4.1.1	Clause-final <i>also</i> and <i>only</i> .....	82
4.1.2	Invariant tag <i>isn't it</i> .....	84
4.1.3	Omission of direct objects.....	85
4.1.4	Summary of the answer to research question 1 .....	86
4.2	Research question 2.....	87
4.3	Research question 3.....	88
4.3.1	Properties shared with the superstrate.....	88
4.3.2	Properties shared with the substrate .....	90
4.3.3	History of Indian communities in Southeast Asia .....	96
	4.3.3.1 Singapore .....	97
	4.3.3.2 Hong Kong.....	102
	4.3.3.3 The Philippines .....	105
4.3.4	Summary of the answer to research question 3 .....	108
4.4	Research question 4.....	111
5	Discussion.....	116
5.1	Theoretical implications .....	116
5.2	Practical implications .....	118
5.3	Recommendations for future research.....	119
6	Conclusion .....	122
	References .....	123
	Appendix A.....	135
	Appendix B .....	141

## List of Figures

Figure 1.	Kachru's Three Circles Model (Crystal 2003: 61) .....	30
Figure 2.	A World System of Standard and Non-Standard Englishes (Mair 2013: 264) .....	33

## *List of Tables*

Table 1.	Articles utilizing the theoretical framework(s) of the models .....	39
Table 2.	Structure of the ICE corpora. Bracketed figures present the number of 2000-word files in each subsection ( <a href="https://www.ice-corpora.uzh.ch/en/design.html">https://www.ice-corpora.uzh.ch/en/design.html</a> ).....	74
Table 3.	Corpora used in the four articles .....	78
Table 4.	Order of varieties from most frequent to least frequent use in data from the ‘Private conversations’ section of the ICE-corpora .....	87
Table 5.	Summary of key aspects related to the Indian minorities in Singapore, Hong Kong and the Philippines (*estimates, no official statistics available on the numbers and languages spoken by the Indian minority) .....	109



## ORIGINAL PUBLICATIONS

- Publication I Parviainen, Hanna. 2012. Focus particles in Indian English and in other varieties. *World Englishes* 31(2), 226–247.
- Publication II Parviainen Hanna. 2016. The invariant tag *isn't it* in Asian Englishes. *World Englishes* 35(1), 98–117.
- Publication III Parviainen, Hanna. 2017. Omission of direct objects in New Englishes. 2016. In Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds.), *Changing English: Global and Local Perspectives* 92, 129–153. Berlin, Boston: Mouton de Gruyter.
- Publication IV Parviainen, Hanna & Robert Fuchs. 2018. 'I don't get time only': An apparent-time investigation of clause-final focus particles in Asian Englishes. *Asian Englishes* 21(3), 285–304.



# AUTHOR'S CONTRIBUTIONS

Publication IV The author was responsible for identifying the studied features in the data, whereas Fuchs was responsible for their statistical analysis. For the article, the sections were written as follows: the author is the primary writer of sections 2, 3.1, 5.1 and 5.2, Fuchs is the primary writer of sections 4, the rest of section 5, and section 6, whereas sections 1 and 3.2 were written in collaboration. Despite this division of work, it should be noted that both authors also contributed to the editing of the other's sections during the writing process.





# 1 INTRODUCTION

South and Southeast Asia have a long history of serving as melting-pots for hundreds of indigenous cultures and languages, and thus, it is not surprising that when the English language was first introduced to the region at the dawn of British colonialism, it took root quickly. As the language began to adapt to the needs and preferences of its local users, it also diversified and became nativized, giving rise to the present situation where Asia is now home to several different varieties of English, which are sometimes called New Englishes.<sup>1</sup> Today, there is a great body of work that focuses on mapping the numerous local features found in these New Englishes, and while some features are unique and the result of English, the ‘foreign’ language, adapting to its new cultural and linguistic environments, others can be found in a number of English varieties spoken around the world – the explanations for their origins can vary greatly, ranging from substrate influence to learner features and language universals. In addition to the diversification of the varieties spoken around the world, signs of another important trend can be observed regarding the numbers of their users; although many native speaker varieties such as British (BrE) and American English (AmE) have traditionally been considered to set the norms for other varieties, especially for New Englishes, the number of people who now speak English as a second or a foreign language has already surpassed the number of native speakers (Crystal 2003a: 108). As a consequence, some non-native varieties, such as Indian English (IndE), a variety which Crystal (2003a; 108–9) notes to have one of the largest speaker populations in the world, could be expected to become more influential in the future, even becoming linguistic epicentres with the power to influence the development of other English varieties they are in contact with.

This study will explore these two aspects, the study of local innovative features in New Englishes and the potential epicentric status of IndE through a detailed study of the frequency and origins of use of three features that previous studies have

---

<sup>1</sup> The term refers to varieties of English that are spoken in countries where English has the status of an official second language, often as a result of colonialism, and where it is not the first language of the majority of the population.

argued to be syntactic innovations that are commonly used in (but not necessarily restricted to) IndE, Hong Kong English (HKE), Philippine English (PhiE) and Singapore English (SinE). The hypothesis of the study is that if there is evidence of IndE having contributed to the rise of the use of the three features in HKE, PhiE and SinE, this would lend support to the argument that IndE is an emerging epicentre in the region. The three syntactic features examined here are the use of clause-final focus particles *also* and *only*, the use of the invariant tag *isn't it*, and the tendency to omit direct objects. The main focus of the study is on the use of these features in spoken language, which is where syntactic innovations occur most frequently, but for the first two features, samples of written language have also been examined.

The research reported here is based on four peer-reviewed papers, three of which have been published as articles in journals (A1 (Parviainen 2012), A2 (Parviainen 2016), and A4 (Parviainen & Fuchs 2018)) and one as a chapter in a book (A3 (Parviainen 2017)). Together, the four publications help determine the spread and frequency of the three innovative syntactic features in IndE, HKE, PhiE and SinE, while also providing suggestions for their potential origins. Although articles A2 and A3 also include some New Englishes spoken outside Asia<sup>2</sup>, these varieties will not be discussed further in the present study, which focuses on exploring the dynamics between IndE and Southeast Asian Englishes. However, two other non-Asian varieties, BrE and AmE (studied in A1–A3), have been retained in this study, since they function as points of reference for the four Asian Englishes.

## 1.1 Background and research environment

The English language is becoming increasingly pluralistic, and since a growing number of its users are now located in Asia, the continent can be expected to have a significant impact on the development of the language in the future. This view is supported by Lim and Ansaldo (2012: 260), who argue that

[i]n future decades, scholars considering the history of English will surely view the twenty-first century a crucial era for developments in the language, in particular in

---

<sup>2</sup> The non-Asian varieties included in articles A2 and A3 are Fijian English (FjE), Kenyan English (KenE) and Jamaican English (JaE).

Asia. This is not only because currently the region is geopolitically the site of economic power but also because it is the venue for the largest and most quickly growing number of English users...

Indeed, some (see, for example, Smith 1998, Kachru 2005) have gone even further by arguing that English has already become an Asian language. To support this argument, Kachru (2005: 15) presents the following four demographical facts:

1. the total English-using population of Asia is more than that of the Inner Circles, including Australia and New Zealand.
2. India, in the Outer Circle, is a major English-using country along with the UK and the US
3. English is the main medium in demand for acquisition of bilingualism/multilingualism in the whole Asian region.
4. in parts of Asia (e.g. in Singapore) English is gradually acquiring the status of the dominant language or the *first* language, whatever we mean by that term.

The growth in the number of English speakers in Asia can also be expected to have an effect on the dynamics between different English varieties, resulting in a shift where some of the power held by the traditional “core” L1 varieties such as BrE and AmE is transferred to the L2 varieties with the largest number of speakers. The fact that such a shift might already be taking place is also reflected in the more recent models that have been developed to describe the dynamics between different varieties of English (see section 2.1). Since IndE has one of the largest numbers of speakers in the world (Crystal 2003a: 108–9), examining whether the variety could extend some of its influence on other varieties spoken in the region – that is, if it could function as a linguistic epicentre – provides an interesting case study that will help shed light on the role that Asia will play in the future development of English. Signs of IndE’s growing influence in South Asia have already been found by, for example, Gries and Bernaisch (2016) and Hundt et al. (2012) and the aim of this study is to see whether IndE’s reach could extend even further, to Southeast Asia.

## 1.2 Objectives and scope of the study

The potential epicentric influence that IndE could have in the region will be examined through the use of three syntactic innovations that previous studies have argued to be frequent in IndE, HKE, PhiE and SinE. The study will first determine

the frequency in which the features are used in the varieties and then seek to trace their origins, considering various possible explanatory factors such as superstrate and substrate influence. Furthermore, the study will attempt to provide an approximate timeline of the development of IndE's influence in Southeast Asia, in addition to making some tentative suggestions regarding the future role that the variety could have in the region. The study will seek to provide answers to these questions with the following research questions:

1. Can the use of clause-final *also* and *only*, invariant tag *isn't it* and the omission of direct objects be argued to be syntactic innovations in IndE?
2. How frequently are these features used in IndE, HKE, PhiE and SinE?
3. What could explain the differences between the varieties?
4. What is the future role of IndE in Southeast Asia?

The rationale for these questions is the following. Firstly, before it can be argued that any feature used in the three Southeast Asian varieties is the result of IndE influence, it must be established that the feature is indeed a local innovation in the variety. For this argument to hold true, the possibility of superstrate influence must be excluded, while a similar structural pattern should be located in any of the major substrates of the variety. After this has been established, the second step is to examine how frequently the same syntactic feature is used in HKE, PhiE and SinE. For a feature to be considered as a possible example of IndE's epicentric influence, it could be expected to be used most frequently in IndE, followed by SinE and/or HKE since all three regions were once part of the British colonial empire and hence, their connections have been stronger when compared with other Southeast Asian countries that were not colonised by the British. By the same token, PhiE could be expected to have the lowest use of the feature, since the country was colonised by the US. Even if the results of the analysis of the three syntactic features follow this order, considering other explanatory factors in addition to the influence of IndE is paramount, and because of this, the third step involves identifying and evaluating other possible factors – such as substrate influence – that could explain the emergence of the feature(s) in the varieties. Based on these findings, the study will fourthly present some suggestions for the role IndE could have in Southeast Asia in the future. Furthermore, in addition to providing answers to the four research questions, this study will also present a detailed investigation of the sociolinguistic histories of India, Hong Kong, the Philippines and Singapore, while also briefly

commenting on how the results relate to some of the different models that have been created to describe World Englishes and their normative orientations.

### 1.3 Structure of the dissertation

The remainder of the study is organised as follows: chapter two begins by introducing the theoretical models used in the four articles, which is followed by a description of the histories and linguistic ecologies of India, Hong Kong, the Philippines and Singapore. In addition, the chapter includes a discussion on terminological issues and an overview of the previous research that has examined the role of IndE in Asia. Chapter three introduces the materials and methods used in the four articles (A1–A4). Chapter four focuses on the research questions (presented above) and the answers provided by the combined results of A1–A4. This discussion is then followed by a more detailed investigation into the presence and role(s) Indians have had in Hong Kong, the Philippines and Singapore. Chapter five presents a brief discussion on the theoretical and practical implications of the study, in addition to commenting on the reliability and validity of the results and providing some suggestions for future research. The study then concludes with a few closing remarks in Chapter six.

## 2 THEORETICAL FOUNDATIONS

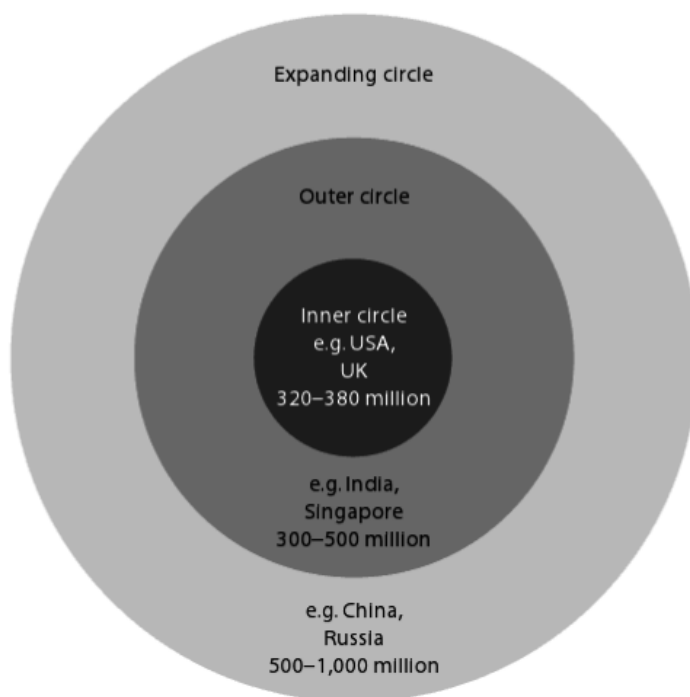
This chapter presents the theoretical models that have been used in articles A1–A4, paying special attention to how the models comment on the normative orientations of different English varieties. This is followed by a description of the histories and contact environments of the Englishes included in the present study. Furthermore, definitions for some core terminology is also presented and their strengths and possible drawbacks are commented on briefly. The section concludes with a discussion on the findings of previous research that has been conducted on this and related topics on epicentric influence in World Englishes.

### 2.1 Modelling World Englishes

Several models have been developed to describe and analyse the English language complex that exists in the world today. Though models such as Stevens's (1980) *Tree model*, McArthur's (1987) *Circle of world English* and Görlach's (1990) *Circle model of English* have in many ways laid the foundation for the more recent models used in the field, the current section will focus on three models that are also referred to in articles A1–A4: Kachru's (1985) *Three Circles model*, Mair's (2013) *World System of Standard and Non-standard Englishes* and Schneider's (2003, 2007) *Dynamic Model*. This choice is based on their level of establishedness in the field of World Englishes studies and on their applicability to the research questions at hand. The order in which the models are discussed is partially based on chronology, but factors related to their extensiveness have also been taken into consideration: Kachru's (1985) model is both the oldest of the three and also the most extensive in its scope, as it encompasses all users of English be they native (ENL), second language (ESL) or foreign language (EFL) speakers. Although Mair's (2013) model is the newest of the three discussed in this study, it is introduced after Kachru's, because it has the second widest scope of English users (ENL and ESL). Schneider's (2003, 2007) model, though predating Mair's (2013), is discussed last, as it has the narrowest focus of the three, focussing only on postcolonial varieties of ENL and ESL.

### 2.1.1 Kachru's Three Circles Model

One of the most influential models describing and categorizing the large number of Englishes spoken around the world is Kachru's (1985) Three Circles model, which divides English varieties into three concentric circles (see Figure 1) based on their linguistic, sociolinguistic and acquisitional qualities (Kachru 1992a: 232). The Inner Circle is formed by native speaker (L1) varieties of, for example, Britain, Ireland, New Zealand and the United States, which Kachru (1992b: 356) argues to be "the traditional cultural and linguistic bases of English". Estimates of the number of speakers in the Inner Circle vary between 320 and 380 million (Crystal 2003a: 107). The Outer Circle consists of English varieties spoken in such countries as India, Fiji and Kenya, where the language functions as an institutionalised second language (L2) (Kachru 1992a: 356). For Outer Circle Englishes, the range of estimates of their speakers is even wider, between 200 and 500 million (Crystal 2003a: 107). The final, Expanding Circle contains the remainder of English speakers in the world in countries such as Germany, Guatemala and China, where English has no official status and where it is spoken as a foreign language. The estimates of the number of speakers in the Expanding Circle are even more tentative, ranging from 500 to 1,000 million (Crystal 2003a: 107).



The three 'circles' of English

**Figure 1.** Kachru's Three Circles Model (Crystal 2003b: 61)

The classification of English varieties on the three concentric circles is also closely linked to the concept of two English diasporas.<sup>3</sup> In the first diaspora, large numbers of monolingual people from Britain moved to areas such as present-day Australia, Canada, New Zealand and the United States, where they replaced much of the local population; in these countries, English functions as the *de facto* language of education, government and the society as a whole and therefore, varieties that emerged from the first diaspora can all be placed in the Inner Circle (Kachru & Smith 2008: 4-5).

Varieties located in the Outer Circle are the result of the second diaspora, when English spread to areas of modern-day India, Kenya and Singapore, to name but a

---

<sup>3</sup> Here it should be noted that the two diasporas do not represent two different time periods, but two different types of migration patterns.



few. In these regions, the patterns of immigration differed from the previous in many ways, as the proportion of English-speaking colonisers remained low, mostly consisting of officers, businessmen and missionaries. Even though English never superseded the local languages in these regions (Kachru & Nelson 2006: 9–10), after the English-speaking minorities gained military, economic and political dominance over these regions, English became established as the only official language used in education, commerce and governance, which, in turn, resulted in a situation where “[English] alone could open up careers and other paths to social and economic advancement”, as Leitner (1992: 205) notes. As these colonies (re)gained their independence, the role of English had usually been rooted in the societies to such an extent that it was retained as one of the official languages, which subsequently led to the nativization of the local varieties during the decades that followed.

A case for a third diaspora has been argued for the countries in the Expanding Circle (see, for example, Ho 2008). In contrast to the Inner and Outer Circle varieties, which can trace their origins to the movement of native speakers, English has spread to the countries in the Expanding Circle because of the economic and political importance of two major English-speaking countries, the UK and the US (Kachru & Nelson 2006: 28), which has led to the present situation where English functions as a global *lingua franca*. Kachru and Nelson (2006: 28) argue that the use of English in the Expanding Circle is still mostly restricted to the spheres of technology and higher education, although they acknowledge the possibility that English will also be increasingly used in the personal domain in the future. Indeed, signs of such development have already been observed in some countries such as Finland (Leppänen et al. 2009) and the Netherlands (Edwards 2016). Here it should be noted that Kachru’s (1985) model is not static and varieties from the Outer Circle can move towards the Inner Circle, while corresponding movement from the Expanding Circle to the Outer Circle is also possible. An example of the former would be SinE, an Outer Circle variety which now has a growing number of L1 speakers (Wee 2013; Tan 2014), whereas a case for the latter has been made for Dutch English (Edwards 2016).

A further important aspect of Kachru’s model concerns the normative orientation of the varieties in the three circles; according to Kachru (1985), varieties in the Inner Circle are considered ‘norm providing’, whereas varieties in the Outer and Expanding Circles are ‘norm developing’ and ‘norm dependent’ respectively. However, when the normativity of the Inner Circle varieties is examined closer, it becomes apparent that this has often been synonymous with the norms of just two

varieties, BrE and AmE.<sup>4</sup> This can be explained by the historical developments over the past centuries: when the British Empire expanded to new continents, it established BrE as the norm-providing high prestige variety in all its colonies (e.g. India, Hong Kong, Singapore), and many BrE features have thus been retained in the local ‘standard’ Englishes that emerged during the following centuries. A corresponding development can also be observed in the English spoken in the Philippines, a former US colony, which has retained many features from AmE. For centuries, the global reach of the British Empire promoted the role of BrE as the sole global prestige variety, a status that remained unchallenged up until the 20<sup>th</sup> century and the dissolution of the British Empire. This coincided with the emergence of the US as the new superpower, which in turn elevated the status of AmE as the new prestige standard with global spread.

There is one further aspect regarding Kachru’s (1985) model which should be discussed, as it is of special interest concerning the topic of the present study. Kachru (1985: 28) suggests that Inner Circle Englishes are not the only varieties that can provide norms for the speakers in the other two circles, since people from the Outer Circle can also transmit their locally developed norms to the speakers in the Expanding Circle. This influence, according to Kachru (1985: 28), is spread by people in various professions, though the most significant group is formed by those working in the field of education. Kachru (1985: 28) also mentions the presence of South Asian teachers in Southeast Asia but he does not explicitly state whether this has resulted in the introduction of features from South Asian Englishes to the other Outer Circle varieties spoken in Southeast Asia.<sup>5</sup>

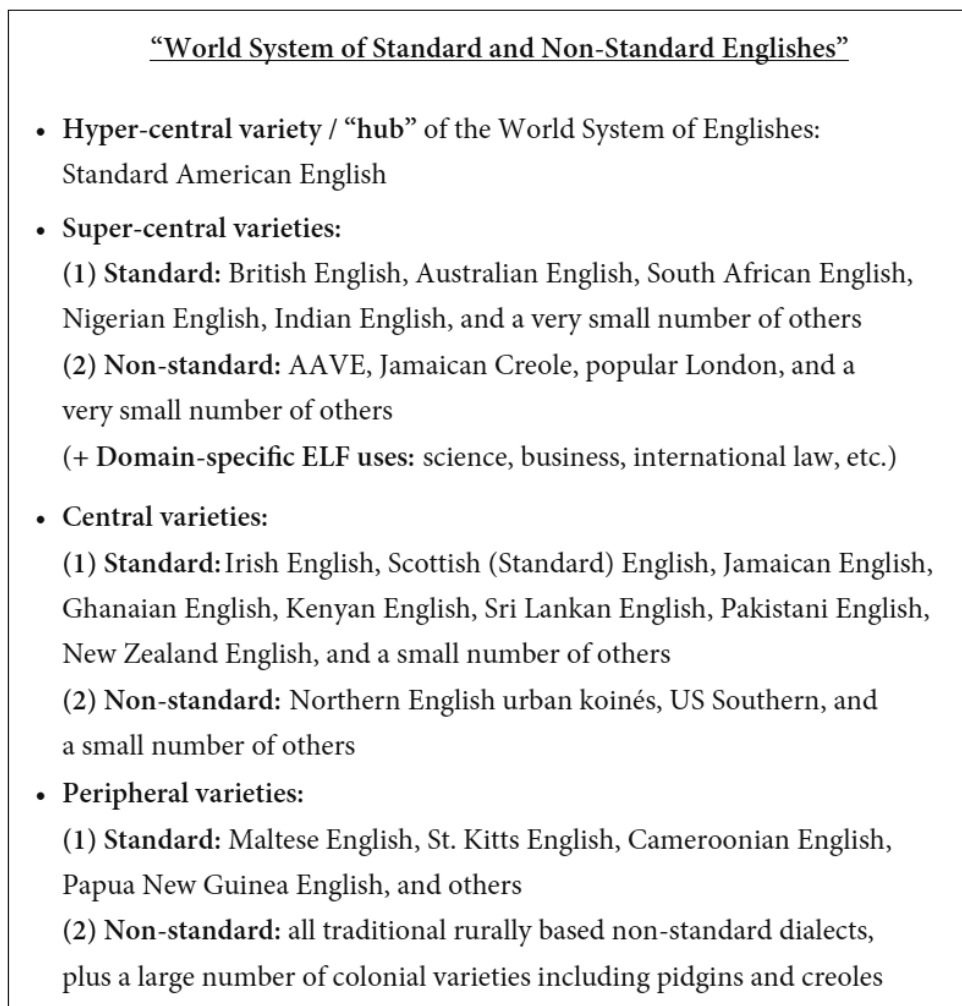
---

<sup>4</sup> Mair (2013: 257–8) criticizes Kachru’s description of the Inner Circle varieties as ‘norm providing’ by pointing out the asymmetric relationship between these varieties where features from AmE are more likely to spread to other Inner Circle varieties than vice versa.

<sup>5</sup> While it is possible that Indian teachers participated more actively in the spreading of IndE features in their new communities abroad due to their norm-providing roles as teachers, an alternative explanation that operates below the level of conscious language choices is also plausible: if a particular feature was prevalent in both the IndE spoken by the teachers and in the local variety of English spoken by their students (due to the feature existing in their respective substrates), it is possible that the teachers simply failed to detect that the feature was “non-standard” in the English(es) of the locals and therefore they did not discourage its use in the way they might have done for other local features they perceived as non-standard. This could have then indirectly contributed to the emergence of shared features in the varieties, even though technically the roots of these features lie in different substrates.

## 2.1.2 Mair's World System of Standard and Non-Standard Englishes

One of the most recent models introduced to the field of World Englishes studies is Mair's (2013) World System of Standard and Non-Standard Englishes, which is an adaptation of de Swaan's (cited in Mair 2013: 259) model for the World Language System. The model consists of four main levels, most of which can be further divided into standard and non-standard varieties (see Figure 2).



**Figure 2.** A World System of Standard and Non-Standard Englishes (Mair 2013: 264)

At the top of the hierarchy is AmE, the Hyper-central variety (or the “hub’ of the World System of Englishes” (Mair 2013: 261)), which is followed by a small number of standard and non-standard Super-central varieties such as standard BrE, Australian English (AuE), Nigerian English and IndE, in addition to the non-standard varieties of African-American Vernacular English, Jamaican Creole and popular London (Mair 2013: 264).<sup>6</sup> The third level is formed by Central varieties, which include such standard varieties as Irish English, Pakistani English (PkE) and New Zealand English (NZE), whereas non-standard varieties include, for example, Southern US English (Mair 2013: 264). The fourth and final level includes a large number of Peripheral varieties such as the standard Englishes of Malta and Cameroon, which have smaller speaker populations and which, according to Mair (2013: 264) are “all traditional rurally based Non-Standard dialects, plus a large number of colonial varieties including pidgins and creoles”.

The order in which the varieties on different levels influence one another is, according to Mair (2013: 261), the following: the Hyper-central variety has the potential to influence all other varieties, whereas influences flowing in the opposite direction will be more limited. Similar tendencies can be seen in all levels of the model, so that, for example, PkE (a Central standard variety) could be expected to receive more influences from IndE (a Super-central standard variety) than vice versa. Speakers of Peripheral varieties, in contrast, are expected to be familiar with the Hyper-central variety, in addition to having some level of knowledge of a number of Super-central and Central varieties.

The strength and appeal of Mair’s (2013) model lies in the fact that it discards the view of native English speakers as the ‘owners’ of the language, who can define its rules for the other speakers around the world. A good example of this is the addition of IndE, an Outer Circle variety in Kachru’s Three Circles model, into the group of standard Super-central varieties. The motivation for this, as stated by Mair (2013: 263), is the following:

There is (as yet) anecdotal evidence that hundreds of thousands of expatriate Indians working abroad in business and information technology are beginning to leave their mark on British and American English. It would be interesting to note whether migration, modern communication and media technology, combined with a craze for Bollywood-style entertainment, will be sufficient to establish Indian English norms as one relevant factor in the future development of the varieties spoken in the Indian diaspora communities in Asia, the Pacific and the Caribbean, whose varieties of

---

<sup>6</sup> In addition to standard and non-standard Englishes, Mair (2013: 264) also adds domain specific ELF uses under this category.

English developed on the basis of mother-tongue substrates but without significant input from Indian English (as it was at the time of emigration).

A further advantage of Mair's (2013) system is that it takes into account both the standard and non-standard varieties of English, and the fact that the spheres of influence are not restricted to geographically proximate varieties. Indeed, one of the consequences of the increasing movement of people, recent developments in the IT sector and global trends in popular culture is that the spheres of any variety's influence can extend to speaker populations that are physically distant from one another, a factor that has frequently been neglected in the older models.

### 2.1.3 Schneider's Dynamic Model

The Dynamic Model by Schneider (2003, 2007) is based on the notion that all postcolonial Englishes evolve through a "fundamentally uniform process" (Schneider 2007: 32), which can be traced by following the developments in the identities and languages used by the indigenous (IDG) population and the English-speaking settlers (STL). According to the model, the (slowly) converging identities and language practices of the two groups are reflected in the local linguistic landscape, which can ultimately lead to the emergence of a new variety of English. Schneider (2003, 2007) divides the development of postcolonial Englishes into five consecutive phases which are called 'foundation' (phase 1), 'exonormative stabilisation' (phase 2), 'nativization' (phase 3), 'endonormative stabilisation' (phase 4) and 'differentiation' (phase 5). For each phase, Schneider (2007) provides details of the developments that are taking place in the two populations, ranging from their identity constructions and the linguistic features used to the socio-political and sociolinguistic conditions of the whole colony.

In the foundation phase (phase 1), according to Schneider (2003: 244), a group of English speaking settlers move to a non-English speaking area for an extended period of time, often following the foundation of a military fort, a trading station or an emigration settlement.<sup>7</sup> At this stage, the relations between the IDG and STL

---

<sup>7</sup> Whether the settlement became a settler or a trade/exploitation colony was often determined by the climate: as Evans (2014: 577) notes, the (sub)tropical climate in Asia was badly suited for European agriculture and as a consequence, these "malign, diseased environments" received only a fraction of Europeans when compared with such settler colonies as the US and Australia, where the Europeans marginalised the indigenous populations quickly.

populations can vary from hostile to friendly, and the identities of the two groups remain separate, the settlers considering themselves as expatriate citizens of their home country, while members of the indigenous population view themselves as the only “true” inhabitants of the country (Schneider 2007: 33–4). This coexistence gives rise to a complex new contact situation which Schneider (2003: 244) notes to function on two levels: the first involves the new dialect contacts among the settlers, who frequently come from various parts of the old motherland, whereas the second level refers to the linguistic interaction between the STL and IDG populations. During the foundation phase, the language of the settlers becomes more uniform (koinéization) and some lexical items related to, for example, flora, fauna and place names are adopted from the indigenous language(s) (Schneider 2007: 35–6). Since only a very small minority of the IDG population will be bilingual at this stage, usually those who work as mediators between the IDG and STL populations, pidginization can emerge, especially in trade colonies (Schneider 2007: 34–6).

During the exonormative phase (phase 2), the role of English is established further as it becomes the primary language of all major institutions such as government, education and law (Schneider 2007: 36). At this stage, according to Schneider (2003: 246–7), the identities of the IDG and STL populations begin to expand and the settlers acquire a new, positive ‘English-cum-local’ identity which separates them from their compatriots who do not share their experiences of living overseas. Also, the IDG strand experiences some significant changes; as the English language has now been established in all major institutions of the colony, the ability to speak English is taken up by the higher social strata of the IDG population, who now view the language as a means to improve their economic, cultural and social standing in the colony (Schneider 2003: 246). As Schneider (2007: 37) importantly notes, this newly acquired skill will become a “source of some pride, and at least amongst the higher echelons of the indigenous society at this stage we find the beginnings of the segregational elitism that characterizes English in some PCE-speaking countries to the present day”. As a growing number of the IDG population adopts English, signs of structural nativization also begin to emerge, predominantly in the spoken vernaculars, though most of the changes remain undetected by the speakers at this stage (Schneider 2003: 246, 2007: 40).

The nativization phase (phase 3) is marked by increased economic, political and/or linguistic independence from the colonial motherland, and, according to

Schneider (2007: 41), many colonies gain their independence during this time.<sup>8</sup> The emerging independence also brings the identities of the IDG and STL populations closer together and a new sense of ‘us’ begins to develop. The growing interaction between the IDG and STL strands leads to increased mutual accommodation and, eventually, a shared variety which is spoken as an L1 by some and an L2 by others (Schneider 2007: 45).<sup>9</sup> At this stage, a growing number of local forms are accepted in the emerging new variety, many of these features originating from the English spoken by the IDG population, which is heavily influenced by the substrate(s) (Schneider 2007: 42–5). Although lexical borrowing might be the most visible result of this process, also morphological and syntactic restructuring occurs, giving rise to structural innovations in the emerging variety (Schneider 2007: 44–5). The nativization phase is also marked by the emergence of the ‘complaint tradition’, which in reality are, as Schneider (2007: 43) notes, “class struggles in disguise”.<sup>10</sup>

Although political independence frequently functions as a prerequisite for a variety to enter the endonormative stabilization phase (phase 4), a feature even more fundamental to this development is what Schneider (2007: 48) calls the “cultural self-reliance” of the population. Though the transition from phase three to four can be peaceful, it can also be instigated by “some exceptional, quasi-catastrophic political event which ultimately causes the identity alignment of STL-strand speakers to switch from a self-association with the former mother country... to a truly independent identity” (Schneider 2003: 250), which Schneider (2007: 49) labels as ‘event X’. Following the disenchantment of the STL population with the former mother country, the strict boundaries between the IDG population and the STL strand become less pronounced and a new, local identity begins to emerge, which Schneider (2007: 49) notes to signal the birth of a nation. These changes are also reflected on the local variety, where the use of a growing number of local (mostly lexical, fewer structural) features from the IDG population are accepted and evaluated positively as signs of linguistic independence (Schneider 2007: 49–50). During the endonormative phase, the term “English in X”, which was used to

---

<sup>8</sup> In fact, Greenbaum (cited in Schneider 2003: 247) argues political independence to be “a precursor of linguistic independence”.

<sup>9</sup> However, Schneider (2007: 45) notes that the difference between L1 and L2 speakers of the variety will eventually disappear and that the differences between the two populations frequently become more sociolinguistic.

<sup>10</sup> According to Schneider (2003: 248), the complaint tradition refers to the criticism presented by the more conservative speakers of English towards the local variety of English, which, they claim, has falling standards and ‘corrupt’ usage.

describe the language spoken in the region during the first three phases, becomes “X English”.<sup>11</sup> This phase is marked by a heightened emphasis of the new variety’s homogeneity, which, as noted by Schneider (2007: 51), results from the needs of the speakers to strengthen their new identity. During this time, there is also a growing trend to publish dictionaries (followed by user guides and grammar books) of the local form of English, which, according to Schneider (2003: 252), responds to the community’s need to codify the new variety.

In the final, differentiation phase (phase 5), the local form of English has reached ‘maturity’ and is now considered a full-fledged variety with norms and rules of its own. During this time, the differences in the type of English used by members of the STL and IDG populations often reappear, which, according to Schneider (2007: 54), “depends upon the amount of bi- or multilingualism that has survived phase 4 developments”. This phase is also marked by the birth of new dialects within the variety as the processes of “group-internal linguistic accommodation” (Schneider 2007: 57) result in the emergence of newly formed subgroups that align themselves according to various sociolinguistic factors such as age, gender, social status and regional background (Schneider 2003: 254). Therefore, it is important to remember that even in cases where the local variety has reached ‘completion’ in the dynamic model, it does not mean that the population speaking the local variety of English has become monolingual, as Schneider (2007: 55) notes; in fact, the IDG strands in many multilingual countries continue to speak English as an “ethnic” dialect or as an L2.

---

<sup>11</sup> Here it should be noted that many postcolonial varieties which have not yet reached the endonormative phase are actually already commonly called ‘X English’ (e.g. Hong Kong English and Fiji English) instead of English in X (e.g. English in Hong Kong and English in Fiji).



### 2.1.4 Summary of models and their relevance to the present study

The references made to the three models in the four articles are summarised in Table 1.

	A1	A2	A3	A4
Kachru	-	-	-	✓
Mair	-	-	✓	✓
Schneider	-	✓	-	✓

**Table 1.** Articles utilizing the theoretical framework(s) of the models

Although direct references to Kachru’s Three Circles model are given only in the last of the four articles, the model is important as it in many ways provides the foundation on which the present-day conceptualisation of the dynamics of World Englishes has been built. The model is useful for the purposes of the present study also on a more concrete level, as it addresses the possibility of Outer Circle Englishes (institutionalized L2s) providing norms for other varieties, Kachru (1985) specifically mentioning the effect IndE has had on other Englishes in the region. Some of the shortcomings of the model in turn concern the rather general description of the dynamics of the different varieties with one another; while the case of IndE is addressed, the model does not provide answers to other more concrete questions of what might, for example, the hierarchies between different Southeast Asian varieties included in the present study be and whether some would be more receptive to influences from IndE than others. As noted in section 2.1.1, the model also fails to specify whether the norm-providing abilities of Outer Circle Englishes are restricted to the Englishes spoken in the Expanding Circle or if they can also extend their influence to other Outer Circle varieties. Therefore, although the model lends some support to the idea of IndE being a potential epicentre in Asia, it fails to provide any concrete descriptions of the factors that should be taken into consideration when examining this phenomenon closer.<sup>12</sup>

---

<sup>12</sup> A defence to this critique should also be provided, as it has always been clear that when creating the model, Kachru was more interested in challenging the inequalities that existed between the Inner, Outer and Expanding Circle Englishes and less concerned with details related to, for example, the proficiency levels of the speakers of different varieties, or the dynamics that govern the interactions between different varieties.

As with Kachru's Three Circles model, some of the strengths of Mair's (2013) World System of Standard and Non-Standard Englishes are ideological; by eradicating such preconceived notions as "postcolonial" and "native" often present in earlier models, Mair creates a more egalitarian framework for the study of World Englishes. A further benefit of using Mair's (2013) model is that it provides a more detailed description of the hierarchies between different English varieties, which helps to position the potential "weight" that IndE has in relation to many other Outer Circle varieties spoken in Asia. Therefore, in theory, it should be possible to use the framework offered by the model to establish the different dynamics that exist between IndE and SinE, HKE and PhiE. Unfortunately none of the three Southeast Asian varieties are included in the model and thus, even though the model lends support for the argument that IndE, a standard Super-central variety, could function as a model for some varieties, it remains unclear where in the model SinE, HKE and PhiE would be in relation to IndE and each other, and how receptive they would be to influences from IndE. Furthermore, Mair (2013) does not lay out the exact criteria according to which different varieties are assigned to the different levels, making it challenging to offer suggestions as to where exactly the three Southeast Asian varieties included in the present study would be located in the model.<sup>13</sup>

The key motivation for including Schneider's (2003, 2007) Dynamic Model in the study is that it provides a detailed investigation into the normative orientations of all the postcolonial varieties included in the present study. Firstly, it helps to determine which varieties have already become endonormative, a prerequisite for any variety to function as a linguistic epicentre (see also discussion in section 2.3). Secondly, the location of a variety in the model, especially whether it has already reached or passed the structural nativization phase, can provide some indication as to the level of acceptability of nativized features among its speakers. While this does not automatically result in all localised features becoming accepted by speakers of the variety, structural nativization increases the overall chances of local features being used even in more formal situations, which in turn increases their chances of spreading also internationally. Thirdly, the model is useful since it helps to determine which varieties are still in the exonormative phase and thus more likely to be receptive to influences from other more established varieties they are exposed to. The locations of IndE, HKE, PhiE and SinE in the Dynamic Model will be discussed further in the subsections 2.2.1–4.

---

<sup>13</sup> Again, a defence to Mair's model could be offered by noting that it has been developed fairly recently and hence, it is possible that such additions will be introduced to the model in the future.

## 2.2 Contact environments

The contacts between South and Southeast Asian people and languages have been diverse throughout centuries and thus warrant further investigation. The following subsections will provide brief summaries of the histories of Hong Kong, India, the Philippines and Singapore, their linguistic ecologies and of the processes through which the local English varieties have developed over the centuries.

### 2.2.1 Hong Kong

#### 2.2.1.1 History of Hong Kong

The Hong Kong Special Administrative Region of the People's Republic of China is a former British colony covering some 1,104 km<sup>2</sup> of land at the Pearl River delta by the South China Sea. The name Hong Kong is Cantonese for 'Fragrant Harbour' (*Heung Gong*) (Ingham 2007), but the exact origins of the name have been lost in history. Even though the area was never at the centre of Chinese politics or trade in the precolonial era, the domestic conflicts of the 17<sup>th</sup> century caused major setbacks to the region and the area remained in a desolate state until the arrival of the British.<sup>14</sup>

Hong Kong became a British Crown Colony during the First Anglo Chinese war (also known as the First Opium War (1839–42)). At the time of the annexation, Hong Kong consisted of small villages and its population totalled under 7,500, most of whom were fishers and farmers (Tsang 2004: 16). According to Tsang (2004: 22), the interests of the British in East Asia were mostly related to "trade and economic benefits rather than territorial acquisition" and to further these aims, the British declared Hong Kong a free harbour (Sivonen 2006: 49; Tsang 2004: 21–2).

The area today known as Hong Kong is the result of three different treaties signed between Britain and China: the Treaty of Nanking (1842), which included the Hong Kong island, the Convention of Peking (1860), which added the Kowloon peninsula and, finally, the second Convention of Peking (1898), which incorporated the Lantau island and the New Territories to the colony with a 99-year lease. With a secured

---

<sup>14</sup> The emperor of the newly established Manchu dynasty, unable to control the pirate infested waters around Hong Kong, ordered the Great Clearance (1661–2) of Hong Kong, which moved some 16,000 people into forced exile – once the order was overturned, only a fraction of the original population returned (Ingham 2007: 6–7).

future for the next 100 years, Hong Kong's role in the regional trade grew and by the interwar periods, the colony had outgrown its old role as a mere transit harbour for Britain's trade with China, as noted by Sivonen (2006: 147).

During the Second World War, Hong Kong was occupied by the Japanese from the end of 1941 to 1945. Despite the approaching invasion, Hong Kong's military defence had remained undersized and thus when the offensive began, Hong Kong fell under the Japanese occupation within days (Sivonen 2006: 152). This, as Tsang (2004: 142) suggests, strengthened the anti-British and anti-colonial spirit of the locals, especially among the Chinese population.<sup>15</sup> The post-war years, however, witnessed Hong Kong quickly rising back on its feet, as China's political instability and the subsequent rise of the Communist party caused an influx of immigrants from mainland China to move to Hong Kong. Among these people were many investors from Shanghai who also brought their businesses with them, thus making a significant contribution to Hong Kong's rise among the key economic centres in Asia during the following decades (Sivonen 2006: 204, 206; Bolton 2002: 33).

In 1982, the British government commenced official negotiations concerning the future of Hong Kong with the People's Republic of China (PRC). The talks resulted in the declaration that Hong Kong would be ceded back to the PRC in 1997, and that the colony would continue as a special administrative region<sup>16</sup> for the following 50 years under the 'one country, two systems' policy. Therefore, the fate of Hong Kong differs from the other former British colonies in the region, as it is the only colony which did not become independent after the end of the British colonial rule. In 1997, Hong Kong was ceded to China as planned, and despite the conflicts with Peking after the handover, Hong Kong has so far retained its status as one of the leading economic hubs of Asia and the world.

#### 2.2.1.2 Linguistic ecology of Hong Kong

Due to the constant influx of people from the adjacent province of Canton, the majority of Hong Kong's population has always been formed by speakers of Cantonese. In 2011, almost 90% of Hong Kong's population reported Cantonese as

---

<sup>15</sup> The evacuation of women and children before the attack was accused of being racially biased and the unprepared state of the local military forces was seen as a strong message of the value the British colonial government gave to Hong Kong and its non-white population (Sivonen 2006: 152–3).

<sup>16</sup> During this time, Hong Kong would have its own government, laws, legal praxis and a capitalist economy (Sivonen 2006: 213)

their ‘usual language’, whereas the second most commonly spoken language, English, was spoken only by 3.5% of the population, although 42.6% of the respondents claimed to speak it as another language (Census and Statistics Department 2011). The third most commonly spoken language is Putonghua, the official language of the PRC, which is used as the primary language by only 1.4% (Census and Statistics Department 2011). While the proportion of English and Putonghua speakers has slightly increased over the past 10 years, the percentages for other Chinese varieties have declined during the same period. Since the handover, the aim of the language policy introduced by the PRC has been that Hong Kong people be educated as bi-literate (Chinese and English) and tri-lingual (Cantonese, Putonghua and English), English being spoken mainly as an L2. As the focus of this study is on local Asian varieties of English, the remainder of this section will focus on describing the development of HKE over the years using Schneider’s (2003, 2007) Dynamic Model.

As Schneider (2007: 133) notes, the foundation phase of HKE can be traced back to the years 1841–2 when the first British arrived in the area. From the first decades of the colony, the population began to grow exponentially, from 7,000 to 90,000, of whom less than three per cent were non-Chinese as Evans (2009: 284) notes.<sup>17</sup> Interestingly, Evans (2008: 387; 2009: 284–6) argues that the concept of STL and IDG strands cannot be easily applied to HKE, as most of the Chinese and British were not permanent residents of the colony – in fact, many still considered themselves as sojourners who only came to Hong Kong for trade.<sup>18</sup> As a consequence, the identities of the two transient groups were still strongly connected to their respective motherlands and no clear ‘settler’ and ‘indigenous’ groups were formed in the colony (Evans 2009: 284).<sup>19</sup>

The first decades of the history of Hong Kong were marked by a deep division between the Chinese and the Europeans (Evans 2009: 287–8). Bilingualism in the population was rare (Schneider 2007: 135), and though English was the sole official language of the colony, much of the communication between the British and the Chinese was conducted in Pidgin English (Evans 2009: 288). The interests of the British were predominantly commercial in the region and thus, with the exception

---

<sup>17</sup> After the British, the three other significant non-Asian groups that resided in Hong Kong during the 19<sup>th</sup> century were Parsees, Portuguese and Indians (Tsang 2004: 65).

<sup>18</sup> This also raises the question whether Evans’s (2008) problematization of these terms in the context of HKE should be extended to other trade colonies such as Singapore?

<sup>19</sup> Only the few thousand fishers and farmers who had lived in the area already before the British would meet the description of the IDG strand, but they were quickly pushed to the margins by the vast number of Chinese who moved to the colony as Evans (2009: 284) notes.

of training some local interpreters, the colonial government was not interested in the education of the locals (Evans 2008: 393). Despite this, some missionary schools<sup>20</sup> offered English-medium teaching already in the 1840s, and by the 1850's, some public schools also took up English-language education, although the number of Chinese students remained low at first (Evans 2008: 406–7). However, as Hong Kong's economy grew during the latter half of the 19<sup>th</sup> century, public opinions also changed, and soon a growing number of Chinese were learning English at both government and mission schools, as the language was seen as a means to economic and social advancement (Evans 2008: 407).

Schneider (2007: 135) locates HKE's transition to phase two to the year 1898 when the Second Convention of Beijing was signed, leasing the New Territories to Hong Kong for 99 years and creating a heightened sense of stability in the colony. This phase, according to Schneider (2007: 135), was characterized by the spread of bilingualism, though mostly only among the elites as they had better access to English education. Schneider (2007: 135) also argues that during this time, the STL and IDG strands began to develop an emerging local identity, which they viewed positively. Evans (2014: 582), however, challenges this view by stating that “only a minority of [the Chinese and British] families put down roots in the city” and thus neither of the two strands showed any signs of an emerging local identity. Furthermore, Evans (2014: 295) argues that the first signs of phase two of HKE cannot actually be detected before 1920s and 1930s, a development that came to an abrupt end at the beginning of WW II.

Interestingly, Evans (2014: 595) argues that HKE experienced a second foundation phase in the 1950s and 1960s, when 2.5 million Chinese moved to Hong Kong to escape the communist revolution in China (see also Bolton 2002: 33), as this was the first time when a large number of Chinese families lived in the colony permanently (Evans 2014: 595). This change in the structure of the population also increased the pressure the colonial government faced regarding its educational policies. This led to the introduction of a series of reforms<sup>21</sup> in the 1970s and 1980s, which, according to Evans (2014: 596), resulted in a situation where the children of the new immigrants often completed their primary studies in Chinese-medium

---

<sup>20</sup> The tradition of religious schools has continued strong to this day, as “a number of Catholic, Protestant and other religious organizations still run significant numbers of primary and secondary schools in the society” (Bolton 2012: 225).

<sup>21</sup> Some of these reforms included compulsory schooling on primary and secondary levels which gave most pupils the opportunity to learn English at secondary schools (Bolton 2002: 34).

schools, while most chose to complete their secondary studies in English. Indeed, as Bolton (2002: 34) notes, these policies “have contributed more than any other factor to the spread of a knowledge of English within modern Hong Kong society” and over the following decades, they changed the linguistic landscape of Hong Kong from “elitist bilingualism” to “mass bilingualism”. Other major changes during this time included the declaration of Chinese as a co-official language of Hong Kong in 1974.<sup>22</sup>

Schneider (2007: 133, 135) argues that HKE entered the nativization phase already in the 1960s, with some traces of phase two still being detectable in the variety. Evans (2014: 591) however, contests this by suggesting that nativization did not take place until the 1980s and 1990s and the rise of “common bilingualism” (see also Bolton 2002: 34, above). Though there are some discrepancies between the two suggested timelines for the evolution of HKE<sup>23</sup>, both Schneider (2007) and Evans (2014) agree that the variety is now in phase three.

After the handover in 1997, Hong Kong’s language policies concerning the medium of education first changed to favour Chinese-medium instruction even in secondary schools, whereas 2010 saw the implementation of a more lax policy that gave schools more freedom to choose between English- and Chinese-medium instruction. The pressure for these reforms came, according to Bolton (2012: 232), from the local parents, businesses wishing to maintain Hong Kong’s role as one of the centres of international trade, and the “pragmatic need to prepare students for a university education”. Hung (2012: 114) notes that today primary schools in Hong Kong, which “universally adopt Chinese as the medium of instruction”, teach English as a subject from grade one, whereas “English is the principal medium of instruction in some secondary schools and all universities”. As in other British colonies, the English taught in Hong Kong has traditionally been oriented towards BrE models, but as the majority of Hong Kongers speak Cantonese as their L1, many elements from the substrate ranging from phonological features (Hung 2000) to morphosyntax (Gisborne 2009) have also influenced the local variety of English. Here it should be noted that the strong presence of Cantonese in the IDG strand has also made some question the status of HKE as a variety of English (see

---

<sup>22</sup> According to Evans (2014: 589), another change taking place during the 1970s was the increased sense of a distinct Hong Kong identity, which did not manifest itself in the population until the children of the new Chinese refugees reached maturity; this change can be detected in the growing use of the terms ‘Honkonger’ and ‘Hong Kong people’ after the 1960s.

<sup>23</sup> Unfortunately, Evans (2014) does not clearly state the timeframe when HKE was in the endonormative phase after the relaunch of the foundation phase in the 50s and 60s.

discussion and references in Evans 2011: 294–6 and Bolton 2003: 41–7). These views are based on the argument that for a localized form of English to develop into a nativized variety, it must function as a lingua franca in the community – a prerequisite that HKE has been argued not to meet (Evans 2011: 294; for a critique of the ‘monolingual myth’, see Bolton 2002: 41–3). However, others (e.g. Bolton 2003: 47; Schneider 2007: 135–9; Evans 2011) have argued that HKE fills many of the central criteria<sup>24</sup> set for New Englishes and thus should be considered to be a distinct variety of its own. This view is also taken by the author of the present study.

## 2.2.2 India

### 2.2.2.1 History of India

India is located in South Asia, covering over 4.4 million km<sup>2</sup> of land on the Indian peninsula. First traces of human civilization in India can be traced back to 2600 BCE, when speakers of Dravidian languages made their first permanent settlements in the basin of the Indus River (McNeill & McNeill 2003: 62). The cultural and linguistic landscape of the country changed dramatically around 1500 BCE, when the Indo-Aryan tribes from Middle Asia entered the Indian subcontinent from the north and pushed much of the Dravidian-speaking population further south creating a division that has existed in the Indian peninsula to this day. Other lasting legacies brought by the Indo-Aryans include the caste system, a religion which would over the millennia develop into modern-day Hinduism, and the Sanskrit language, which has both remained relatively unchanged in the holy scriptures of Hinduism and evolved into modern-day Hindi and Urdu.<sup>25</sup>

The first Islamic influences on the Indian subcontinent can be dated back to the 7<sup>th</sup> century and the Arab conquest of the area of modern-day Pakistan, but the era of the first major Islamic state in India, the sultanate of Delhi, was not established until the 13<sup>th</sup> century (McLeod 2002: 33–5). In the 16<sup>th</sup> century, the sultanate was followed by the Mughal Empire, which ruled vast areas of modern-day India for

---

<sup>24</sup> The criteria include having a recognisable pronunciation and having words or phrases that are peculiar to HKE. For full list of criteria, see Butler (cited in Bolton 2003: 46–7, 206–18).

<sup>25</sup> The two languages originated from Sanskrit, which due to various socio-political reasons began to be written with two different systems (Hindi is written in the Devanagari script while Urdu is written in the Arabic script) (Kachru 2006: 2; see also King 2001).



over 300 years (Kulke & Rothermund 1986: 197); during these centuries, Persian replaced Sanskrit as the language of administration, court as well as literary creativity and India flourished financially, culturally and politically.

The British East India Company (EIC) first arrived in India at the beginning of the 17<sup>th</sup> century and they soon established several trading stations along the Indian coastline. During this time, they also became increasingly involved in the local power politics, and the situation began to escalate by the 1750s as the Mughal Empire weakened and numerous rivalling states across the Indian subcontinent began competing for power (McNeill & McNeill 2003: 240). By this time, the EIC had numerous trading stations across India and it began defeating the new emerging states one by one, until the company had subjugated the entire Indian subcontinent (Kulke & Rothermund 1986: 269).<sup>26</sup> Despite this, the enterprise struggled to remain profitable and before the end of the 18<sup>th</sup> century, the company was finally taken over by the British government (McLeod 2002: 69–70). To secure their trade, the British began expanding their territorial control in India and by the mid-19<sup>th</sup> century, the British controlled (directly and indirectly) an area that included not only the present-day India, but also Pakistan and Bangladesh (McLeod 2002: 79).

The relations between the British and Indians had always been precarious, but it was only after the 19<sup>th</sup> century when a series of uprisings<sup>27</sup> broke out that demands for independence became more widely expressed. The next major unrest of the same scale did not emerge until almost a century later, in 1942–3, when Mahatma Gandhi's 'Quit India' campaign spread across the country (McLeod 2002: 123). Finally, the pressure to end the British Raj became too great and preparations for the region's independence were drafted. By this time, the 'divide and rule' tactic used by the British had created new rifts between the Hindus and Muslims in many parts of the country, and the breach grew larger as the moment of independence drew closer (Reeves & Rai 2006: 26–8) – a development that finally resulted in the partition of India. On the 14<sup>th</sup> of August 1947, the Dominion of Pakistan (including modern-day Bangladesh) was declared independent, which was followed by India's declaration of independence on the following day, the 15<sup>th</sup> of August. Although the movement of Hindus and Muslims to their new states had begun already before the partition, the migration flows grew after independence and the outbursts of violence

---

<sup>26</sup> The victory of the British in the Battle of Plassey (1757) is often seen to mark the moment when the role of the British changed from a business enterprise into a dominant political power in South Asia (Schneider 2007: 163).

<sup>27</sup> The most famous of these was the Great Revolt of 1857.

that took place during this time left a deep trauma in the minds of the two nations (Reeves & Rai 2006: 59).

After independence, India faced many challenges, but as the century neared its close, India's economy began to recover as the global growth of the IT markets enabled India to utilize one of its greatest assets, its human capital. Reeves and Rai (2006: 31) credit this trend to the "availability of an educated middle-class workforce fluent in English [which] provided a basis for the development of 'outsourcing' of 'back-office' operations". During the past 30 years, India has successfully established itself as a "part of the coming Asian powerhouse of world economic development", as Reeves and Rai (2006: 31) note and thus, its importance and impact is not likely to diminish in Asia or beyond in the near future.

#### 2.2.2.2 Linguistic ecology of India

India is one of the most multilingual nations in the world: there are 121 different languages with over 10,000 speakers according to the 2011 Census (Registrar General and Census Commissioner 2011), in addition to the hundreds of smaller languages spoken in the country. The majority of these languages (e.g. Hindi, Bengali and Marathi) belong to the Indo-Aryan branch, which dominates the linguistic landscape of Northern India. The second largest group of languages belongs to the Dravidian language family (e.g. Telugu, Tamil and Kannada), which in turn is strongly represented in the South.<sup>28</sup> Furthermore, some Austro-Asiatic, Sino-Tibetan, Austronesian, Kam-Thai and Andamanese languages are also spoken in the country, though their speakers form only a fraction of the population (Asher 2008). After independence, Hindi was declared as the official language of the country, while English was retained as an 'associate language'. In addition to Hindi, there are 21 other major local languages, which the constitution lists as 'scheduled languages'. According to Graddol (2010: 51), the status of these languages is mainly rhetorical, as states are free to select the official language of their regional governments. In fact, as Reeves and Rai (2006: 30) point out, many of the state borders in India were redrawn following linguistic lines in 1956, and therefore, as suggested by Graddol (2010: 51), the choice for the official language was uncontroversial for most states. However, there are some states, especially in northeast India, where the linguistic

---

<sup>28</sup> Some pockets of Dravidian languages can still also be found in parts of northern India (Graddol 2010: 50).

situation is more complex, and which have chosen English or some other regional lingua franca as the language of local government (Graddol 2010: 51).

Although India is one of the biggest English-speaking countries in the world today (Crystal 2003b: 109), the country's relationship with English has always been complicated. The beginning of phase one of the dynamic model (Schneider 2007) in the history of IndE can be traced back to the first permanent contacts made by the British EIC in the beginning of the 17<sup>th</sup> century (Bhatt 2000: 71). During this time, Indians were exposed to the English language mostly through the trading stations of the EIC<sup>29</sup> and the missionary schools, which were the first to teach the locals English (Schneider 2007: 162–3). A third channel that facilitated the spread of English was the army of the EIC, which had begun hiring Indians, most of whom were Sikhs (Mukherjee 2007: 164). Despite these various channels of contact, Sedlatschek (2009: 9) suggests that the impact of English in India remained limited during the first 150 years after the arrival of the British.

IndE's transition to phase two began soon after 1757, when the EIC won the Battle of Plassey; as Schneider (2007: 163) notes, the victory not only changed the role of the company from a trade enterprise to the most powerful political player in the country, but it also changed the status of India from a trade colony to an exploitation colony. As English was now the language of the most powerful group in India, it was soon viewed as the language of prestige and social mobility and hence, the following decades saw a noticeable increase in the spread of missionary schools teaching English (in metropolitan areas) and in the number of Indians speaking English as an L2 (Schneider 2007: 163–4; Mukherjee 2007: 164–5). Like in many other British colonies, the ability to speak English correlated closely with social class. However, the use of the language was not solely restricted to the upper social strata of the IDG population, since, as Schneider (2007: 165) points out, there was also a growing number of Indians who worked in different income-level jobs or the colonial government<sup>30</sup> and who had a functional knowledge of English (see also Sedlatschek 2009: 12).

The exonormative phase also witnessed the emergence of a small group of influential Indian intellectuals who supported the use of English over local languages

---

<sup>29</sup> According to Methora (cited in Sedlatschek 2009: 9), the first Indians to speak English were likely those who worked as translators for the EIC.

<sup>30</sup> Davydova (2012: 369–70) argues that the most well-known pidginised varieties of IndE are Butler English and Babu English, the former being frequently used by kitchen servants and the latter by office workers.

in all fields of intellectual inquiry (Bhatt 2000: 71). This view was also gathering support in Britain<sup>31</sup>, though their rationale was typically colonial in spirit, as demonstrated in the famous Macaulay's minute from 1835:

We must at present do our best to form a class who may be interpreters between us and the millions whom we govern, --a class of persons Indian in blood and colour, but English in tastes, in opinions, in morals and in intellect. To that class we may leave it to refine the vernacular dialects of the country....

The ideas presented in the minute were soon implemented, which initiated a new phase of “systematic and widespread bilingual education in India” (Schneider 2007: 164). Indeed, already towards the end of the century, some 60% of elementary schools in India used English as the medium of instruction, while it was the only language used in both secondary schools and the first three universities which had been founded in Mumbai, Kolkata and Chennai in 1857 (Schneider 2007: 164; Mukherjee 2007: 165). Furthermore, it should be noted that already during this time, the majority of teachers providing linguistic models for Indian pupils and students were also Indians (Bhatt 2000: 71; Schneider 2007: 167).

Interestingly, Mukherjee (2007: 165) argues that IndE entered the nativization phase already in the 19<sup>th</sup> century after Macaulay's minute, whereas Schneider (2007: 165) suggests that the variety did not reach phase three until the beginning of the 20<sup>th</sup> century. The rationale for Mukherjee's (2007: 165) differing timeline lies in the fact that the decision to introduce bilingual education eventually gave rise to a new group of educated English-speaking Indians, which in turn contributed to the increasing nativization of the English used by the IDG strand.<sup>32</sup> The STL strand in turn now found itself in a country which was increasingly English-speaking, and thus less alien, drawing the STL and IDG strands closer to one another (Mukherjee 2007: 166). This did not, however, mean that the relations between the British and Indians became less constrained; the educated English-speaking Indians had now also been exposed to western ideas of democracy and the nation-state, which, as Mukherjee (2007: 166) suggests, enabled them to use “the English language as a pan-Indian

---

<sup>31</sup> There were two competing groups in Britain, the ‘Orientalists’ who supported the use of local languages as a medium of instruction in India and the ‘Anglicists’ who supported the use of English (Sedlatschek 2009: 12–3).

<sup>32</sup> This included all aspects of the variety, ranging from lexicon (Schneider 2007: 165) to phraseology, grammar and phonology (Mukherjee 2007: 167). In addition, the variety became increasingly viewed as a suitable medium for personal communication and creative writing among educated Indians (Sedlatschek 2009: 16).

communicative device in multilingual India that made it possible for Indian intellectuals from all over the subcontinent to agitate jointly against British rule and, thus, form an all-Indian political identity”.

When India became independent in 1947, English had permeated the Indian society to such an extent that even though still viewed as the language of the colonisers by many, it was declared an ‘associate language’ in the new constitution which came into effect in 1950. The original plan was to phase the use of English out of official contexts over the following 15 years, while Hindi, the official national language, would be modernized so that it could take on the various pan-Indian functions previously reserved for English (Sedlatschek 2009: 18; Gargesh 2006: 94). Although Hindi is the largest L1 spoken in the country<sup>33</sup>, it is mostly spoken in the northern parts of the country. Therefore, as Gargesh (2006: 94) points out, Southern Indians, most of whom are native speakers of Dravidian languages, became worried over the possible disadvantageous position they would be put vis-à-vis the native speakers from the North if Hindi were declared the only official language of the country. As the 15-year transition period for English was drawing to a close, these concerns escalated and a series of language riots broke out in the South, which led to the decision to retain English as an official associate language (Gargesh 2006: 94). This, according to Mukherjee (2007: 168), is also around the time when IndE entered the endonormative phase, although, as he notes, it is difficult to establish a precise year when this happened. Furthermore, the decades following India’s independence were marked by the virtual disappearance of the STL strand (Schneider 2007: 167).

A further compromise introduced during the language conflicts was the Three Language Formula, which established that all pupils must learn one regional indigenous language, Hindi (if not L1 speaker) or a south Indian language (if L1 speaker of Hindi) and English (Sedlatschek 2009: 20); it should be added, however, that since India is home to hundreds of languages<sup>34</sup>, many might speak an additional indigenous language at home which is not taught at school because it is not the official (indigenous) language of their home state. Although not compulsory as the medium of education anymore, English-medium schools are popular in India, but these are private schools as Davydova (2012: 370) points out, further contributing to the unequal distribution of English skills in the Indian society. According to

---

<sup>33</sup> Graddol (2010: 51) notes that in 2001, 41% of Indians spoke Hindi as their L1, leaving the second (Bengali 8%) and third (Telugu 7%) largest language groups far behind.

<sup>34</sup> Graddol (2010: 50) mentions that the “2001 census recorded 6,661 ‘mother tongues’ – but many of these are simply different names for, or dialects of the same language” (see also Lange 2012: 52-3).

Sedlatschek (2009: 23–4), English is today used “as an intranational link language by middle- and upper-class Indians in public domains such as education, administration, business and the media, but also in more private domains among family and neighbors” and thus, it does not function as an identity carrier for Indians (see, for example, Schneider 2007: 167; Mukherjee 2007: 174). This does not, however, mean that speakers of IndE (especially younger, more educated speakers) would not view their own variety in an increasingly positive light (see also Leitner 1992: 208; Fuchs 2015, 2016) – a further sign that IndE has indeed reached the endonormative phase in the Dynamic Model.

## 2.2.3 The Philippines

### 2.2.3.1 History of the Philippines

The Republic of the Philippines is an island nation spread over 7,600 islands (some 300,000 km<sup>2</sup>) located between the South China Sea and the Philippine Sea in Southeast Asia. The first people to inhabit the Philippine archipelago some 30,000 years ago were the Negrito people, who were followed by the proto-Malay and Malay people from various parts of Asia and Polynesia over the following centuries (Dolan 2003: 27). Unlike many other parts of Southeast Asia, the Philippine islands remained relatively isolated until the colonial era, which SarDesai (1997: 71) notes to be the result of two factors: firstly, the waters surrounding the Philippine archipelago are characterised by deep-sea valleys and yearly typhoons, which restricted the number of traders visiting the islands, and secondly, the people in the islands did not sell any products, such as spices, which would have made the dangerous sea journey profitable for the merchants from the continent. Despite this, the Filipinos were not completely cut off from the trade connections of Southeast Asia and the influences thereof. Islam, for example, was introduced to the islands by Indonesian traders and missionaries, and by the beginning of the 16<sup>th</sup> century, the religion had spread from the most southern islands to the area of modern-day Manila (Dolan 2003: 37).

The first contacts between the Filipinos and the Spanish can be dated back to 1521, but the first permanent Spanish settlement was not established until 1565 (SarDesai 1997: 70). According to Dolan (2003: 38), Spain had three major objectives for its only colony in Asia: to obtain a share of the Asian spice trade, to develop relations with Japan and China, and to convert the local Filipinos to Christianity. Although the first two objectives were never achieved, as SarDesai (1997: 63) notes,

“the degree of religious and cultural penetration by the Europeans in the Philippines was greater than anywhere else in Asia”.<sup>35</sup> This did not, however, restrain the locals from organising a series of revolts over the years, all of which were successively put down by the Spanish colonisers (SarDesai 1997: 150).

The era of Spanish colonial rule did not come to an end until 1898, when Spain, having lost the Spanish-American war<sup>36</sup>, agreed to sell the Philippines to the US for 20 million dollars. The Filipinos, who had considered “their relationship with the United States ... [as] that of two nations joined in a common struggle against Spain” (Dolan 2003: 48), did not accept this turn of events and, as a consequence, the rebellions continued (SarDesai 1997: 157–8). In 1934, first steps towards independence were taken with the establishment of the Commonwealth of Philippines; both parties agreed on a transition period of 10 years, during which the Commonwealth would create a stable and independent structure of governance, while the US retained its administrative authority over foreign policy issues (Dolan 2003: 59). The plans were temporarily put on hold in 1942 when the Philippines fell under Japanese occupation during WWII, but the situation came to an end in 1945, when Japan officially surrendered, and the Philippines gained its independence from the US the following year.

The post-independence years of the Philippines have been tumultuous and characterised by fights against corruption and problems related to the country’s large foreign debt, which has encumbered the nation’s development significantly (Bolton & Bautista 2004: 2). As Bolton and Bautista (2004: 3) note, the most important export of the Philippines today is its English speaking people, many of whom work abroad as domestic helpers, nurses and seamen – the money they send back to their families makes a significant contribution to the country’s struggling economy (see also Martin 2012: 195); the consequences that this has had on the country’s language policies will be discussed in the following section in more detail.

---

<sup>35</sup> To this day, Catholicism has retained its status as the most widely practiced religion in the country (Philippines in Figures 2016: 28).

<sup>36</sup> Miller (1984: 14) argues that the US’s interest in the war was due to its desire to obtain “strategically located colonies to provide communication links and logistic support for commercial and missionary enterprises”. However, in the course of events, the US decided to take over the entire Philippine archipelago.

### 2.2.3.2 Linguistic ecology of the Philippines

The Philippines has two official languages, English and the Tagalog-based Filipino, although in reality, the language situation is much more complex since, depending on the source (see, for example Lewis et al. 2015; McFarland 2004: 60, 64), between 116 to 182 local languages of the Austronesian language family are currently spoken in the country. The languages with the largest number of speakers are Tagalog (24%), Bisaya/Binisaya (11%), Cebuano (10%), Ilocano (9%), Hiligaynon/Ilonggo (8%), Bikol (7%) and Waray (4%) (Philippine Statistics Authority 2014: 27).<sup>37</sup> Although the remaining languages have smaller speaker populations, many are small only in relative terms, as some of these languages such as Pangasinan, Pampangan and Maguindanao still have over a million speakers (Lewis et al. 2015). Due to the multilingual nature of the nation, many Filipinos are minimally trilingual according to Dayag (2012: 92), who notes that people frequently use their mother tongue at home, while Filipino and English are used in schools and at work.

Even though Spanish was the official language of the Philippines for over 300 years, the language was never widely used among the population (Beebe & Beebe 1981: 322). The reason for this, according to SarDesai (1997: 72), was that most of the Spanish priests and friars, who had full control over the education offered in the country, were opposed to the idea of teaching the Filipinos any Spanish, because they were worried that access to European literature might give rise to civil unrest. When the Philippines was sold to the US, the Americans quickly took measures to supersede the Spanish language with their own; the motivation for this was, as Beebe and Beebe (1981: 322) note, “to educate the Filipinos and uplift and civilize and Christianize them to fit the people for the duties of citizenship”, thus echoing the colonial spirit of the times. Interestingly, although the already low number of Filipinos speaking Spanish dropped even lower after the introduction of English<sup>38</sup>, Spanish was retained as one of the official languages until 1987 (Sibayan & Gonzalez 1996: 139–140).

The beginning of phase 1 in the history of PhiE can be traced back to the year 1901 and the arrival of the first American English teachers (the ‘Thomasites’). Although the number of English teachers was initially low, their figures grew rapidly,

---

<sup>37</sup> These also constitute the major ethnic groups in the Philippines according to the Philippine Statistics Authority (2014: 27), although racially, Filipinos are Austronesian and thus are related to the populations of Malaysia and Indonesia (Bolton & Bautista 2004: 2).

<sup>38</sup> According to the 1980 census, only 3 per cent of Filipinos could speak English (Sibayan & Gonzales 1996: 139)



since some of the more advanced students were allowed to begin teaching English to the lower classes while they continued with their own studies (Gonzalez 2004: 8). Therefore, though PhiE is a relatively young variety, it has been taught to Filipinos by other Filipinos for most of its history, as noted by Gonzalez (2004: 10) whereas the STL strand always remained marginal in size (Schneider 2007: 140). The Filipinos were eager to learn English as they considered it to be a “socioeconomic ‘equalizer’” (Sibayan & Gonzalez 1996: 140)<sup>39</sup> and thus, the language spread through the population at a pace unparalleled in colonial history (Gonzalez, cited in Bolton & Bautista 2008: 4).<sup>40</sup> This, according to Schneider (2007: 140), led to a situation where the first two phases of the development of PhiE “have practically merged”. In 1939, Tagalog, the language spoken widely in the Northern parts of the country was declared ‘Wikang Pambansa’, the national language, and during the Japanese occupation, it was elevated to an official language in 1942 (Bolton & Bautista 2004: 3; Sibayan & Gonzales 1996: 139). This did not, however, have an influence on the prominent status English had in the Philippines (Martin 2012: 191) and, according to Schneider (2007: 141), the language even became a symbol of resistance during the years of Japanese occupation.

Schneider (2003: 262) dates the beginning of the nativization phase to the years following independence (1946) and to the implementation of a bilingual teaching scheme, which introduced Tagalog on a national scale as a subject, whereas English remained the language of instruction (especially in the sciences, see also Martin 2012: 191). The original plan was that Tagalog would eventually replace English in such institutions as schools, the government and business so that the language would become the sole national language of the country (Dolan 2003: 82). To strengthen the image of Tagalog as a national language, the name of the language was changed to ‘Pilipino’ in 1959, which in turn was changed to ‘Filipino’ in 1987 (Martin 2012: 191). Despite this, many people remained concerned over the position of their own regional languages, while others considered English vital for maintaining connections with the rest of the world (Dolan 2003: 83) and as a consequence, the use of English spread even further during the decades following independence.

---

<sup>39</sup> Schneider (2007: 140) argues that although the Filipinos craved for education, the number of students dropping out of school were high. Furthermore, when the upper social strata of the Philippine society abandoned Spanish and began learning English, the language eventually changed from a “socioeconomic equalizer” to a “socioeconomic stratifier”, as noted by Sibayan and Gonzalez (1996: 140; see also Bailey 1991: 88).

<sup>40</sup> The phenomenal spread of the language is also evident from the statistics: in 1918, only 8.7% of the Filipinos could speak English, whereas by 1980, the figures had grown to 64.5% (Gonzales 2004: 9).

According to Schneider (2007: 141), this development reached its culmination during the 1970s, when English was “deeply entrenched in the national self-projection, continued to move into home environments, and developed audibly indigenous forms”.<sup>41</sup>

Towards the end of the decade, differing opinions began to be voiced as new generations of students began to view English as the language of the colonisers that was hindering the development of both the national language and the nation (Thompson 2003: 35; Gonzales 2004: 12). In 1973, the new constitution maintained that English and Pilipino be retained as the official languages of the country, thus reaffirming both the status of English and the intention of developing Pilipino into a national language (Thompson 2003: 37). Interestingly, Peters (2009: 107) argues that PhiE is temporarily fossilized, a view supported also by Schneider (2003: 263), who notes that even though some preliminary signs of the endonormative phase can be detected in the variety, “the Philippines appears to be an example of a country where the predictive implications of the dynamic model may fail”.

The language situation in the Philippines is indeed a complicated one. Gonzales (2004: 11) argues that “the Filipino speaks his or her vernacular at home as the language of intimacy, Filipino in the neighbourhood if it is in the Manila area, urban center, or in the Tagalog region, the vernacular of the community if non-Tagalog-speaking, English in school and in the office, English for travel”. Furthermore, a study commissioned by the Linguistic Society of the Philippines in 1993 shows that proficiency in English correlated positively with three factors: higher socioeconomic class, young age and urban areas (Thompson 2003: 72–3) and based on these findings, Thompson (2003: 74) argues that the Philippines can still be regarded as an “ESL country, at least in the urban setting.”

Although the role of English is still contested by many Filipinos, the language’s economic value to the nation is indisputable. Bernardo (2004: 29) argues that “English now has a much more circumscribed role as a language of access, and the primary role of Filipino and other local languages in the more basic educative functions is being underscored”. Indeed, as Martin (2012: 192) points out, overseas Filipino workers are considered “the single most significant force in promoting a high demand for English in Philippine society”, and the remittances sent by Filipinos working abroad form a sizeable portion of the (unofficial) national income (Gonzales 2004: 14–5). Furthermore, the government is trying to attract investments

---

<sup>41</sup> It should be noted, however, that traits from AmE can be detected in the variety even today (Bailey 1991: 88; Hassan et al. 2012: 329).

from the call centre industry, for which English-speaking labour force is needed (Martin 2012: 193). Unfortunately, as Martin (2012: 194) points out, the *kind* of English spoken by the Filipinos frequently does not match the needs of the industry, as the English used by many is actually closer to Taglish than ‘standard’, international English.

## 2.2.4 Singapore

### 2.2.4.1 History of Singapore

Singapore is a small city-state located at the southern tip of the Malay Peninsula, covering only some 700 km<sup>2</sup> of land distributed over Singapore Island (Pulau Ujong) and some 60 smaller islands in its vicinity. The first historical records of human habitation on the Singapore Island come from 14<sup>th</sup> century scriptures that refer to a settlement called Temasek, which Turnbull (2009: 19) argues was a small but prosperous port.<sup>42</sup> At the end of the 14<sup>th</sup> century, the name of the settlement changed to Singapura (‘Lion city’), but the origin of the name has been lost in history (Turnbull 2009: 21–2). Over the next centuries, Singapore was part of many larger Southeast Asian kingdoms, such as the great maritime power Srivijaya, the Majapahit Empire and the Sultanate of Malacca (Sivonen 2006: 244–5; SarDesai 1997: 44–5), but by the beginning of the 19<sup>th</sup> century, Singapore Island was only inhabited by some 500–1000 Malays and a few dozen Chinese, most of whom were fishers (SarDesai 1997: 93; Turnbull 2009: 25), while Lim and Foley (2004: 2) also mention the presence of some Indonesians, Indians and possibly some Eurasians.

In 1819, Sir Thomas Stamford Raffles of the EIC leased Singapore from the sultan of Johore and founded a trading station on the island (SarDesai 1997: 92). Because Britain’s interests in the region were commercial, Singapore was declared a free harbour, and its favourable location between Britain’s India-China trade, its free trade policy and the (relatively) well-organized society that developed in the colony all contributed to the exponential growth of its commercial importance and population (Turnbull 2009: 5; see also Newbold, cited in Bao 2015: 16). The

---

<sup>42</sup> It should be noted that the island was not uninhabited prior to the 14<sup>th</sup> century, but, as Turnbull (2009: 20) notes, it was only after the changes that took place in the regional trade during this time that gave rise to the emergence of small ports in the Malay archipelago, including the one on Singapore Island.

composition of the population became fixed early, Chinese forming the biggest ethnic group which was followed by Malays and Indians in descending order (Turnbull 1977: 36–7; Wells 1982: 645), whereas the number of Westerners always remained low, mostly consisting of British (Sivonen 2006: 104).<sup>43</sup>

After the Anglo-Dutch treaty of 1824, the EIC founded the Straits Settlements, which consisted of Singapore, Penang and Malacca, but in 1830, the Straits was placed under the authority of British India.<sup>44</sup> Therefore, as Rai (2006: 177) notes, “Singapore, for nearly half a century, was ... effectively a part of India” and this can be seen, for example, in the temporary spike in the number of Indians living in the colony (see, for example, figures presented in Bao 2015: 17). In 1867, the Straits Settlement was declared a separate Crown colony and Singapore’s population and economy continued to thrive. When the colony was approaching its 100<sup>th</sup> year at the early decades of the 20<sup>th</sup> century, the Chinese still formed the largest ethnic group (over 70%); they were followed by Malays (under 14%) and Indians (approximately 9%) (Bao 2015: 17).

During the Second World War, Singapore fell under Japanese occupation in 1942 and although the Singaporeans welcomed the return of the British in 1945, something had irreversibly changed during the war: as Turnbull (1977: 218) notes, “the only ultimate justification for a colonial power was its ability to protect and in this the British colonial regime had been tried and found wanting.” In 1943, while the Japanese occupation was still ongoing, Britain promised that preparations towards establishing Singapore’s self-governance were to be made as soon as the war was over, and one of the first steps the British took after the war was to declare Singapore a separate Crown colony, thus giving it independence from the Malaccan peninsula (Sivonen 2006: 183–4). During the years that followed, Singapore gained more independence over its internal affairs and in 1958, Singapore’s status was elevated from a colony to a state (Turnbull 1977: 267). Despite this, Singapore’s ties with Malaysia were still strong, and in 1963, Singapore (together with Sarawak and Sabah) joined the Federation of Malaysia. However, this arrangement was short-lived, as Malayan nationalism, which frequently targeted the Chinese, was on the rise in Malaysia, causing conflicts with Singapore’s Chinese-majority population (Sivonen 2006: 186). Major conflicts over the rights of the Chinese citizens soon escalated and

---

<sup>43</sup> Despite their low numbers, this group held most of the key seats in the local government and they were also an important source of capital in the trade colony (Sivonen 2006: 104).

<sup>44</sup> The motivation for this, according to Rai (2006: 177), was purely financial, as the EIC was trying to cut down on its costs.

as a consequence, Singapore was expelled from Malaysia and thus it became independent in 1965 (Dixon 2005: 27).

During its early years, the Republic of Singapore faced multiple challenges ranging from securing its economic prosperity to creating a feeling of national consciousness among its multi-ethnic and multilingual population (Turnbull 1977, xiv). Since Singaporeans did not have any common cultural legacy on which to build their identity, the governing bodies decided that the nation's cohesion was to be based on a "multi-racial, multi-lingual secular society" (Turnbull 1977: 300). Furthermore, because Singapore did not have any natural resources it could exploit, it had to rely on the only resource at its disposal, its people (Deterding 2007: 2). This strategy proved to be successful, and today Singapore has established itself as one of the leading centres of trade and business in Asia and it is one of the wealthiest countries in the world.

#### 2.2.4.2 Linguistic ecology of Singapore

Singapore has four official languages: Mandarin, Malay, Tamil and English, the first three being assigned as the official mother tongues of Singaporeans based on their ethnicity, while English functions as an intranational linking language between the groups. The ethnic profile of Singapore has remained fairly stable for the past 100 years (Bao 2015: 16) and today, the population of Singapore is approximately 5,6 million, consisting of three major ethnic groups, Chinese (74.3%), Malay (13.4%) and Indian (9%), while the remaining 3.2% is formed by 'other' (Singstat 2017). Although each ethnic group is assigned to an "official mother tongue", it should be noted here that in reality the language situation in Singapore is more complex, and many do not actually speak their official mother tongue as an L1.

Phase one for SinE in the Dynamic Model began in 1819, when the EIC established its trading station on the island (Schneider 2007: 153). As with most colonies of this type, the proportion of the STL strand remained marginal throughout Singapore's history.<sup>45</sup> Although the 'true' IDG strand in Singapore was formed by the local Malays (who spoke Malay) and more Malays moved to the region after the foundation to the colony (Lim & Foley 2004: 2), they were quickly

---

<sup>45</sup> It could be argued that the same critique presented by Evans (2008) regarding the applicability of the terms STL and IDG strands in the case of Hong Kong could also be applied here, since Singapore too was a colony where for the first 100 years or so people mostly came for trade without the intention of settling down.

superseded by the influx of Chinese who mostly spoke Cantonese, Teochew, Hokkien and Hakka (Turnbull 1977: 36). The third largest ethnic group in Singapore was formed by Indians, many of whom were Tamils and Malayalees, but also Sikhs, Punjabis, Sindhs, Gujaratis, Bengalis and some Parsis came from Northern India (Sivonen 2006: 120) and therefore languages from both Dravidian and Indo-European language families were spoken within the group.<sup>46</sup> When the British arrived, the local community was already divided based on ethnic lines (the ‘capitan’ system), which the British decided to maintain (Lim & Foley 2004: 2). Bao (2015: 21) argues that this division was based on both ethnic (Malays, Indians) and dialectal (Chinese) affiliations and, as Lim and Foley (2004: 3) note, the groups used Bazaar Malay, a lingua franca spoken widely in the Malay archipelago, to interact with one another.<sup>47</sup> As in many other colonies, the British were mostly interested in teaching English to the local elite, who could then function as intermediaries between them and the remainder of the local population, which is why English-medium schools were founded already in 1816 (Bao 2015: 25; Lim & Foley 2004: 3), whereas education in the native languages was not offered until 1834 (Lim & Foley 2004: 3).

SinE reached phase two rapidly, Schneider (2007: 154) dating the transition around the year 1867 when Singapore was declared a crown colony. Bao (2015: 21–2) notes that there are indications that a pidginized form of English was emerging as a local lingua franca already during the latter half of the 19<sup>th</sup> century, though it did not yet replace Bazaar Malay as an interethnic linking language. According to Schneider (2007: 154), bilingualism was also becoming more common during this time, but it still had “an elitist touch”. The exonormative phase also saw English becoming established as a key element for socio-economic mobility, and even though Chinese-medium schools still had the highest rates of enrolment (Bao 2015: 27)<sup>48</sup>, the number of Chinese children enrolling to English-medium schools rose significantly as the 19<sup>th</sup> century drew to a close (Lim & Foley 2004: 3–4). This, as Schneider (2007: 154) suggests, could also have set off the development that

---

<sup>46</sup> For a more detailed description of the substrates spoken in Singapore, see Lim (2007).

<sup>47</sup> Lim (2007: 452) notes that also Hokkien was used as a lingua franca at the time, but only among the Chinese whom spoke numerous Southern Chinese dialects.

<sup>48</sup> According to Bao (2015: 26, 31), Mandarin was taken up as the medium of education after the Chinese revolution of 1911, thus enabling the schools to offer education for the whole Chinese community – prior to this, people speaking different Chinese dialects had attended different schools. Despite this, Mandarin did not displace the other Chinese dialects, which were still used at home (Bao 2015: 51).

eventually led to the emergence of colloquial form of SinE called Singlish.<sup>49</sup> A further aspect that likely contributed to this development was the fact that many Asian migrants were now settling down in Singapore and began starting families (Bao 2015: 17), thus giving rise to a growing number of locally born Singaporeans.

The development of SinE during the 20<sup>th</sup> century was rapid; Schneider (2007: 155) dates the beginning of the nativization phase to the year 1945 when Britain regained control of Singapore after the Japanese occupation. At the time, the ability to read and write in English was slightly higher among the Malays and Indians when compared with the Chinese, who had traditionally favoured Chinese-medium schools (Bao 2015: 27), but in the 1950s, the percentages of students enrolling to English-medium schools rose to over 40%, thus overtaking those of Chinese-medium schools (Lim & Foley 2004: 4). The decade also marked the introduction of a new language planning policy, which presented many suggestions that have shaped the linguistic landscape of Singapore to this day. Firstly, the document recommended that Malay, the language of the original IDG population, be declared Singapore's national language (Lim & Foley 2004: 4). Secondly, it recommended that the country have four official languages: Mandarin, Malay, Tamil and English. The first three languages were designated as the "heritage languages" of the three major ethnic groups (Alsagoff 2012: 143), thus "[anchoring] the people to their ethnic roots and cultural traditions" (Low 2012: 36). English, in turn, was to function in the role it has often had in other postcolonial societies – as that of a neutral interethnic language that is equally foreign to all, while at the same time being a key element in Singapore's strategy to increase its competitiveness in the global markets (Lim & Foley 2004: 4). The third major recommendation of the document was that a bilingual education policy be set up where all pupils are taught two languages (English and their "assigned" ethnic L1<sup>50</sup>) in primary school and three languages (English, Malay and their ethnic L1) in secondary school (Alsagoff 2012: 141). The aim of these policies was, as Lim and Foley (2004: 4) suggest, "to create national

---

<sup>49</sup> While some (see Schneider 2007: 155 and references therein) suggest that due to the 'founder principle' (see Mufwene 2001), Malay, the local lingua franca which has also been influenced by various Chinese dialects, has had a disproportionately strong influence on the development of the local variety of English, some (see, for example, Bao 2015: 35) have suggested that the influences are more likely to come directly from Chinese dialects.

<sup>50</sup> Low (2012: 36) notes that these are not restricted to the four official languages of Singapore, as the Indians could, for example, also study Hindi or Gujarati. In addition, Low (2012: 36) claims that people could choose to study another official language so that, for example, a pupil of Malay ethnicity could also study Mandarin instead of Malay.

unity and forge a national identity and consciousness that transcended ethnic boundaries”.

SinE entered the endonormative phase in the 1970s, which Schneider (2007: 157) notes to be a time of rapid economic growth, modernization and nation building in the history of Singapore. The decade was also marked by two other changes in the linguistic ecology of Singapore: the emergence of colloquial SinE, Singlish<sup>51</sup>, and Bazaar Malay being superseded by English as the local lingua franca (Schneider 2007: 158; Bao 2015: 23).<sup>52</sup> The spread of English was also noticeable in the educational system, where the number of pupils enrolling to English-medium schools grew, leading to a decision to make English the only medium of instruction used in schools from 1987 onwards, which in turn changed the status of English to a “first language” (Alsagoff 2012: 146).<sup>53</sup> Schneider (2007: 158) argues that SinE “has gone through a vibrant process of structural nativization, more visibly on the basilectal level but also in formal styles” and describes the language situation in Singapore as diglossic. As the 20<sup>th</sup> century neared its close, Singapore’s economy soared, but the government became worried over the growing use of Singlish among its population and the negative effect it might have on the country’s economy. As a solution to the perceived problem, the Singaporean government launched the “Speak Good English Movement” in 2000, a campaign which has continued to this day despite the lack of support from the general population who view Singlish as a positive symbol of their national identity (Schneider 2007: 158; Alsagoff 2012: 148–9). Today, English is also increasingly used at home, especially by the younger generations, some 46% of 10-to-14-year-olds reporting it to be the language most frequently spoken at home, which is a noticeable increase to the 34% reported by the age group 10 years their senior (Singstat 2010; see also Lim & Foley 2004: 6; Schneider 2007: 157). Therefore, it is not surprising that signs of differentiation, the final phase in the Dynamic Model, have already been detected in SinE (Schneider 2007: 161).

---

<sup>51</sup> Schneider (2007: 158) links this development to the emergence of second-generation speakers of English in Singapore.

<sup>52</sup> According to Lim (2007: 456), the older generations of Singaporeans still use Bazaar Malay.

<sup>53</sup> Alsagoff (2012: 146) importantly notes that the term in this context simply “reflects the primacy of the English language as the medium of education” and therefore the term should not be viewed as any type of indication of it being a pupil’s mother tongue.



## 2.3 Defining epicentres

The term *epicentre* originates from the field of geology, and the Oxford English Dictionary (OED) defines the word as “[t]he point on the earth's surface that overlies the subterranean focus of an earthquake”. Hence, an epicentre forms the core from which seismic waves appear to spread to the surrounding environment. The second definition given by the OED is more figurative, but echoes the original, as the word can also refer to a “center of activity, energy or disturbance”. When the term is used in the field of World Englishes, its meaning is not far removed from the original(s): when a variety is referred to as a linguistic epicentre, it is generally considered to describe a situation where the variety has some influence on the development of other English varieties – the figurative waves of influence thus spreading from one country/community to another.

One of the first to use the term to describe the emerging pluricentricity of English varieties around the world is Leitner (1992: 225), who argues that “[e]pi-centres [among L2 varieties] have already been recognized in India, Singapore and a few other areas.” The notion that multiple centres already exist for English is also supported by Kachru (1998: 97), who argues that they

1. provide the norms and models for its acquisition;
2. develop methods and materials for appropriate localized pedagogical goals;
3. use innovations in literary creativity, genre development, and region-specific ESPs;
4. develop linguistic materials for authentication and local and regional codification;
5. recognize convergence of English with local languages (e.g. Chinese, Malay, Tamil, Hindi, Tagalog, Thai) as a natural process of convergence and acculturation; and
6. consider the formal processes of nativization as an integral part of the linguistic variety and incorporate these features in the local dictionaries and teaching materials of the variety.

While the above list includes many key elements that could indeed be expected to feature in a variety that functions as an epicentre, Leitner (1992) and Kachru (1998)

seem to accept the existence of these multiple centres without much criticism.<sup>54</sup> Hundt (2013), in contrast, warns against the lax use of the term, reminding that since the field still has not reached a consensus as to the precise meaning and applicability of the term, it should be used with caution. According to Hundt (2013: 185), only the following features have hitherto been agreed upon: for a variety to function as an epicentre, it must firstly be endonormatively stabilized and secondly, it should have the “potential to serve as a model of English for (neighbouring?) countries” – two requirements that IndE can be argued to meet (see discussion in 2.2.2 and 2.4). Hundt’s (2013) observations on the theoretical problems caused by the lack of a more precise definition will be addressed here, whereas her observations concerning the methodological challenges that come with the use of the term will be discussed in the Methods section (3.2).

The first theoretical problem according to Hundt (2013: 183) is that the term epicentre has not been defined sufficiently for its linguistic purposes. While the idea of waves emanating from a single source to its surroundings might be more applicable in some contexts such as South Asia (see Gries & Bernaisch 2016, also section 2.4), Hundt (2013: 189) notes that the earthquake metaphor excludes the type of influence that spreads from one urban centre to another which often leaves “gaps” between the two regions (see also, for example, Trudgill 1974; Milroy & Milroy 1985). Consequently, Hundt (2013: 189) suggests that the term ‘additional’ centre would be more appropriate, as the term lacks any implication of geographical proximity between the varieties. Though the appeal of the term suggested by Hundt is obvious to a study that focuses on varieties geographically distant from one another, the present study nevertheless settles on the more commonly used term ‘epicentre’ for two reasons. Firstly, after decades of use, ‘epicentre’ can be argued to be an established term in the field and secondly, because the two aspects of the term that are generally agreed upon (the endonormativity of the variety and its potential to function as a model for other varieties) are not seen to presuppose the geographical proximity of the varieties involved. Therefore, when calling IndE a potential epicentre, the present study does not assume that the variety’s influence would spread evenly from India to the surrounding areas; instead, it recognises that for this to happen, different conduits need to have been established between the speakers of the varieties, and that the quality and strength of these connections all

---

<sup>54</sup> Indeed, the fact that Kachru (1998: 97) mentions Thai among the local languages implies that he might be willing to consider one of these centres to exist in Thailand, a country where English does not even have the status of an official L2.

contribute to the (potential) influence that IndE has on other English varieties spoken in the region. Because of this, the historical, political, social, cultural and economic connections between speakers of IndE and the three Southeast Asian varieties are examined in detail in Section 4.3.3.<sup>55</sup>

The second problem Hundt (2013: 190) raises regarding epicentric influence is the fact that it can be categorized either as conscious adoption of a feature or as a case of unconscious convergence and therefore, “[s]peakers may thus adopt certain lexical or grammatical features from other varieties of English without necessarily consciously aspiring to sound like speakers of that variety”. For example, as Hundt (2013: 183) points out, in many Outer and Expanding Circle countries, materials for teaching English might come from an Inner Circle country, which can result in a situation where the educational target, especially concerning grammar and lexis, also follows that particular variety. However, considering the long tradition of Indian teachers of English working in South and Southeast Asia (see also Kachru 1985: 28), Hundt’s question could be extended to cover also the potential influence that these teachers could have (had) on establishing specific features as prestigious in their respective speaker communities. This issue will be discussed in greater detail for each variety in Section 4.3.3.

The third problem Hundt (2013: 183) raises is connected to the application of the term, which has traditionally been used when referring to standardized English varieties, but since “[n]ational varieties of English also have an influence on each other on the non-standard level”, as Hundt (2013: 183, see also 187–8) notes, the present study takes the view that the scope of the term should be extended to cover also vernacular norms. This decision is further supported by one of the models used in this study, the World System of Standard and Non-Standard Englishes (Mair 2013), which recognises the possibility of non-standard varieties influencing other English varieties.

---

<sup>55</sup> Hundt (2013: 190-1) also suggests that there is more flexibility in the scope that the term epicentre has, as it can refer to individuals (e.g. celebrities), collectives (e.g. peers) or abstract entities (e.g. nation state) – in short, anyone/thing that acts as a source of authority in situations where speakers interact with one another. These aspects will also be examined closer in the Discussion.

## 2.4 Previous research on IndE influence in Asia

As mentioned in section 2.1, the possibility of non-native varieties influencing one another has been addressed in many of the models describing English varieties and there is already a small but growing body of work that indicates that IndE could indeed function as a linguistic epicentre in parts of Asia. Hoffman et al. (2011: 259, 260) note that “[t]here is a rich body of literature indicating that IndE has developed local norms at virtually all linguistic levels” while there are also “sociolinguistic studies [which] indicate that IndE is increasingly accepted as a full-fledged variety by Indian speakers” (see also Fuchs 2015; Bernaisch & Koch 2016). This view is also supported by Mukherjee (2007), who claims that IndE has already reached the endonormative stabilization phase in Schneider’s (2003, 2007, 2014) dynamic model and therefore meets some of the central criteria for a variety to function as a linguistic epicentre. At present, there are some studies (e.g. Koch & Bernaisch 2013) which address IndE’s potential epicentre status in passing, whereas others (Hoffmann et al. 2011; Hundt et al. 2012; Gries & Bernaisch 2016) have explored this issue in greater detail. The remainder of this section will be dedicated to studies belonging to the latter group.

Hoffmann et al. (2011) examine the potential role of Indian English as an emerging epicentre through the use of light verbs constructions containing the verbs *give*, *have* and *take* in IndE, Bangladeshi English (BaE), PkE, Sri Lankan English (SLE) and BrE. Though comparable diachronic data is lacking for the five varieties, Hoffmann et al. (2011: 261) suggest that a variety’s potential development into an epicentre can be examined using large-scale synchronic data, which might reveal some of the similarities between dominant (IndE) and peripheral (e.g. SLE) Englishes. The data used in the study come from the online archives of South Asian national newspapers, which are compared with the corresponding section of the British National Corpus (BNC). While some of the findings of Hoffmann et al. (2011: 270) indicate that light-verb constructions are used most frequently in IndE, the remaining Asian varieties seem to follow BrE’s lead in using these constructions less frequently. This, according to Hoffman et al. (2011: 271), might suggest that “India is a fully institutionalised variety, whereas the other L2-varieties are still in the process of becoming institutionalised”. However, when the different patterns of use are examined closer, a more complex picture emerges. According to Hoffmann et al. (2011: 273–4), it is possible that some individual divergent patterns in the data could represent local South Asian innovations, while others might simply be (so-called) learner mistakes in the data, but since differentiating between the two is

frequently difficult, Hoffmann et al. (2011: 276–7) conclude their study by noting that though the majority of their results support the notion of IndE providing a model for other South Asian Englishes (SAEs), further studies focusing on different lexicogrammatical features, language attitudes and sociolinguistic factors are needed before anything conclusive can be said about the role of IndE in South Asia.

Another study focusing on the role of IndE in South Asia comes from Hundt et al. (2012), who examine the use of the modal *would*, the subjunctive *were* and the indicative *was* in hypothetical *if*-clauses in IndE, BaE, PkE, SLE and BrE. According to Hundt et al. (2012: 149), previous studies on the use of the subjunctive indicate that speakers of L1 varieties use the subjunctive more than speakers of L2 varieties, whereas speakers of younger L1 varieties fall somewhere between the two.<sup>56</sup> Based on this observation, Hundt et al. (2012: 149–50) hypothesize that IndE, the most established institutionalised variety in the region, could have developed norms of its own, which could also be reflected in the increased use of non-standard *would* in other SAEs. Both the composition and size of the data for Hundt et al.’s (2012) study mirror those used by Hoffmann et al. (2011) discussed above. Interestingly, although their qualitative analysis indicates that *would* indeed functions as a variant for *was* and *were* in SAEs, the quantitative analysis of the three variables does not support Hundt et al.’s (2012) hypothesis, as they cannot detect any systematic pattern of distribution in SAEs which would differentiate them from BrE. Hundt et al. (2012: 159) attribute their findings to three main factors: the exonormative influence of AmE on SAEs, the level of institutionalisation of each SAE, and the exonormative influence of the old superstrate BrE. Despite this, Hundt et al. (2012: 160–1) do in fact discover one interesting feature, the use of *on if* in contexts where BrE would use *whether*, which they argue to be a genuine “nativized syntactic pattern in IndE”. In all of the other SAEs studied, the use of *on if* had also been extended to new contexts, which causes Hundt et al. (2012: 162) to speculate that this particular lexicogrammatical innovation could have spread from IndE to other SAEs – this in turn would support the argument that IndE indeed extends some influence over the other varieties spoken in South Asia.

A third study examining the role of IndE in South Asia comes from Gries and Bernaisch (2016), who use Multifactorial Prediction and Deviation Analysis with Regression (MuPDAR) to study the patterns of preference that can be found in the

---

<sup>56</sup> Some scholars (Hundt et al. 2012: 149 and references therein) have suggested that this is due to the tendency of L2 varieties to favour patterns of lesser grammatical complexity and clearer meaning.

dative alternation<sup>57</sup> of SAEs when compared to their superstrate, BrE. The data for Asian Englishes come from the South Asian Varieties of English Corpus (SAVE), which contains data from English language newspapers from India, Pakistan, Sri Lanka, Nepal, and the Maldives, whereas data for BrE was obtained from the BNC.<sup>58</sup> Gries and Bernaisch (2016: 8–9) began by retrieving a random sample of 500 hits of the verb GIVE from each variety, of which they were able to code almost 40% for the following variables: variety, paper, transitivity of the verb and the semantic class of the patient. In addition to these, both the recipient and patient were coded for length, animacy, accessibility, and pronominality. By using the MuPDAR approach, Gries and Bernaisch (2016) are able to take into account the possible effects the different variables have on the choice made by speakers of SAEs when they differ from the predicted choice of BrE speakers. The results of the study indicate that the accessibility of the recipient and patient, in addition to the pronominality of the recipient are among the most prominent factors contributing to the South Asian patterns that differ from BrE (Gries & Bernaisch 2016: 21). In addition, Gries and Bernaisch (2016) tested which of the varieties' patterns corresponds most closely with those of the others and discovered that “the underlying norms of the IndE model for the dative alternation with GIVE were most compatible with the constructional choices of the other varieties” (p. 21) and as a consequence, they consider it “reasonable to profile IndE as the linguistic epicentre of SAEs at least for this alternation” (p. 22).<sup>59</sup>

The final study discussed in this section differs from the previous in both the methods employed and the selection of varieties examined: Hogue (2001) uses historical lexical data to investigate if lexical influences from Anglo-Indian English<sup>60</sup>

---

<sup>57</sup> The two alternating patterns are the prepositional dative (e.g. ‘Sally gave a letter to Mike’) and the double object (e.g. ‘Sally gave Mike a letter’).

<sup>58</sup> Although all of the newspapers that provide the data for Hoffmann et al.’s (2011) and Hundt et al.’s (2012) studies are also represented in the SAVE corpus, the data for SAVE comes from a wider selection of sources, as data for each variety was collected from two local English newspapers instead of just one.

<sup>59</sup> This argument is also supported by the results of Heller et al. (2017), whose examination of the genitive alternation in BrE, IndE, SLE, SinE, HKE and PhiE shows that IndE indeed appears to function as an epicentre for SAEs. In addition, a study by Sedlatschek (2009) shows that some lexical features that are used most frequently in Indian newspapers written in English also appear in the corresponding newspapers published in Pakistan, Bangladesh and/or Sri Lanka. Furthermore, Koch & Bernaisch (2013: 85) mention that IndE could have influenced acrolectal Nepalese English through education.

<sup>60</sup> Hogue (2001: 167) notes that the term Anglo-Indian refers to the type of English used by the expatriate British who were living in India; this form of English also functioned as the model for

spoken in the Indian subcontinent spread to Southeast Asia during colonial times. With the help of several dictionaries (e.g. *the Hobson-Jobson* and *the Macquarie Dictionary Corpus of Asian English*), Hogue (2001) examines the extent to which 45 words of Anglo-Indian origin (e.g. *coolie*, *bungalow*) were used in International English (IE)<sup>61</sup> and the emerging English varieties that developed into HKE, PhiE, SinE and Malaysian English (MyE). According to Hogue (2001: 167, 189), most of the studied lexis entered Anglo-Indian English from Hindi(-Urdu), Tamil or even Sanskrit, but since words from, for example, Sanskrit spread to many languages spoken in Southeast Asia already during precolonial times (see, for example, Milner 2008: 24, 38; Scott 1994: 129, 196, 213), determining whether a particular lexical item entered one of the Southeast Asian Englishes via one (or more) of the local language(s) or via Anglo-Indian English is often difficult.<sup>62</sup> Indeed, Hogue's (2001) discussion on the channels of influence in the region provides an illustrative description of the complex connections that have existed between the people and the languages of the Indian peninsula and Southeast Asia for centuries. The highest number of Anglo-Indian loan words can, according to Hogue (2001: 190), be found from MyE (43 words), followed by SinE (37), HKE (26), PhiE (17) and IE (8). Interestingly, Hogue (2001: 193) argues that the majority of these lexical items entered the (developing) Southeast Asian varieties "as a direct result of British colonialism – either through the British themselves or the English-speaking Indians they brought to other colonies". Since many Indians left their homes to work for the colonial government in Southeast Asia, it is not surprising that most of the Anglo-Indian lexis used in MyE and SinE (and to some extent also HKE) are related to work as well as administrative and social practices as noted by Hogue (2001: 191). This also explains the smaller number of loan words found in PhiE, an American colony, which was never part of the British Empire (Hogue 2001: 193).

---

Indians learning English at the time and which they used when they moved to other British colonies in Southeast Asia. When speaking of the connections between present-day IndE and Anglo-Indian, Hogue (2001: 167) suggests that "while they are not exactly the same, there is clearly a close relationship".

<sup>61</sup> The term International English is not defined further by Hogue (2001).

<sup>62</sup> For example, the word *mandarin* which is commonly associated with China, is actually of Sanskrit origin (*mantrin*) from where it spread via Hindi (*mandri*) to Portuguese (*mandar*), which was spoken as a lingua franca in parts of South and Southeast Asia (Hogue 2001: 181, 189). From Portuguese, the word then entered into the English used by Anglo-Indians, who then spread the word to other Englishes spoken in the Asian region and, eventually, the word *mandarin* became established even in International English (Hogue 2001: 182).

Though Hogue's (2001) results clearly show that linguistic influences spread from India to Southeast Asia already during colonial times, he notes that the results are "a direct result of British colonialism" and that possible contemporary influence was excluded from the focus of the study (p. 193). Thus, the fact that the young uncodified type of English (if we are disinclined to call it a variety at such an early stage) spoken in India could already extend some of its influence to Southeast Asia should be discussed in greater detail, as it challenges the view that only endonormative varieties can influence other varieties in their vicinity. Peters (2009: 122) suggests that in cases where early-stage influence can be detected, "it is more likely to come from settler than indigenized varieties of English", which might also explain the results of Hogue's (2001) study. Even though IndE was still too "young" to function as a linguistic epicentre (see also Leitner 1992: 203), many of the speakers of Anglo-Indian English were British (STL strand in Schneider's Dynamic Model), which could have temporarily strengthened the status and influence of the developing Indian variety spoken in the region. A relevant example of this is provided by Peters (2009) who shows that words from AuE lexis were able to spread to NZE before the Australian variety had reached endonormativity. The reason for such an unusual development is, according to Peters (2009: 121), that "[l]oanwords help to cover lexical gaps in domains of common pioneering experience, and may pass between nations long before the donor variety is fully codified". Since the people moving between India and Southeast Asia were frequently employed by the colonial government, a 'common pioneering experience' seems unlikely; a more plausible explanation is the overrepresentation of workers from South Asia (be they of British or Indian origin) in certain professions, which in turn is reflected in the types of words that spread from the Anglo-Indian lexis to the emerging varieties spoken in Southeast Asia.<sup>63</sup>

The results of Hogue (2001) and Peters (2009) give rise to one further terminological question – what term should be used when referring to a situation where a variety's influence can be detected in another variety prior to the former having reached endonormativity? Peters (2009: 121–2) presents two alternatives: firstly, the early influences could be termed as *pre-epicentric influence*, or secondly, all instances of cross-varietal interaction could be categorized as *epicentric influence*. If we consider the critique presented towards the lax use of the term in the past (Hundt

---

<sup>63</sup> Also Schneider (2007: 165) has noted that the 19<sup>th</sup> century was a time of strong lexical borrowing from the Indian languages to the local variety of English and that some of these lexical items eventually spread to international English.



2013), it is more logical to categorise these instances as pre-epicentric influence. Thus, it could be said that previous studies on the role of IndE in South and Southeast Asia show that the variety was able to extend its pre-epicentric influence as far as Southeast Asia already during colonial times, whereas other studies based on contemporary data show that IndE seems to function as an epicentre for other SAEs, at least for some linguistic features. However, what the previous studies have not addressed is the current role that IndE has in Southeast Asia, which is the aim of the present study.

## 3 DATA AND METHODOLOGY

### 3.1 Data

The selection of varieties examined in this study was partially determined by the availability of comparable corpora, and for some varieties, the source of data changed as newer, more comparable corpora became available. The data used for the four articles (A1–A4) come from three different sets of corpora, the *International Corpus of English* (ICE), the *Santa Barbara Corpus of Spoken American English* (SBCSAE) and the *Freiburg-Brown Corpus of American English* (Frown). A description of each will be provided in the following sections together with observations of the benefits or any possible drawbacks that they might have.

#### 3.1.1 The International Corpus of English

The *International Corpus of English* (ICE) is a family of corpora that contains both spoken and written data from different varieties of English around the world. The seeds of the ongoing project were sown in 1988, when Sidney Greenbaum first suggested that a family of corpora be compiled, which would enable the comparative study of English varieties that are used in countries where the language is either spoken as an L1 by the majority of the population, or where it has the status of an official second language (Greenbaum 1991: 3–4). Since then, the number of ICE siblings has grown steadily, with corpora for 13 varieties already available, 14 new corpora in the making at the time of writing and new ICE teams continuously joining the project. Furthermore, new syntactically annotated editions have also been

published of some of the existing corpora<sup>64</sup>, which expands the usability of the corpora into new areas of research. The following ICE corpora are examined in the four articles included in the dissertation: ICE-Great Britain (ICE-GB), ICE-Hong Kong (ICE-HK), ICE-India (ICE-IND), ICE-Philippines (ICE-PHI) and ICE-Singapore (ICE-SIN). In addition, ICE-Jamaica (ICE-JA), sections of spoken Kenyan English from ICE-East Africa (ICE-EA), the written component of ICE-USA (ICE-USA) and spoken data from the yet unpublished ICE-Fiji (ICE-FJ) are also included in some of the articles.<sup>65</sup> Since tagged versions were not available for all of the aforementioned corpora at the time the articles were written, the original untagged versions were used, since differences in the search methods for the two types of corpora could have compromised the comparability of the results.

The strength of the ICE corpora lies in their unified structure<sup>66</sup>, which enables the detailed comparison between different Englishes across various genres. Each corpus contains one million words of which 600,000 words represent spoken and 400,000 words written English. A more detailed structure of the corpora is presented in Table 2.

---

<sup>64</sup> Annotated versions are now available for the following corpora: ICE-CA, ICE-GB, ICE-HK, ICE-IND, ICE-IRL, ICE-JA, ICE-NZ, ICE-NG (written), ICE-PHI, ICE-SIN, ICE-SL (written) and ICE-US (written) (<https://www.ice-corpora.uzh.ch/en.html>).

<sup>65</sup> I would like to thank Professor Marianne Hundt and her team for the opportunity to assist in the compilation of ICE-FJ and for granting me access to the unpublished data.

<sup>66</sup> While the majority of corpora follow this structure, there are some such as ICE East Africa (ICE-EA) which differ from it noticeably (for further discussion, see Hudson-Ettle & Schmied 1999).

<b>SPOKEN (300)</b>	
<b>Dialogue (180)</b>	
Private (100)	Face-to-face Conversations (90) Telephone Calls (10)
Public (80)	Classroom Lessons (20) Broadcast Discussions (20) Broadcast Interviews (10) Parliamentary Debates (10) Legal Cross-examinations (10) Business Transactions (10)
<b>Monologue (120)</b>	
Unscripted (70)	Spontaneous Commentaries (20) Unscripted Speeches (30) Demonstrations (10) Legal Presentations (10)
Scripted (50)	Broadcast News (20) Broadcast Talks (20) Non-broadcast Talks (10)
<b>WRITTEN (200)</b>	
<b>Non-printed (50)</b>	
Student Writing (20)	Student Essays (10) Examination Scripts (10)
Letters (30)	Social Letters (15) Business Letters (15)
<b>Printed (150)</b>	
Academic Writing (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)
Popular Writing (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)
Reportage (20)	Press News Reports (20)
Instructional Writing (20)	Administrative Writing (10) Skills & Hobbies (10)
Persuasive Writing (10)	Press Editorials (10)
Creative Writing (20)	Novels & Stories (20)

**Table 2.** Structure of the ICE corpora. Bracketed figures present the number of 2000-word files in each subsection (<https://www.ice-corpora.uzh.ch/en/design.html>)

In addition to their uniform structure, the data collected for each corpus must also meet the following criteria: all language samples must have been produced during or after the 1990s by informants who are at least 18 years of age, who have gone

through English-medium education in the country for which the corpora is compiled<sup>67</sup> and who were either born in that country or moved there while young (<https://www.ice-corpora.uzh.ch/en/design.html>). One shortcoming in the data collected for the corpora concerns its demographical representativeness; as noted on the ICE homepage (<https://www.ice-corpora.uzh.ch/en/design.html>), the selection of informants is not required to be balanced according to such demographic factors as age, gender, ethnicity, mother tongue or home region. Even though compilers of corpora frequently strive to collect data from a balanced and varied group of informants, this is not always possible, resulting in some demographic groups becoming over/underrepresented in the data. Fortunately, this problem can be minimized by consulting the metadata files that are provided with the corpora.<sup>68</sup>

Although ICE is a valuable source of data, the compilation of a family of corpora that is geographically and now also increasingly chronologically diverse is a challenging task. One of the questions this raises is whether a family of corpora that now spreads over three decades can really be said to be a source of synchronic comparable data. While this might be a greater problem for some research topics, this is less of an issue for the present study that focuses on syntax, which is not as susceptible to changes as, for example, lexis, as Mukherjee et al. (2010: 74) note. Another problem with the ICE corpora is also connected to the passing of time; the world has changed significantly from the 1990s when the guidelines for compiling the corpora were laid down, and consequently, the compilers of the newest ICE corpora are facing many new challenges regarding the criteria that should be followed when collecting data. Firstly, people travel and live abroad in greater numbers today, which makes it increasingly difficult to find people who meet the original requirements for informants. For example, as Biewer et al. (2010: 6–7) point out, the citizens of some countries have traditionally moved abroad to receive their secondary education, and thus, some corpora such as ICE-FJ and ICE-HK contain data from informants who have studied abroad (see also Bolt, cited in Biewer et al. 2010: 7).<sup>69</sup> Secondly, the global spread of Western popular culture and advances in

---

<sup>67</sup> For L1 varieties, this refers to the completion of secondary school, whereas “second-language countries might require a university degree” as Greenbaum (1991: 3) points out.

<sup>68</sup> Here it should be acknowledged that many of the metadata files have gaps in the speaker information provided and in the case of ICE-SIN, the files have never been made available for the academic community.

<sup>69</sup> During the analysis of data for the articles, it became apparent that this also applies to some speakers in ICE-PHI who, based on their personal descriptions of their past, have also studied abroad for longer periods of time.

social media have resulted in a situation where many people around the world are now exposed to different native varieties of English on a daily basis (see also Mukherjee et al. 2010: 68). This problem of “foreign” influences also extends to many of the written genres included in the corpora, since the global trend of using native English proof-readers can at times challenge the representativeness of the published material available for a variety (Biewer et al. 2010: 13). Thirdly, some categories in the original corpus design such as ‘social letters’ are now becoming obsolete, and collecting compensatory material from emails is not always unproblematic (see, for example Mukherjee et al. 2010: 67). However, as Biewer et al. (2010: 5-6) point out, following the original guidelines too meticulously could result in a situation where the compiled corpus is no longer representative of the variety and therefore, the compilers must sometimes strike a balance between representability and comparability.

Fortunately, the main focus of all four articles included in the present study is on spoken language and thus, many of the issues discussed above are of minor importance as they deal with issues related to written language. There are, however, other challenges brought on by the choice of data that became apparent during the research process. Since the focus of this study is on syntactic innovations, which occur most frequently in informal spoken language, the data used in the articles was often restricted to a single subcategory of the corpora, ‘Private conversations’, which contains only 200,000 words per corpus. This, when combined with the other requirements placed on the features (substrate-induced innovation in IndE and attested use in HKE, PhiE and/or SinE), narrowed the set of features that could be included in the present study and thus, only those that occur relatively frequently in the data could be studied.<sup>70</sup> Another problem was caused by the lack of annotated data for some varieties, which meant that only such features could be examined that could be retrieved relatively easily from the corpora (e.g. using search words). A third problem was related to the discrepancies between the corpora that became apparent during the research process – the feature-specific solutions to these issues have been discussed in articles A1–A4.

---

<sup>70</sup> In fact, there were other promising innovative IndE features such as the presentational use of *itself* (see, for example, Lange 2007, 2012) which eventually had to be left out of the present study due to the feature’s low level of use in the data.

### 3.1.2 The Santa Barbara Corpus of Spoken American English

The *Santa Barbara Corpus of Spoken American English* (SBCSAE) contains some 249,000 words of data collected from 60 different naturally occurring interactions recorded in various parts of the US at the turn of the millennium. Since the spoken component of ICE-US has not been released yet, SBCSAE was included in the study to represent spoken AmE (articles A1, A2 and A3) – a decision that can be justified by the fact that SBCSAE will also form a part of the spoken component of ICE-USA (<http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>). Although the files in SBCSAE mostly consist of private face-to-face conversations, other genres such as telephone conversations and classroom lectures are also represented, and thus, the SBCSAE data correspond with the ICE data quite extensively.<sup>71</sup> However, there are also some files in SBCSAE, such as the public story-telling event with interactional elements, which do not neatly match any of the categories of the ICE corpora. Despite this, using data from SBCSAE to compensate for the lack of ICE-USA is common practice among scholars working with ICE corpora (see, for example, Collins 2005, 2009; Rautionaho 2014), and it could be argued that issues such as the one mentioned above are only problems if they are not taken into consideration when comparing the results obtained from the ICE and SBCSAE corpora.

### 3.1.3 The Freiburg-Brown Corpus of American English

The *Freiburg-Brown Corpus of American English* (Frown) is part of the Brown family of corpora and it contains one million words of written American English published in the year 1992. Frown consists of 500 2000-word files, which represent 15 different text categories. At the time of the first article (A1), the written component of ICE-US was not yet available and therefore data from Frown was used to represent written American English. Although the structures of Frown and ICE do not fully correspond to each other in size and selection, they nevertheless have many parallel categories such as ‘Press editorials’, ‘News reports’ and ‘Fiction’.<sup>72</sup> Furthermore,

---

<sup>71</sup> For a detailed list of the types of interactions included in the SBCSAE data, see <https://www.linguistics.ucsb.edu/research/santa-barbara-corpus#Contents>. To see how the SBCSAE data correspond to the ICE categories, see Appendix A.

<sup>72</sup> For a complete list of the categories of Frown and how they correspond to the categories of ICE, see Appendix B.

scholars such as Collins (2005, 2009) have also used data from Frown to compensate for the lack of the written component of ICE-USA in the past.

### 3.1.4 Summary of the data

Table 3 summarises the distribution of the various corpora used in the four articles.

	ICE-Corpora									SBCSAE	FROWN
	EA	FJ	GB	HK	IND	JA	PHI	SIN	US		
A1			✓	✓	✓		✓	✓		✓	✓
A2			✓	✓	✓		✓	✓	✓	✓	
A3	✓	✓	✓	✓	✓	✓	✓	✓		✓	
A4				✓	✓		✓				

**Table 3.** Corpora used in the four articles

As can be seen, the ICE corpora provide the main source of data for all varieties, and SBCSAE and Frown are only used to compensate for the lack of ICE-US. Although the selection of varieties differs from article to article to some extent, the focus of the present study is on HKE, IndE, PhiE and SinE, whereas the superstratal varieties BrE and AmE are included for points of reference.<sup>73</sup>

## 3.2 Methods

Since descriptions of the methods used in each of the four articles are already given in A1–A4, elaborations on the individual choices made therein will not be provided here. There are, however, some broader questions regarding the methodological choices made when studying emerging epicentres which will be discussed here.

Hundt (2013: 191, basing her argument partly on Meyerhoff & Niedzielski) suggests that when studying emerging epicentres empirically, the following four criteria should be met: firstly, it should be established that the variant has “structural and functional equivalence” in the varieties examined. Since the focus of the present study is on syntax, the structural likeness of the three features has already been established (for details, see A1–A4), whereas meeting the requirement of functional

<sup>73</sup> For explanations for the choice of varieties included in each article, see articles A1–A4.



equivalence was more difficult to determine for one of the three features studied here; although previous studies have shown that *also* and *only* have acquired new functions as presentational focus markers in IndE, the data contained too many instances of particles that could have been given both the canonical (additional/restrictive) and the innovative (presentational) reading. As a consequence, a decision was made to focus solely on their location within the clause. This decision was further justified by the fact that the innovation in IndE also has a structural aspect, that is, *also* and *only* are always placed after their centre of focus (for a more detailed discussion of this, see A1 and A4). Furthermore, studying the semantic functions of the two particles in the corpora might have been redundant, since it is also possible that only the structural aspect of the innovation has been copied to the receiving varieties, as the post-referent location of the particles is already licensed (but not as frequently used) by the English grammar, and because structural features, especially on a clausal level, are borrowed more easily from one variety to another (see, for example Thomason 2001: 69; Lange 2012: 246). For the other two features examined in this study, the use of *isn't it* as an invariant tag and the omission of direct objects, the question of functional equivalence was easier to prove, as the innovativeness of the features is more connected to their form than function: although the invariant tag *isn't it* has lost its grammatical markedness, it has nevertheless retained its (primary) original function as a question marker, while the tendency to omit direct objects is not so much connected to new functions as it is connected to the licensing of omitting elements of low information value from the sentence even in cases where their presence would be grammatically required. While the argument could be made that both of these features are fairly universal tendencies in languages, to claim that this is the sole explanation for their existence in the varieties studied would be too reductionistic. For example, Sharma (2012) argues that when the different factors that could explain the existence of shared features in New Englishes are examined, they should be considered in the following order: properties shared with the superstrate(s) (including epicentres), properties shared with the substrate(s), acquisitional universals and finally, general universals, and therefore, other explaining factors should be explored more thoroughly before resorting to more generalistic explanations such as language universals. It should also be noted here that while epicentric influence is included in the first explaining factor on Sharma's (2012) list, IndE's epicentric influence in Southeast Asia has not been established yet and consequently, other factors such as substrate influence should also be taken into consideration when evaluating the explanatory power of the different factors involved. A further aspect included in these considerations is the

location of the varieties on the three different models (Kachru 1985; Schneider 2007; Mair 2013) used in this study.

The second criterion for studying potential epicentres according to Hundt (2013: 191) is that “the study should distinguish external influence of the (potential) epicentre on another variety from independent regional (parallel) developments”. While this is often difficult to prove beyond doubt, the likelihood of these two explanations can be delineated by following the order provided by Sharma (2012), that is, by first looking at the superstrate(s), followed by the substrates. For the three features examined for this study, superstrate influence was not considered to be a likely (sole) explanatory factor in any of the varieties, and therefore, a more detailed investigation into the local substrates and linguistic ecologies was needed. Because the results of all three features could be explained with a combination of answers related to superstrate/substrate/IndE influence, the two final aspects on Sharma’s (2012) list, acquisitional and general universals, were not examined further in the articles.

The third requirement Hundt (2013: 191) raises regarding the methods that should be used in a study focusing on potential epicentric influence is that they should “assess the evaluation of the variable in the variety that is adopting it, i.e. whether speakers use the feature consciously or unconsciously”. Since the data used for this study come from anonymised corpora that were compiled two to three decades ago, it is impossible to interview the speakers on how consciously they used the three features studied here, or how prestigious / stigmatised they perceived them to be. However, some indirect conclusions of the features’ acceptability can be drawn by examining their spread across different genres (articles A1 and A2), and the age and gender of the speakers who use them (article A4). If the results show that a feature is used in more formal contexts, it is likely to be more widely accepted among the speaker population. Also, if there are signs that the feature is used more in such written genres as student essays and examination scripts (which provide the data for the category ‘Academic writing’ in the ICE corpora), it gives cause to suspect that the younger generations perceive the feature to be more acceptable than the older generations and thus, it is possible that the use of the feature will continue to spread to other written genres over time. Although it could of course also be argued that the appearance of innovative features in students’ writing is caused by their inability to adhere to the conventions of academic writing, the mere fact that the feature exists in their written language too and that they do not perceive it to be a strongly stigmatised feature supports the argument that it could be on its way to becoming more accepted among the younger speakers of the variety (see also Balasubramanian

2009: 225). The last article included in this study (A4) also looks into the differences related to the age and gender profiles of the users of one feature, the use of clause-final *also* and *only*, with the hypothesis that if the feature is used more by younger/female than older/male speakers, the results would support the argument that the feature is spreading in the variety, since younger and female speakers are often the first to use a feature before it spreads to the whole speaker population (for detailed discussion on the methods used, see A4). Moreover, it was assumed that if the feature has (proportionally) more older/male users in IndE than in the other varieties studied, the results would indicate that the feature is more established in IndE, which in turn lends support to the argument that IndE could have contributed to the growing use of the feature in other varieties. Unfortunately, the other two features examined for this study, the invariant tag *isn't it* and the omission of direct objects could not be examined using this method, as their frequencies in the data were too low for this type of a statistical analysis.

Fourthly, Hundt (2013: 191) states that “any verification of the epicentre status of a variety has to take the social, economic and cultural context into account and evaluate whether it fosters or hinders epicentric influence of one variety on another”. While these issues have been discussed in all four articles to some extent, a more detailed investigation into these aspects will be presented in section 4.3.3, where they will be examined in connection with the results of the four articles.

## 4 RESULTS

In this section, the results of the four articles will be discussed in connection with the four main research questions of this study. The subsections in this chapter will present summaries of the main findings of the articles, while more detailed discussions of the results can be found in the articles themselves (A1–A4, above).

### 4.1 Research question 1

The first research question asks whether the use of clause-final focus particles *also* and *only*, invariant tag *isn't it*, and the tendency to omit direct objects could be argued to be syntactic innovations in IndE. In order to establish this, their frequency and level of acceptability in IndE and Standard English (StE, represented here by BrE and AmE) must be discussed.

#### 4.1.1 Clause-final *also* and *only*

The first feature, the use of clause-final focus particles *also* and *only*, is examined in articles A1 and A4, but before elaborating on the innovativeness of this feature in IndE, the canonical uses of these focus particles in English should be described in greater detail. König (1991: 32–3) states that the function of the additive focus particle *also* (1) and restrictive *only* (2) is to draw attention to a particular part of the utterance (*creativity*, *flowers*) which is called the particle's centre of focus.

- (1) Creativity is *also* important. (ICE-IND:S1A-044#146:1:A; emphasis added)
- (2) There was uh no body there was *only* flowers. (ICE-IND:S1A-012#141:1:B; emphasis added)

Although there are multiple locations within the clause that the focus particle can occupy, as exemplified in (3), their location signals which element in the clause is

focused and therefore, any changes to the location of the particles within a sentence will also often affect their meaning (Biber et al. 1999: 781).

(3) ‘(Only) they (only) fed (only) the cats (only)’. (Nevalainen, 1991: 39)

Despite the multiple locations licensed by the English grammar, the most common location for these particles is between the operator and the main verb (Quirk et al. 1985: 605; Fuchs 2012: 37) and while clause-final location is an option even in StE, it is not very frequently used (see, for example Nevalainen 1991: 33; Fuchs 2012: 35), because the final position in the clause is usually occupied by the focused element itself (Carter & McCarthy 2006: 778). However, when the particles are placed in clause-final position, this tends to occur in more formal and/or written contexts (A1, Fuchs 2012; see also Huddleston & Pullum 2002: 590).

Interestingly, the usage pattern for clause-final *also* and *only* has been reversed in IndE, as the highest frequencies of their use can be found from the most informal and spoken contexts (A1). The origins of the feature in the Indian variety can be traced back to the substrates (see, for example, Balasubramanian 2009; Fuchs 2012; Lange 2007) – especially to two Hindi particles which can signal inclusive (*bhi*) and exclusive (*hi*) focus (Sharma 2003: 60–2) and which thus resemble the canonical functions of *also* and *only* in StE.<sup>74</sup> There are, however, two aspects by which the Hindi particles differ from the English *also* and *only* and which have presumably contributed to the rise of their innovative use in IndE. Firstly, unlike their English equivalents, the Hindi particles always follow their referents (Sharma 2003: 64) and secondly, they have acquired new meanings as presentational focus markers in the variety (Lange 2007; Bhatt 2000; Fuchs 2012) where they are used to signal the presence of “new, nonpresupposed information” (Kiss 1998: 246). Examples of this new non-canonical use in IndE are given in (4) and (5):

(4) That’s now you’re offering offering me [a party] without my guests *also*.  
(ICE-IND:S1A-003#27:1:B, emphasis added)

---

<sup>74</sup> Here it should be noted that Sharma’s (2012) list of explanatory factors was originally meant to be applied when looking at features shared by new Englishes (and not when looking at individual varieties like here), but since the superstrates cannot explain the presence of the feature, the next step is to look at substrates.

- (5) Means ma- calculation should be finished of you know first *only*  
(ICE-IND:S1A-071#83:1:A, cited in Lange, 2007: 109, emphasis added)

Therefore, the tendency of speakers of IndE to place *also* and *only* in clause-final position can be explained with substrate influence, which has given the English particles new functions, in addition to creating new patterns of preference regarding their location within the sentence.

#### 4.1.2 Invariant tag *isn't it*

The second innovative feature (studied in A2) is the use of *isn't it* as an invariant question tag and again, before proceeding with a more detailed description of the feature in IndE, a brief description of the canonical uses of the tag should be given. Tag questions are shortened yes-no questions located at the end of statements, where their function is to request the listener to confirm the truth-value of the proposition presented (Leech & Svartvik 2002: 132, 382–3). The tag consists of a pronoun which refers back to the subject of the main clause (Huddleston & Pullum 2002: 893; Quirk et al. 1985: 810) and a verb which is also determined by the main clause – if the main clause contains a modal verb, the lexical verb *be*, or an auxiliary *be*, *do* or *have*, the verb is repeated in the tag, but if the verb in the main clause does not meet these criteria, an appropriate form of *do* is used instead (Carter & McCarthy 2006: 547–8). The two different types of tag questions relevant for a discussion on the invariant tag *isn't it* include exclamation tags, which follow exclamative *wh*-clauses (Carter & McCarthy 2006: 551) (6) and question tags, which can be divided further into reversed (7) and constant polarity tags (8). Of these two, the former is the most common type (Huddleston & Pullum 2002: 892; Quirk et al. 1985: 810), while negated versions of the latter are exceedingly rare and are even considered ungrammatical by some as (9) shows (see, for example, Quirk et al. 1985: 813; Huddleston and Pullum 2002: 787):

- (6) What a beautiful painting it is, *isn't it*? (Carter & McCarthy 2006: 551)

- (7) She is a teacher, *isn't she*? (Carter & McCarthy 2006: 547)

- (8) I haven't seen you before, *have I*? (Quirk et al. 1985: 810)

- (9) \*So he doesn't like his job, *doesn't he*? (Quirk et al. 1985: 813)

The findings presented in A2 show that while the use of *isn't it* as an invariant tag is rare in BrE and AmE, it is very common in IndE (see also Balasubramanian 2009; Columbus 2010; Lange 2012). The origins of the feature can be traced back to substrates such as Hindi, which has two invariant tags, *hai na* ('is not') and *na* ('no'), which can be placed at the end of statements to form questions as in (10):

- (10) vah kal baazaar gayaa thaa, *hai na*  
 'He went to the market yesterday, *didn't he*' (Agnihotri 2007: 30, italics added)

Furthermore, as Agnihotri (2007: 31), notes, "many speakers of English in India tend to replace the whole range of Standard English tag questions by either 'isn't it?' or just 'no?'" and therefore, substrate influence provides a likely explanation for the innovative use of the invariant tag *isn't it* in IndE.

#### 4.1.3 Omission of direct objects

The final feature examined in this study is the tendency to omit direct objects, discussed in A3, which has been noted to be common in IndE (in addition to many other New Englishes). In order to establish the extent in which direct objects can be omitted in English, issues related to how transitivity is expressed in the language should be discussed in greater detail. Kittilä (2002: 78–9) notes that transitivity in English is fundamentally determined by the number of core participants, and even though verbs and clauses can both be categorised as transitive, Huddleston and Pullum (2002: 216) argue that it is the verb which ultimately also determines the transitivity of the clause. For instance, in the examples below, the verb *eat* permits both the omission (11) and the presence (12) of the direct object.

- (11) He ate. (SV)

- (12) She was eating an apple. (SVOd)

A verb can often have both an intransitive (11) and a transitive (12) function, and the latter group can be further divided into monotransitives, which can be ordinary (12) (SVOd) or complex (13) (SVOdAo, though also SVOdCo is possible), and ditransitives, which can only be ordinary (14) (SVOiOd) (Huddleston & Pullum 2002: 218).

(13) The neighbours kept the dog inside the house. (SVOdAo)

(14) She sent him a letter. (SVOiOd)

Even though the intransitive/transitive uses of the same verb are common even in canonical varieties of English such as BrE and AmE (see, for example, Platt et al. 1984: 117; Subbārão 2012: 28; Baskaran 2004: 1080), a number of New Englishes such as HKE (Platt 1982: 410), MyE (Baskaran 2004: 1080), SLE (Platt et al. 1984: 117) and IndE (Subbārão 2012: 28) have been noted to omit direct objects more frequently – an argument which is also supported by the results of A3. Furthermore, as the results of A3 show, IndE has the highest tendency to omit direct objects when compared with the other New Englishes included in the study, and this can again be explained with the influence of IndE substrates such as Hindi, where the morphology is rich enough that object omission is sometimes allowed as in (15).

(15) mēne                      kahi:.  
1SG-erg                      said-fs  
'I said (it) to him/her.' (Koul 2008: 214)

Another factor that has likely contributed to the elevated tendency to omit direct objects in IndE is the topic prominence (Tp) of the substrates<sup>75</sup>; in Tp languages, objects refer to the topic of the sentence more frequently than in subject-prominent (Sp) languages such as English, and therefore, they can be omitted more freely as their referents are clear from the context (for a more detailed discussion on this, see A3).

#### 4.1.4 Summary of the answer to research question 1

As the discussion in the above sections shows, the three features studied in the articles can indeed be argued to be linguistic innovations in IndE. The first piece of evidence to support this argument is that the features are used very rarely – if at all – in the superstrates (BrE and AmE) as the results of A1, A2 and A3 indicate and

---

<sup>75</sup> Although there are some differing opinions of the classification of Hindi since some scholars such as Kidwai (2004: 255) argue Hindi to be a Sp language, others (see, for example, Junghare 1988: 316; Sato 2011: 362) argue that the language is actually more Tp.



therefore, it is unlikely that their input has given rise to the use of these features in the Indian variety. The second piece of evidence that strongly supports the argument that these features are local innovations is the fact that all three features have close equivalents in the local Indian substrates.

## 4.2 Research question 2

The second research question asks how frequently these innovative features are used in the Asian varieties. Unfortunately, since all the articles either analyse different sections of the corpora and/or use different statistical means, their results cannot be neatly combined into a single table. The closest point of comparability can be reached by focusing on the results from the most informal category, ‘Private conversations’, where innovative features occur most frequently, and which was examined for all articles. Table 4 below presents the order in which the innovative features are used in the four Asian varieties from the most frequent to the least frequent according to the results of A1 to A3.<sup>76</sup>

Variety	<i>also</i>	<i>only</i>	<i>isn't it</i>	Od omission
IndE	1.	1.	1.	1.
SinE	2.	2.	2.	2.
HKE	4.	3.	3.	4.
PhiE	3.	4.	4.	3.

**Table 4.** Order of varieties from most frequent to least frequent use in data from the ‘Private conversations’ section of the ICE-corpora

It is important to note here that the above table only provides a simplification of the results of the original articles included – it does not show how much the varieties differ from one another, nor does it give any indication of the factors (IndE or substrate influence etc.) that could explain the results for different varieties (for more detailed information, see articles A1–A3). What the results summarized in Table 4 do show, however, is that for all three features studied here, IndE shows the highest

---

<sup>76</sup> Since the focus of A4 differs from those in A1 to A3 to some extent (number of varieties included, focus on mapping the demographic profiles of the users of clause-final *also* and *only*), the results of A4 have been left out from Table 4 and will be discussed to some extent in 5.3.2 and in greater detail in section 4.4.

frequency of use and the second highest frequencies can be found from SinE, whereas for the last two varieties, HKE and PhiE, the order varies from feature to feature. Furthermore, since all four Asian varieties show clearly elevated uses of the features when compared to their superstrates BrE and AmE (see A1–A3), it is unlikely that superstrate influence could explain the results for any of the Asian varieties examined here. Another interesting point worth mentioning concerning the order of varieties in Table 4 is that though not directly comparable, they do mirror the order of varieties presented by Hogue (2001, see also section 2.4), who studied the presence of Anglo-Indian loan words in Southeast Asian Englishes.

### 4.3 Research question 3

The third research question asks what could explain the differences between the varieties examined in the articles. As mentioned in the Methods (section 3.2), the present study uses Sharma's (2012) theoretical framework when estimating the likelihood of different explanatory factors, which are 'Properties shared with the superstrate' (Type A, section 4.3.1), 'Properties shared with the substrate' (Type B, section 4.3.2), 'Acquisitional universals' (Type C) and 'General universals' (Type D). The order in which these factors are listed is not arbitrary: as Sharma (2012: 221) argues, "[i]t is logical to first examine the languages in contact [superstrates and substrates] for proximate causes of shared features before appealing to more general motivations [of acquisitional and general universals]".<sup>77</sup> Because potential explanations for all three features could be established from the groups of proximate causes (discussed below, but see also A1–A3), explanations related to the general motivations will not be discussed in separate subsections here.

#### 4.3.1 Properties shared with the superstrate

According to Sharma (2012: 222), explanations in this category include 'Founder effects', 'Exogenous prestige-driven mimicry' and 'Diffusion from one variety to another'. While some hits for the features studied can often be found even in the

---

<sup>77</sup> Both Hundt (2013: 191) and Sharma (2012: 219) comment on the importance of establishing genuine similarity in the use of the features studied before exploring the different alternatives – for a discussion of the features examined in the present paper, please see Methods.

superstrate varieties, as the results of A1–A3 indicate, their frequencies of use are significantly lower when compared with the Asian varieties. Therefore, even though some of the features, such as clause-final focus particles, might have entered the feature pools (Mufwene 2001) of the Asian varieties already via the superstrate, the new patterns of uses they have acquired in the Asian varieties suggest that the primary cause for their growing use is located elsewhere.

One further superstrate-related explanation that should be discussed here is the possibility that BrE and AmE could still have given rise to the features in (some of) the Asian varieties included in this study, albeit less directly. As noted above, clause-final *also* and *only* are already licensed by the grammar of StE, and therefore, it is possible that the increased use found in the four Asian varieties is actually the result of each variety having picked up this feature in disrupted transmission independently, which could have then developed into the more prominent use of feature we see in the Asian varieties today.<sup>78</sup> While this is indeed a plausible explanation, it also gives rise to some questions. Firstly, since this would be a case of disrupted transmission, the question that should then be asked is if this explanatory factor, though connected to the superstrate, is actually closer to explanations related to language acquisition. If this is indeed the case, the explanation would be better placed under category C ‘Acquisitional universals’ (Sharma 2012), which in turn means that this explanation should be considered only after explanations related to category B ‘Properties shared with the substrate’ have been excluded. Secondly, it remains unclear what is the likelihood that such parallel but independent developments in the four varieties would have given rise to the particular pattern that we see repeated in the results of this study. These issues will be discussed further after the possibility of substrate influence has been addressed.

The second explanation, ‘Exogenous prestige-driven mimicry’ is also unlikely as the two superstrates included in the studies are also the old (BrE) and new (AmE) prestige varieties (see also Mair 2013). Because IndE has significantly higher frequencies of use for these features when compared with the Southeast Asian varieties included in this study and because the variety has been shown to function as an epicentre for other English varieties spoken in South Asia, the possibility of ‘Diffusion from one new variety to another’ is a potential explanatory factor for HKE, PhiE and SinE. However, since IndE’s epicentric influence over Southeast Asian varieties has not been established yet, this issue will be examined further after

---

<sup>78</sup> Considering the structural aspects of canonical tags in StE, the same explanation could also be applied to the use of the invariant tag *isn’t it* in the four Asian varieties.

another proximate explanation, influence from the substrates, has been discussed in greater detail.

### 4.3.2 Properties shared with the substrate

According to Sharma (2012: 223), the two explanatory factors under ‘Properties shared with the substrate’ are ‘Accidental resemblance’ and ‘Areal convergence’. One of the challenges for the present study concerns the linguistic diversity of the South and Southeast Asian region, which often makes it difficult to present estimates on which of the two explanatory factors would be more likely even though it is clear that some substrate effect is at play. For example, for the tendency to omit direct objects in a number of Asian Englishes, the results could, at least partially, be explained with the dominance of topic-prominent (Tp) languages in the region. Because of this, some local languages belonging to subject-prominent (Sp) language branches have also developed more Tp features such as object omission, which in turn increases the likelihood of direct objects being omitted also in the local varieties of English. Therefore, for the tendency to omit direct objects, ‘Areal convergence’ seems the likelier explanation of the two. For more definitive arguments, more detailed explorations into the dynamics between the numerous South and Southeast Asian substrates would be required, but this would exceed the scope of this study. Furthermore, such an investigation would not provide information that would be of central importance for the present topic, for which establishing substrate influence in the varieties is sufficient. As a consequence, the remainder of this section will focus on discussing the results of the Asian varieties only in relation to their substrates and whether the three features studied exist in them or not.

For clause-final *also* and *only*, substrate influence can explain the results in the Asian varieties to varying degrees (see also A1 and A4). As was discussed in 5.1.1, substrates provide a very likely explanation for the emergence of the feature in IndE, but since the results in A1 and A4 do not include a semantic analysis of the data (for the rationale for focusing on syntax, see section 3.2 ), there are a few issues that should be discussed concerning the effect this might have had on the results on *also* and *only*. As previous studies show (see, for example, Lange 2007; Fuchs 2012), these particles have both retained their old meanings in addition to acquiring new uses as presentational focus markers in IndE. Therefore, it is possible that the higher frequencies reported here for IndE could actually be caused by these new additional uses elevating the total figures for clause-final *also* and *only* in the variety, whereas the

lower frequencies found in SinE, HKE and PhiE could merely signal that the new uses have not spread to the three Southeast Asian varieties. While this could indeed offer some explanations for the differences between IndE and the Southeast Asian varieties, it is unlikely that it could explain them completely; Lange's (2012: 188) study on the use of contrastive and presentational *only* in the category 'Private conversations' in ICE-IND shows that the presentational type forms only 17 per cent of the total number of cases in the Indian data.<sup>79</sup> Similar results have been presented for *also* by Fuchs (2012: 33), who shows that presentational uses form only 4 per cent of the total number of hits for *also* in IndE, but since this figure is based on both the spoken and written components of ICE-IND, it is likely that this percentage is actually higher in the category "Private conversations".<sup>80</sup> Despite this, it could be argued that the results of Lange (2012) and Fuchs (2012) indicate that while the new presentational uses could indeed have elevated the total frequencies for clause-final *also* and *only* in IndE to some extent, these new uses do not occur frequently enough in the ICE-IND data to solely explain the differences between the Indian and Southeast Asian varieties. Because of this, a more detailed discussion of the possible explanatory factors for each Southeast Asian variety is given below.

While particles resembling clause-final *only* can also be found in Hong Kong Cantonese, the dialect lacks a particle that would be equivalent to clause-final *also* (Hiramoto 2015: 644) and therefore, substrate influence does not provide a satisfactory explanation as to why both *also* and *only* occur equally frequently in HKE (see also A4). Furthermore, as the results of A4 indicate, both clause-final *also* and *only* are used more frequently by women and younger speakers of HKE, which supports the argument that the feature has begun to spread in the variety more recently when compared with IndE, where the tendency is more established. This in turn lends support to the argument that IndE could have contributed to the emergence of the feature in the variety, but the likelihood of this explanation should be reassessed only after the connections between Hong Kong and India have been discussed in greater detail (see section 4.3.3).

---

<sup>79</sup> Lange (2012: 188–9) also notes that the use of contrastive *only* in IndE is approximately twice as frequent when compared with BrE, which she suggests is due to the fact that "a speaker of BrE would then be more likely to mark a contrastive focus construction by prosodic prominence than a speaker of IndE, who is more likely to prefer the free morpheme *only* as a focus marker".

<sup>80</sup> Interestingly, Fuchs's (2012: 33) study shows that the proportion of additive uses is the same, approximately 78%, in both IndE and BrE, but the frequencies for the additive cases are again over twice as high in IndE when compared with BrE and thus the pattern that emerges for *also* mirrors that of *only*.

As for the results of PhiE, the substrate could possibly explain the existence of the feature in the variety: Tagalog, the variety's largest substrate also has particles that resemble the canonical uses of *also* and *only* (Ramos 1971: 49–50; Ramos & Cena 1990: 99) and while they can occupy clause-final positions (see, for example, Ramos 1971: 49–50), this is not their usual position within the clause (Ramos & Cena 1990: 98). As the results of A1 and A4 indicate, PhiE has only a somewhat elevated tendency to use clause-final *also* and *only*, and the use of the former is still possibly growing, while the use of the latter seems to be more stable in the variety. Therefore, it is possible that the results of PhiE are solely caused by substrate influence in the variety.<sup>81</sup>

For the results of SinE, clause-final particles translatable as *also* and *only* can be found from Bazaar Malay, Mandarin Chinese and Cantonese (see, for example, Aye 2005: 70; Matthews & Yip 1994; Lin 2001; Ross & Ma 2006; Bao & Hong 2006: 110). In addition, SinE has a wide variety of clause-final particles (for a more detailed description, see, for example, Wee 2004: 1068; Lim 2007) serving various discourse-pragmatic functions, which could have reinforced the tendency to use clause-final *also* and *only* in the variety. It is also possible that the existence of these particles has caused the speakers of SinE to adopt the use of the feature from IndE more readily. Sharma (2012: 223) suggests that “in situations of overlap, where we find evidence of both Type A and B effects, further varieties may need to be added to the comparison, or reinforcement may be arising from both sources”. For the results of SinE, a further point of comparison is provided by HKE, another former British trade colony in the region with a strong Chinese substrate.<sup>82</sup> Interestingly, even though a close equivalence for clause-final *only* can be found from the substrates of both varieties, the tendency is significantly stronger in SinE than in HKE; one explanation for the higher results of the former variety is that it has been influenced by IndE to some degree. The same explanation could also be offered for the results on *also*, which are again significantly higher for SinE when compared with HKE.

---

<sup>81</sup> Though it is interesting that the order in which the particles seem to have become more established in the PhiE variety (*only* preceding *also*) follow that of IndE, the results of PhiE are not statistically significant.

<sup>82</sup> Even though using HKE for this purpose might be deemed problematic (SinE being more advanced than HKE in the Dynamic Model, or the differences in the feature pools of the two varieties) HKE is nevertheless the only other New English variety that is spoken in the region which also has a Chinese dialect as a major substrate and which is represented in the ICE family of corpora, thus making the variety the best point of comparison for SinE that is available at present.

Substrate influence offers a potential explanation also for the use of *isn't it* as an invariant tag in some of the varieties. As the discussion in A2 shows, all substrates examined for the study use invariant tags, which could be expected, since particles that are attached to the end of leading questions are common in many languages (Dryer 2007: 93). The temptation to categorise the use of invariant *isn't it* in the Asian varieties as a case of general universals should nevertheless be resisted, since, as Sharma (2012: 223) notes, “[t]rue emergence of universals would only be certain if the substrates do not have a particular unmarked feature, and yet the feature arises in offspring varieties of English”. As noted in A2, PhiE shows only slightly elevated frequencies of use of the invariant tag *isn't it* and these could be the result of substrate influence in the variety. Interestingly, although the closest equivalent to the English *isn't it* can be found in Hindi, also Mandarin Chinese and Cantonese have particles that resemble the structure of the English tag to some extent. Despite this, SinE shows a significantly higher tendency to use invariant *isn't it* than HKE, which is surprising, considering that the same structure is represented in the feature pools of both varieties. It should be noted here that *isn't it* is not, of course, the only invariant tag used in the varieties included in this study, and the preferred invariant tag in both SinE and HKE is actually *is it* (Columbus 2009: 408; Hoffmann et al. 2014: 703). Despite this, as the results of A2 show (see also Columbus 2009: 408; Hoffmann et al. 2014: 702), SinE clearly outnumbers HKE in its use of the invariant *isn't it* and one possible explanation for this is that the results of the former variety have been elevated by IndE influence, while the lower figures for HKE might be solely caused by substrate influence.

There is one further aspect related to the results on invariant tags that should be discussed in the contexts of substrates. As noted above, *isn't it* is not the only invariant tag used in the Englishes included in A2, and, depending on the variety, there are significant differences in the number of other options available for their speakers. For instance, some of the most common invariant tags in IndE are *isn't it*, *na* and *no* (Columbus 2009: 408), whereas SinE favours *is it* in addition to using a large number of substrate-based sentence-final particles such as *lah* and *wah* (Columbus 2009: 408). Therefore, it is possible that the lower frequencies of *isn't it* in the results of SinE are simply caused by the fact that a larger selection of other invariant tags exists in the variety. While this might indeed explain the results to some extent, there are a few considerations that should be discussed. Firstly, this could not explain the differences between the results of SinE and HKE; Setter et al. (2010: 78) point out that even though Cantonese is a major substrate for both varieties, the use of substrate-based sentence-final particles is rare in spoken HKE, which is

interesting, since it means that HKE (with its narrower selection of invariant tags) still uses *isn't it* less when compared with SinE.<sup>83</sup> Therefore, it could be argued that a narrower selection of invariant tags in a variety does not necessarily result in higher frequencies for the existing invariant tags in that variety. Secondly, it is also important to remember that the overall frequency for the use of invariant tags is not the same for each variety. For example, Columbus's (2009) study on the use of 17 invariant tags<sup>84</sup> in BrE, NZE, IndE, SinE and HKE shows that there are significant differences between the varieties and that the highest use could be found in SinE, followed by IndE, NZE, HKE and BrE (Columbus 2009: 408). Interestingly, somewhat similar results for the use of canonical tags in BrE, SinE, HKE and IndE have also been reported by Hoffmann et al. (2014), whose study shows that for the three Asian varieties, the highest frequencies of canonical tags can be found in SinE, followed by HKE and IndE.<sup>85</sup> Even though the results of these two studies do not easily compare with those presented in A2 as they all examine slightly different phenomena using different methods, it could be argued that the broader picture that emerges from them is that the existence of additional invariant tags in a variety might not necessarily result in significantly lowered frequencies for a particular tag (such as invariant *isn't it*) and that the wider selection of tags available could also give rise to an elevated tendency to use tags in the variety in general. This does not of course mean that a wider selection of invariant tags in SinE would not have any effect on the use of *isn't it* in the variety, but it is difficult to estimate if and to what extent the increased use of tags could have counterbalanced the effects of the reduction in the proportion of *isn't it* in SinE.

For the tendency to omit direct objects, substrate influence (complex morphology and topic prominence) offers the most plausible explanation for most varieties included in A3 (including PhiE), but the differences between SinE and HKE are noteworthy as they both have strong Tp substrates and yet, the results for SinE are significantly higher when compared with HKE. Again, a potential

---

<sup>83</sup> The work on establishing the exact origins of the particles in SinE is still ongoing, and while many of these particles might actually come from other substrates such as (Bazaar) Malay or Hokkien (see, for example, Lim 2007), many such particles also exist in Cantonese.

<sup>84</sup> These included both English tags (such as *you know* and the invariant *isn't it*) and substrate-based tags (such as *lah* and *na*).

<sup>85</sup> The study by Hoffmann et al. (2014) also includes an interesting analysis of the pragmatic functions of canonical tags in IndE, SinE and HK, and conducting a similar analysis on the different functions of invariant *isn't it* in the varieties included in A2 could help determine whether the patterns found in the Southeast Asian varieties mirror those of IndE not only in form but also in function.



explanation for differing results between the two varieties is that the Singaporean variety could have been influenced by IndE.

In addition to the feature-specific discussions presented above, there are also two other factors that could help shed more light on the results presented in A1–A3. The first involves analysing the frequencies and proportions of corresponding features in the substrates, as this could help identify those features that are more likely to be reflected in the local variety of English in some form. Such an analysis could bring forth new important information that might help explain the results presented here more fully, even causing the explanation related to IndE’s potential influence become redundant in some cases. However, since comparable data that would enable such an analysis is not available yet, the discussion presented here is based on a simpler approach which examines the various corresponding structures in the substrates and compares them to the structures of the innovative features in the local varieties of English.<sup>86</sup>

The second factor is more specific, as it might help explain the differences between the results of SinE and HKE which have been noted to exist for all three features studied here. Even though the influence of IndE is a factor that could help explain the differences between SinE and HKE, it is also possible that the higher frequencies found in SinE could actually be caused by the variety’s higher level of nativization; as was mentioned in sections 2.2.1.2 and 2.2.4.2, SinE is already showing signs of entering the differentiation phase, the fifth and final stage on Schneider’s (2003, 2007) Dynamic Model, whereas HKE is still in stage three, the nativization phase. However, the extent to which this can be used to explain the results of the present study is challenging to estimate. While there are studies that suggest that the differences in which nativized features are used in New Englishes could in some cases be explained with their location on the Dynamic Model (see, for example, Gries & Mukherjee 2009; Hoffmann et al. 2014; Götz 2017: 277), Gries and Mukherjee (2010: 542) argue that “a neat alignment of evolutionary stages on

---

<sup>86</sup> For example, such an examination for the invariant *isn’t it* would include first finding corpora that contain spoken data on the major substrates of the local varieties of English (preferably private conversations to maximise comparability with the ICE-data), and then conducting multiple searches to identify all invariant tags in the corpora. This would then be followed by a detailed analysis of the frequencies of each tag, clarifying how often they are used in relation to each other, while also possibly conducting analyses of their different functions in the language. The same detailed investigation would then be conducted on the invariant tags in the local varieties of English, and the results from the two sets of corpora would be compared in order to detect any parallel patterns in the substrates and local varieties of English (for an example of a detailed discussion on the different rates of the imperfect aspect in the substrates of IndE and SinE and the varieties themselves, see Sharma 2009).

the one hand and linguistic features, their frequencies and distributions on the other can only be found at the level of rather abstract linguistic configurations based on a wide range of linguistic forms”. Because this study examines only three features, many of which appear rarely in some of the corpora, it seems unlikely that they alone would meet the requirements set by Gries and Mukherjee (2010). Although it is indeed possible that the differences between the results of SinE and HKE could be caused by their different levels of nativization, due to the relatively low frequency of the features included in the study (which makes more advanced statistical testing impossible (see, for example, Gries & Mukherje 2010)), it is challenging to establish the extent to which this particular explanatory factor can be applied for the results of the present study. One way to circumvent the problem of low figures could be to combine the explanatory power of the nativization levels of the varieties with the more detailed examination of the parallel features in the substrates (discussed above). With this approach, the effects of the substrates in the local varieties of English could be compared, which could in turn provide a more comprehensive description of the dynamics between a variety’s level of nativization and the extent to which this is mirrored by the substrate effect.

As the discussion above indicates, there is some evidence that the epicentric influence of IndE could be an explanatory factor in a number of instances, especially for the results of SinE. Hundt (2013: 191, see also Methods) suggests that in cases of potential epicentric influence, one should take “the social, economic and cultural context into account and evaluate whether it fosters or hinders epicentric influence of one variety on another”, which is why the following section will present a more detailed investigation into these aspects in all three Southeast Asian varieties included in the study.

### 4.3.3 History of Indian communities in Southeast Asia

Signs of India’s cultural, linguistic and religious influence in the Southeast Asian region extend widely across both space and time (see, for example, Lamb 1975: 442; SarDesai 1997: 16; Rai 2008: 29–30), but this impact has not remained unchanged over centuries. As Rai (2008: 33) notes:

Although India’s influence on South-east Asia was considerable prior to the arrival of the Europeans, Indians in the region were largely transients, moving to and from the subcontinent to trade their wares. Permanent settlement, while evident, was exceptional, on a small scale and limited to key ports. Of those who settled prior to colonial rule, many married local women and their descendents – known as *jawi pekan*

in Malaya – were largely assimilated into the local culture. It was only after the advent of colonialism that substantial number of Indians arrived in South-east Asia.

The Indians who migrated to the region during the colonial period came from both high- and low-skilled occupational groups, and the specific migration profiles for each colony were strongly shaped by the needs of the local British colonial government(s). This colonial connection also explains why some countries such as Singapore, and, to a lesser extent, Hong Kong have larger Indian minorities than the Philippines, which as a former American colony never experienced government-supported migration from India. Due to the variation that exists between the Indian minorities of Singapore, Hong Kong and the Philippines, country-specific profiles regarding their numbers, histories, forms of employment, level of integration, and linguistic profiles are provided below.

#### 4.3.3.1 Singapore

The first Indian communities in Singapore were not formed until 1819, when Sir Stamford Raffles first arrived in the island. His fleet consisted of, among others, some 120 Indian sepoys and an unspecified number of other Indian maintenance staff, and there are indications that some Indians from the latter group decided to stay and live on the island permanently (Rai 2006: 176). Soon after the founding of the colony, Singapore was declared a free port and a growing number of Indian merchants began to move to the colony. By 1824, the number of Indians, most of whom were now military personnel or traders (Sivonen 2006: 96), had risen to over 10,000, which at the time constituted some 7% of the local population (Rai 2006: 177). During the decades that followed, another group contributing to the growth in their numbers was formed by convicts from India. The British transported them to Singapore where their labour was needed for various local construction projects and after they had served their sentence, many decided to stay in Singapore as Rai (2006: 177) notes.<sup>87</sup>

In the latter half of the 19<sup>th</sup> century, the size of the Indian population had risen to 13,000, which made it the second largest ethnic group in the colony (Turnbull 1977: 37). Around this time, their occupational profile began to diversify and while the majority of North Indians still worked for the military or as traders (Rai 2006:

---

<sup>87</sup> According to Rai (2006: 177), by 1860, the number of these convicts totalled in 2,275 individuals.

176), a growing number of South Indians were now also being employed as railroad and harbour workers (Sivonen 2006: 120). According to Rai (2008: 36), most traders moving to Singapore were Tamil Muslims, Chettiars, Gujaratis, Parsis, Bengalis and Marwaris, whereas Sikhs and Sindhis were prominent in the textile industry. Some Sikhs also came to Singapore to work as police officers, and those who were not employed by the police often found work as caretakers and guards (Yong & Major 1995: 10). At the time, Singapore was also the home of a growing plantation industry, which hired large numbers of Indians (mostly men from South India) as indentured servants or through the *kangani* and *maistry* systems (Rai 2006: 177).<sup>88</sup> These labour contracts quickly formed the main channels through which most Indians came to the region (Sivonen 2006: 120).

Interestingly, although the majority of Asian immigrants in Singapore were Chinese<sup>89</sup>, Europeans consistently favoured hiring Indians, because they were considered to be more ‘docile’ and ‘malleable’ (Rai 2008: 34). Indians were also favoured in other lines of employment by the British: according to Rai (2006: 178), both the public and private sectors in Singapore suffered from a shortage of locally educated employees in the 19<sup>th</sup> and 20<sup>th</sup> centuries and as a consequence, “the majority of workers employed in public projects, serving the budding administrative, development, commercial and defence functions in the settlement, were drawn from India” (Rai 2006: 177).<sup>90</sup> Unlike in other neighbouring countries such as Malaysia where the majority of Indians were uneducated and poor, the relatively high representation of Indians in commerce and administrative positions secured that they “were able to influence . . . the views of the colonial government vis-à-vis Indians” as Rai (2008: 36) suggests.<sup>91</sup> Despite this, Rai (2006: 179) argues that for much of Singapore’s history, the majority of Indians saw the colony as a place where they could try to make their fortune before going back to India and therefore, many considered their stay only temporary. This is understandable, since some 80% of the

---

<sup>88</sup> As criticism towards the harsh terms of indentured servitude grew, it was first replaced by the *kangani* system in 1870, which in turn was followed by the *maistry* system (Rai 2008: 33–4). Despite the new names and minor changes in the specifics of the contracts people signed, they were still basically indentured servants in all but name.

<sup>89</sup> Rai (2006: 177) notes that the Chinese were also the largest ethnic group working in the agricultural sector.

<sup>90</sup> Among these people were also many Indian doctors and teachers (Rai 2006: 178).

<sup>91</sup> During the early decades of the colony, the governmental connections between Singapore and India were even stronger, as between 1826 and 1867 the Straits Settlements, which comprised of Singapore, Melaka and Penang, were governed by the British Indian Government, thus making Singapore “effectively apart of India” (Yong & Major 1995: 4–5)

Indian minority were still uneducated manual labourers who perceived their future in the colony uncertain, whereas the middle- and upper-class white-collar workers formed only 5–7% and 0.5% of the minority respectively (Rai 2006: 180). Another significant factor contributing to the high turnover was the imbalance in the male-female ratio among the Indians (Rai 2006: 179).

The first half of the 20<sup>th</sup> century was marked by significant changes, as the great depression gave rise to tighter migration restrictions being imposed on Indians, in addition to many being repatriated (Yong & Major 1995: 11). Furthermore, many Indians returned to India due to the Japanese occupation of Singapore during the Second World War (Latif 2008: 553). The years following the war (1947–1957) were marked by a significant increase in their numbers, as some 40,000 Indians from Malaysia moved to Singapore in search of a better life (Rai 2006: 183). This trend was, however, curtailed by additional restrictions imposed on Indian migration before the start of the 1960s, after which most Indians who were allowed to immigrate to the country were family members of Singaporean Indians (Rai 2006: 183).

According to Rai (2006: 158), two distinct phases can be identified in the development of the local Indian community after Singapore gained its independence in 1965. During the first phase, from the declaration of independence to the early years of the 1990s, the Indian community became more settled<sup>92</sup> and their allegiances towards the motherland weakened and moved towards Singapore (Rai 2006: 185; Latif 2008: 560). Indians continued to be well-represented in all branches of the government and as a whole, the minority experienced steady upward mobility (Rai 2006: 185).<sup>93</sup>

The second phase, which began in the mid-1990s and has continued well into the 21<sup>st</sup> century, has been marked by significant changes in the attitudes of the Singaporean government, as it has now begun to encourage the immigration of educated workforce from India (Rai 2006: 185). The new situation was brought about by two concurrent changes that took place in India and Singapore as the century was nearing to its close. Firstly, India made significant reforms to its economic policy during the early 1990s and as a consequence, launched its new “Look East” policy, which aimed to strengthen the country’s ties with the states in

---

<sup>92</sup> One factor contributing to this development was the gradual levelling of the ratio between men and women during the first phase (Yong & Major 1995: 13; Rai 2006: 185).

<sup>93</sup> According to Rai (2006: 185), “a growing number of Indians were either in professional, technical, clerical and managerial positions, or in the sales or service industry”.

Southeast Asia. Secondly, the government of Singapore became aware of the declining birth rate of its population and to tackle the problem, decided to begin actively supporting the immigration of white-collar workers from India (and China) (Rai 1006: 186). Moreover, Rai (2006: 188) notes that since the 1990s, the Singaporean government has also “sought to establish a free trade agreement [...] alongside numerous educational and economic initiatives focused on the subcontinent”.

This latest phase has had a significant impact on the size and occupational profile of the Indian minority in Singapore. According to Shantakumar and Mukhopadhaya (2008: 573), the period has “witnessed unprecedented labour immigration from the Indian sub-continent” as unskilled Indian labour has been recruited to work on various infrastructure and construction projects, in addition to more highly educated Indians being hired for various white-collar positions.<sup>94</sup> The increase can also be seen in the size of the Indian minority, which has grown from 7.1% in 1990 to 9% in 2018 (Rai 2006: 188; Singstat 2018). Furthermore, Shantakumar and Mukhopadhaya (2008: 574) argue that the percentage of Indians in administrative, managerial, professional and technical positions has more than doubled between the 1990s and the first decade of the 21<sup>st</sup> century, growing from 18% to 47%.<sup>95</sup> Here it should be noted that the statistics do not include many of the less-educated Indians who work in the country as domestic helpers and manual labourers and for whom permanent residence status or the possibility to applying for citizenship is not offered (Rai 2006: 187). In addition, Shantakumar and Mukhopadhaya (2008: 586) note that it is uncertain whether the new skilled Indian workers will want to settle in Singapore

---

<sup>94</sup> Unfortunately it remains unclear what kind of IndE these immigrants speak. Since many of the blue-collar workers come from South Asia, some of them might use Tamil, one of the official languages of Singapore, when interacting with some communities in Singapore, whereas their English is likely to be closer to vernacular IndE. The white-collar workers in turn are likely to use a more standard form of IndE, which is closer to standard BrE (see, for example, Bhatt 2000: 74). One of the consequences of this could be that the features examined in this study might not actually be as prevalent in the speech of this subgroup of Indian migrants. However, as Lange (2012: 4) notes, “the speakers providing the data [for ICE-IND] have been chosen to represent the standard usage of their respective variety” and therefore, the results of studies using the ‘Private conversations’ section of the ICE-IND corpus “can be taken to be representative of educated spoken Indian English” (Lange 2007: 4). Since the educational background of the white-collar Indians working in Singapore can be expected to be close to the speakers of the ICE-IND data, it could be argued that there are also parallels in the kind of IndE they use, which, as the results of this study show, does contain some vernacular features of IndE.

<sup>95</sup> Even though the statistics show that the Indian minority is faring well by all criteria, Rai (2008: 47) suggests that the ‘old’, local members of the Indian community are not as well off as the expatriate Indians who have moved to the country more recently.

permanently and therefore, it is difficult to estimate how significant a role they will play in the future of the Indian community in particular and the Singaporean society as a whole.

The diversity of the Indian minority is also reflected in the language profile of the community. As was mentioned in section 2.2.4, each of the three major ethnic groups in Singapore is assigned with an official L1, which for Indians is Tamil. However, Rai (2008: 45) argues that since this was (and still is) not the actual L1 of many Singaporean Indians, many families who spoke other Indian languages such as Hindi, Urdu, Gujarati or Bengali rejected Tamil as the second language their children learned at school and chose Malay, the old regional *lingua franca*, instead. Unfortunately, many struggled to learn the language properly (Rai 2008: 45) and to solve the problem, the Singaporean government recognised also Hindi, Urdu, Gujarati, Punjabi and Bengali in the school curriculum in 1990 (Rai 2008: 46). Although the popularity of all of these additional languages has grown since, the number of students choosing Hindi exceeds those of all other non-Tamil Indian languages, which Rai (2006: 187–8) claims to be an indication of the transient nature of the most recent wave of Indian migration.<sup>96</sup> The strong connection between Indians and English has also survived to the present, since, according to the Singapore Census of Population (Singstat 2010), English is the language most frequently spoken at home for 42% of Indians, whereas the respective figures for the Chinese and Malay groups are 33% and 17%. Here it should also be noted that the “official” L1 of Indians, Tamil, was reported to be the language most frequently used at home by only 37% of Indian respondents (Singstat 2010), which further demonstrates the extent to which Indians have adopted English.

When the history of the Indian minority in Singapore is examined closer, it becomes clear that there are many factors that support the possibility that IndE could have influenced SinE. Even though the profile of the minority is diverse, Indians have always enjoyed good social standing in Singapore and they have been well-represented in various branches of the government, the education system and the military. Furthermore, with the exception of the decades following Singapore’s independence, the connections between India and Singapore have always been close and the continuous circulation of migrants from India has ensured that speakers of

---

<sup>96</sup> According to Rai (2006: 187–8), these parents wish their children to be competent speakers of Hindi in case they one day return to India. While this is undoubtedly true for many Indians, it could be suggested that an additional reason for some Indians to choose Hindi over other languages is the same why many Chinese families speaking non-Mandarin languages have opted for Mandarin – because the language is perceived to bring economic advantages to their children once they enter the labour market.

SinE (from all ethnic groups) have continuously been exposed to IndE. This trend has only strengthened since the 1990s with the new wave of Indian migration and the advances in technology (Rai 2006: 188), which now also enable Indians (old and new) to retain closer connections to the old motherland.

#### 4.3.3.2 Hong Kong

The early stages of Indian immigration to Hong Kong mirror those of Singapore's quite closely. Hong Kong was annexed to the British Empire in 1841s and according to Plüss (2005a: 155), Indian soldiers working for the British army were among the first groups to arrive at the new settlement, which at the time, as Tsang (2004: 16) notes, was populated by only a small group of Chinese fishermen and farmers. Soon after the arrival of the British, Hong Kong was declared a free port, and as the economy of the colony grew, a growing number of Indians working as soldiers, police officers, government clerks, traders and employees moved to the colony (Plüss 2006: 206) and even though they formed a significantly smaller proportion of the population when compared with that of Singapore<sup>97</sup>, there was still considerable diversity within the group.<sup>98</sup> Another interesting difference between the Indian minorities of Singapore and Hong Kong is that despite the relatively esteemed professions they had in both colonies, Sivonen (2006: 62) argues that in Hong Kong, Indians were frequently the targets of prejudices of both the British and the Chinese<sup>99</sup> and therefore, it appears unlikely that their social standing in the colony was as high as in Singapore.

The two major occupational groups for Indians at this time were soldiers and police officers. Because Indians were considered to be more loyal to the British, they were frequently hired from India to work for the local police forces<sup>100</sup> and by the end of the 19<sup>th</sup> century, over one third (226 persons) of the police officers working

---

<sup>97</sup> According to Vaid (cited in Plüss 2006: 106), the number of Indians in Hong Kong had not risen by more than 1100 individuals (from 362 to 1453 persons) over the course of the 19<sup>th</sup> century.

<sup>98</sup> Not only were the roots of these people in various parts of India, but they also represented many different faiths: the biggest religious groups were the Sunni and Shi'a Muslims, Sindhis, Jains and Parsees (Plüss 2005a, 2006).

<sup>99</sup> According to Sivonen (2006: 62), the Chinese often saw Indian merchants and police officers as rivals.

<sup>100</sup> One manifestation of this trust was the fact that during the early years of the colony, Indian police officers were given weapons while their Chinese colleagues were not (Plüss 2006: 206).



in Hong Kong were Indians (mostly Muslims or Sikhs from Punjab) (Pluss 2006: 206). Interestingly, despite their decades-long presence in the military and the police, many considered their stay in the colony only temporary, which Pluss (2006: 208) argues was the sum of various factors, such as the following: their contracts were usually temporary, they were not allowed to bring their families to Hong Kong, and, for some time, marriages with Chinese women were forbidden. Furthermore, the imbalance between Indian men and women in the colony was significant (a situation that lasted well into the 20<sup>th</sup> century, see Vaid, cited in Pluss 2006: 206) which meant that most Indian men wishing to marry and settle down had to return to India.

A third major occupational group among the Indians up until the mid-20<sup>th</sup> century was that of traders (Pluss 2006: 206), many of whom had prior experience of doing trade with the British in India as Sivonen (2006: 51) notes. The most Anglicized Indians, according to Plüss (2005a: 157), were the Parsees, who had established themselves as the middlemen between the foreign traders and Indian masses already in India<sup>101</sup> and consequently, this prosperous group had “began to speak English, learned Western manners, and became very loyal to the British”. They maintained this role also in Hong Kong, where their Anglicization continued further, as some Parsee families even Anglicized their family names during the first half of the 20<sup>th</sup> century (Plüss 2005b: 206).<sup>102</sup> Though the Parsees were a closed group who mostly turned to India to find spouses, Plüss (2005a: 158) notes that some Parsee men did also marry local Chinese women over the years. Another major subgroup of Indian merchants were the Sunnis and Shiites and the former also married Chinese women (Plüss 2005a: 159–160).

After the 1950s, the profile of the Indian minority in Hong Kong began to change as Hindus (and among them, especially the Sindhis) became the largest group of Indians, many of whom worked as traders (Chu 2005: 13; Plüss 2005a: 165). Interestingly, Plüss (2005a: 166) claims that Sindhi traders also frequently learned Cantonese in order to trade directly with the Chinese, even though this did not result in greater integration between the two groups. During this time, the occupational profile of the Indian minority also began to diversify and by the turn of the

---

<sup>101</sup> According to Plüss (2005b: 204), the “Parsis held a subordinate position in Indian Society” and thus, it could be argued that trading with Europeans was seen as means to achieve upward social mobility. Furthermore, Plüss (2005b: 208) suggests that their good English skills and friendly terms with the British authorities made it easier for them to enter into more respected professions such as doctors and teachers.

<sup>102</sup> They also had the reputation of philanthropists: for example, the donation of a wealthy Indian Parsee initiated the establishing of the first university of Hong Kong (Pluss 2005b: 206).

millennium, many worked as bankers, investors, doctors, academics, office workers, journalists, traders, security guards, restaurant owners and tailors (Plüss 2006: 206). By this time, the imbalance between the genders also levelled out (Census and Statistics Department 2018), thus enabling the Indian community to become more settled than before.

The number of Indians in Hong Kong has always been marginal and although their absolute numbers have doubled since 2001, they still formed only 0.5% of the total population in 2016 (Census and Statistics Department 2016a).<sup>103</sup> According to Singhvi et al. (2001: 277), there were also some 22,000 non-residential Indians living in Hong Kong at the turn of the millennium who are not included in the official statistics<sup>104</sup>, but even with this addition, their total percentages of the total population remain marginal. However, it should be noted that the significance of a minority to the wider community is not solely determined by its size; for example, Keezhangatte (2008: 212; see also Plüss 2006: 207) points out that at the end of the 20<sup>th</sup> century when Indians formed less than 0.3% of the population, they were actually responsible for nearly 10% of Hong Kong's exports and thus, it could be argued that their importance to the society could be higher than what their numbers suggest.

The Indian minority forms a distinct group from the Cantonese-speaking majority also by its language profile. According to the 2016 Population By-census (Census and Statistics Department 2016b), less than 10% of Indians reported Chinese (any dialect) as their most usually spoken language, while 54% of Indians selected 'Other', which in most cases likely refers to native Indian languages, as they are not specified in the questionnaire. It should also be noted that some 38% of Indians reported English as their most frequently used language, while for another 51% of Indians it was the second most frequently used language (Census and Statistics Department 2016b). The census also reveals that only approximately 35% of Indians reported Chinese (any dialect) as their second most used language, which further supports the notion of English as one of the major languages spoken by Hong Kong Indians.<sup>105</sup>

---

<sup>103</sup> According to the Census and Statistics Department (2011, 2016b), the number of Indians in Hong Kong has risen from 18,543 in 2001 to 36,362 in 2016.

<sup>104</sup> Unfortunately, the publication does not provide exact details of the year(s), which these estimates are based on. As for the type of English they might speak, the same could be assumed as in the case of Singapore (see footnote 94).

<sup>105</sup> The importance of English for the Indian minority is also reflected in their education: after the handover, the majority of schools in Hong Kong are now using Chinese as the language of instruction,

The question whether IndE could have influenced HKE is more difficult to answer. There are some key factors that do not support this argument, such as the low numbers of Indians living in Hong Kong and the fact that their relations with the ethnic Chinese have been somewhat strained in the past. Nevertheless, there are also some factors that might support the idea of IndE influence, since throughout Hong Kong's history, Indians have worked in relatively esteemed and publicly visible professions and many (but not all – see, for example, Plüss 2005b: 214) are competent users of English. Furthermore, since the Indian minority in Hong Kong has always retained its connection with the motherland, they have not lost contact with the variety of English spoken there.

#### 4.3.3.3 The Philippines

Archaeological and anthropological evidence suggest that contacts between India and the Philippine archipelago can be traced back to 300 to 200 BCE (Sridharan & Sevea 2006: 198). Signs of this connection can be found in the historical artefacts and local folktales in addition to the local languages themselves, as traces of Sanskrit and Tamil influence can be detected in the vocabularies and ancient writing systems of some Filipino languages (Scott 1994: 129, 196, 213; SarDesai 1997: 77; Salazar 2008: 500, 522).<sup>106</sup> Despite this, the first Indians known to permanently live on the islands arrived only after the British attacked Manila in 1762 during the Seven Years' War, briefly occupying the city. Among the crew were some 600 Indian soldiers and workers, most of whom decided to stay in the country when the occupation came to an end two years later (Sridharan & Sevea 2006: 198), but after over 200 years of intermarrying with the local Filipinos, the descendants of this group have now fully integrated into the surrounding population (Rye 1993: 713).

Indian migration to the Philippines did not recommence until the US colonized the islands at the turn of the 20<sup>th</sup> century, but unlike in many other Asian countries where the majority of Indians came as indentured servants, most Indians who

---

which has caused problems for the Indian children for whom the medium of instruction has traditionally been English (Plüss 2006: 209).

<sup>106</sup> There is some disagreement as to the method in which these influences arrived on the islands, some arguing for the direct contact between the Filipinos and Tamil traders (Jha 2009: 28; Salazar 2008: 522) whereas others such as Sridharan and Sevea (2006:198) claim that “[w]hatever little Indian influence could be discerned in ancient Philippines was of an indirect kind, possibly from the Hinduised kingdoms of Southeast Asia”.

immigrated to the Philippines did so as free individuals (Rye 1993: 729).<sup>107</sup> The first Indians to arrive were the Sindhis, who had a long tradition of working in trade (Salazar 2008: 501).<sup>108</sup> Interestingly, the Sindhis have to this day actively maintained a separate identity from the rest of the Indians in the Philippines as they usually do not marry outside their own group, they frequently employ their relatives in their businesses and they have even begun using English as their mother tongue (Salazar 2008: 501, 521; Sridharan & Sevea 2006: 199).

The Sindhis were soon followed by the Punjabis, many of whom had previously worked in the Malay and Hong Kong police forces under the British administration, while others had simply bought a ticket to Hong Kong in search of a better life and only upon their arrival learned of the opportunities that the Philippines had to offer (Sridharan & Sevea 2006: 198–9; Rye 1993: 715). Today, the Punjabis have outnumbered the Sindhis and are, as Salazar (2008: 501) notes, “found spread in almost every town and city nationwide”.<sup>109</sup> They too have carved an economic niche for themselves as the sellers of instalment-based consumer goods and as the creditors of micro-loans, which has given rise to the disparaging name of “Bombays 5–6” (Salazar 2008: 501–2).<sup>110</sup> Despite this, the Punjabis have always assimilated extensively with the Filipinos according to Sridharan and Sevea (2006: 198–9) and to the majority of Filipinos, it is the Punjabis who “represent the face of India [...] as they operate at the grass-roots or community level” (Salazar 2008: 502).

Although the Punjabis and the Sindhis form the two biggest groups within the Indian minority of the Philippines, the size and diversity of this community has begun to expand in recent decades. Since the 1990s, there has been a small but steady stream of expatriate Indians who have moved to work in the international organisations and transnational corporations operating in the Philippines. Salazar (2008: 502) estimates the size of this group to be some 3500 individuals, who will, according to Sridharan and Sevea (2006: 199), “draw and deserve more attention

---

<sup>107</sup> As opposed to being sent as prisoners, indentured servants or under the kangani or magistracy systems.

<sup>108</sup> In the Philippines, the Sindhis have traditionally worked mostly in the clothing industry (Salazar 2008: 501).

<sup>109</sup> The Sindhis, in contrast, predominantly reside in the larger cities where they live in “exclusive suburbs” (Salazar 2008: 521).

<sup>110</sup> The term ‘5-6’ comes from the 20 per cent interest that the Punjabis usually collect, so that for a loan of ₱5,000 one has to pay back ₱6,000 (Salazar 2008: 502). Although the Punjabis’ moneylending business is officially considered illegal, they do, as Salazar (2008: 510) points out, provide an important service to many Filipinos who do not have access to bank loans.

over the next few years due to the Philippine government's growing economic interests in India and President Arroyo's expressed desire to make Manila a 'Little Bangalore'".

The total number of Indians in the Philippines is difficult to estimate due to the lack of official statistics and the high number of Indians who stay in the country illegally, but Salazar (2008: 503) estimates that there are some 62,500 Indians, most of whom (approximately 80%) reside in the Metro Manila area. This is a clear increase from the estimations presented at the turn of the 21<sup>st</sup> century by Singhvi et al. (2001: 263), when only 38,000 Indians were believed to live in the country. Despite the increase, they are still a very small minority, forming only approximately 0.07% of the total population.<sup>111</sup> The legal status of the majority of Indians living in the Philippines is still that of immigrants who have permanent residence status and who still hold Indian citizenship and passports, while many of those who have entered the country since the 1990s have the status of temporary residents (Rye 1993: 718).<sup>112</sup> In fact, only some 7% of the Indians living in the metro Manila area have a Filipino citizenship (Salazar 2008: 521) and thus, it could be argued that the identities and allegiances of Indians living in the Philippines are likely to be far from simple.

In the past, the public image of the Indian communities in the Philippines has suffered to some extent from the disparaging attitudes of the Filipinos towards the moneylenders. In addition, Indians have for years voiced their concerns over their difficulties to obtain visas to enter the Philippines (Singhvi et al. 2001: 263), which Salazar (2008: 520) suggests is caused by the government's concern over the possible threats the Indian money lenders pose to the banking sector of the country. There are, however, indications that the attitudes of the Philippine government towards Indians might be changing. Since the turn of the century, some efforts have been made to facilitate the issuance of visas (Salazar 2008: 519–20), which might suggest that the small but steady influx of highly trained Indian work force has opened the eyes of the government to the economic gains that these expatriate Indians could bring to the country.

Although there are no statistics available on the languages used by the Indians living in the Philippines, a general picture of the situation is sketched by Salazar (2008: 521) who notes that "when a Sindhi interacts with a Punjabi, they normally

---

<sup>111</sup> The calculation presented here is based on the figures presented by Salazar (2008: 503) and The Philippines in Figures report (Philippine Statistics Authority 2016).

<sup>112</sup> Only some three to four thousand Indians have a Filipino citizenship (Rye 1993: 718; Salazar 2006: 521).

speak Filipino or Tagalog to each other and not Hindi.” The ability of the younger generations to speak both Tagalog and English is noted by Salazar (2008: 521), whereas the Sindhis, who can also speak Tagalog, actually speak English among themselves (Salazar 2008: 501, 521; Sridharan & Sevea 2006: 199) Unfortunately there are no statistics on the languages used by the white collar Indians who have come to work in the Philippines since the 1990s, but considering their level of education and place of employment, it is probable that they interact with the Philippine society using English.

Based on the results (A1–A4) and the history of Indians in the Philippine islands, the likelihood of IndE having influenced PhiE is weak. Nevertheless, since there is evidence that Indian influences have entered the Philippine islands through many (indirect) paths in the past, it cannot be indubitably argued that the prominent use of some linguistic features in IndE and some Southeast Asian Englishes could not have contributed to the likelihood in which speakers of PhiE consider these same features acceptable in their own variety, be their origin local or not. Despite this, the present study maintains that because the number and social standing of Indians in the Philippine society is not strong, it is unlikely that they could have had any significant direct impact on the development of the local variety of English.

#### 4.3.4 Summary of the answer to research question 3

As the discussion in the above sections shows, providing an answer to the third research question “What could explain the differences between the varieties?” is far from simple. While superstrate influence is an unlikely explanation for the existence of all three features in all four Asian varieties (see section 4.3.1), in many cases, substrate influence provides a plausible explanation for the results presented in A1–A4, even though the differences between the results for, for example, SinE and HKE seem to warrant additional explanatory factors. Furthermore, while HKE functions as a point of reference for the results of SinE, the results of A4 imply that substrate influence might not be able to explain the results for HKE fully either. Since the influence of IndE could help explain some of the results, the connections Singapore, Hong Kong and the Philippines have with India (and Indians) were examined in greater detail, following Hundt’s (2013: 191) suggestion. Table 5 below summarises some of the key features of this investigation.

	Singapore	Hong Kong	The Philippines
Percentage of Indians in the total population	9%	0.5%	0.07%*
Social standing	Good	Moderate	Complicated
Main language of interaction with other ethnicities	English	English	Filipino and English*
Connections with India (migration, culture, education, economy)	Strong and continuous	Moderate and continuous	Weak, not continuous

**Table 5.** Summary of key aspects related to the Indian minorities in Singapore, Hong Kong and the Philippines (\*estimates, no official statistics available on the numbers and languages spoken by the Indian minority)

Based on this information, it can be argued that SinE could indeed have been influenced by IndE. The two countries have maintained stable connections ever since English was introduced to the region, there has been a steady flow of Indian migrants to the island and due to their English skills, they have always been well-represented in professions that have high social prestige. Furthermore, the connections between the states of India and Singapore are close and they have sought to strengthen their collaboration in various fields such as security, trade, education and culture since the 1990s. Because of these multiple facilitating factors, it can be argued that the likelihood of IndE extending some direct influence over SinE is relatively high, which could also help explain some of the results of SinE in articles A1–A3.

For HKE, the situation is more complex. Even though the presence of an Indian minority extends to the early years of the colony when they worked in relatively prestigious professions, Indians in Hong Kong have never been able to establish themselves the way they have in Singapore. Some of the factors contributing to this include their low numbers and the fact that unlike Singapore and its more pluralistic society, the population of Hong Kong has been predominantly Chinese throughout its history. Therefore, even though much of the Indian minority in Hong Kong has retained its direct connections with India, their numbers are too low for them to have likely had any significant impact on the development of HKE. This is interesting, since IndE influence could actually help explain some of the results for HKE in A4. There is one possible path of influence that has not been explored further yet. As Hundt (2013: 189) notes, “[i]nfluence of an epicentre on a spatially

non-adjacent variety of English is...more like the spread of change from one urban centre to another” (see also Trudgill 1974; Milroy & Milroy 1985); both Hong Kong and Singapore are one of the largest metropolitan areas in the region and therefore, it is possible that Singapore could have contributed to the spread of IndE features to other urban centres in Southeast Asia such as HKE – a phenomenon which Vennemann (2002) describes as “the transitivity of language contact”. Unfortunately, in order to establish this, the SinE results from A1 should be examined further using the same apparent-time method employed in A4, which due to the lack of availability of the ICE-SIN metadata is not possible. Interestingly, some signs of SinE’s influence over HKE have been found by Heller et al. (2017: 133), whose study on the genitive alternation in Asian Englishes shows that SinE has slightly higher predictive accuracy over HKE than vice versa. However, since similar predictive accuracy could not be established for SinE over PhiE, Heller et al. (2017: 137–8) do not consider it plausible that SinE could function as an epicentre for the other Southeast Asian Englishes. Therefore, even though there is some evidence to support the argument that SinE could be spreading IndE features to some Southeast Asian varieties, because research on this is still extremely scarce, the present study will refrain from exploring this line of explanation further.

The connections between PhiE and IndE are significantly weaker when compared to the other two varieties included in the study: the size of the Indian minority in the Philippines is marginal and the attitudes of both the local population and the state towards Indians have been more sceptical. In addition, since the Philippines was never part of the British Empire, it lacks any shared colonial history with India, thus setting it apart from Singapore and Hong Kong. To put it briefly, none of the aspects examined here support the notion of IndE having any direct influence over PhiE, and this argument is also supported by the results in A1–A4.

As mentioned in section 4.3.1, the possibility that issues related to disrupted transmission from the superstrates could have caused the emergence of the use of clause-final *also* and *only* and invariant tag *isn’t it* in (some of) the four Asian varieties independently should be reassessed now after explanations related to the substrates have been analysed. As the above discussion shows, the combination of superstrate (Type A) and substrate (Type B) explanations could be enough to explain the results presented in this study and therefore, explanations related to acquisitional universals (Type C) need not be considered further (see Sharma 2012: 224). Despite this, it should be noted that even though the evidence presented here does indeed lend support to the notion that IndE could extend its influence to some of the Englishes spoken in Southeast Asia, this cannot be unequivocally proven to be the case. This



in turn gives rise to the difficult question whether Type C answers should be explored further in a situation where Type A and B related answers *could* be enough to explain the differences, but their explanatory power cannot be established beyond doubt – a question that cannot unfortunately be answered comprehensively in the framework of the present study. Furthermore, since the aim of this study is to see if there is proof of IndE’s possible influence in Southeast Asia, it could be argued that the question of acquisitional universals does not actually need to be explored further here, as it would exceed the scope and focus of the present study. This does not, of course, mean that the issue should not be examined further in any future studies related to this topic.

## 4.4 Research question 4

The fourth and final research question asks what the future role of IndE will be in Southeast Asia. As the discussion in section 4.3 above shows, India’s influence in Southeast Asia extends through much of the region’s history, but when the focus is narrowed down to the role IndE has played in the region, a more complex picture emerges: even if the type of English spoken in India during colonial times had an effect on the other emerging new varieties of English that developed in the other British colonies (see, for example, Mesthrie & Bhatt 2008: 86), this should be considered as pre-epicentric influence (see Peters 2009: 121–2 discussed in section 2.3), as IndE was not developed enough to function as an epicentre at the time. The interesting question is then, what has been the role of IndE in the region since the 1960s when the variety became endonormatively stabilised (Mukherjee 2007: 168) and thus finally met the basic requirements of a linguistic epicentre?

Despite the lack of sufficient diachronic data that would enable a detailed examination of the changes that have taken place in the use of the three syntactic features in IndE, SinE, HKE and PhiE, the apparent-time method used in A4 provides a means to circumvent this problem to some degree. Since the apparent-time method has traditionally been used on established native varieties of English, there might be a justifiable cause to question the applicability of the method when using some of the ICE-corpora which contain data from both native and non-native speakers of New Englishes. There is, however, already a small body of work that has used this method with the New Englishes represented in the ICE-corpora (see, for example, Fuchs & Gut 2015; Hansen 2017, 2018). Furthermore, a study by Huber (2014) which focuses on the sociolinguistic variation of *t*-affrication and relativizer

choice in Ghanaian English (GhE) – a New English variety that has reached the nativization phase on Schneider’s (2003, 2007) Dynamic Model – shows that the two features already show “robust stylistic and gender-related variation on the phonological and morphosyntactic levels of GhE” (p. 103). This leads Huber (2014: 104) to suggest the following: “[i]t appears that sociolinguistic variation, hitherto tacitly assumed to be characteristic of full-fledged varieties only, has to be expected from early on in the development of a New English, possibly even from the very beginning.” Even though the data for Huber’s (2014) study does not come from the ICE-corpora, GhE is a nativized variety (like the Asian varieties included in this study), and the results of the study suggest that the sociolinguistic patterning detected in native speaker data can also be found in New Englishes. Thus, there seems to be some evidence to support the argument that the apparent-time method can indeed be used when studying New Englishes with the ICE-corpora.

As the results of A4 indicate, the use of clause-final *also* and *only* is more frequent among older/male speakers in IndE, whereas in HKE and PhiE, they are used more by younger/female speakers and therefore, the feature appears to be more established in IndE than in HKE and PhiE. This argument is strengthened even further by the fact that although the majority of data for ICE-IND, ICE-HK and ICE-PHI was collected during the 1990s, some data for the two Southeast Asian varieties was collected later, at the beginning of the 21<sup>st</sup> century (see also 3.1.1 and A4). This means that some of the informants in the ICE-IND data who, for example, belong to the 26–35 age group could have actually been born in the same year(s) as people belonging to the 36–50 age groups in the ICE-HK and ICE-PHI data. While this detail does not have any effect on the results when comparing the frequencies of use between different speaker groups *within* a single corpus, the slight asynchronous nature of the age groups *between* the three corpora actually strengthens the argument that the feature is less established in the two Southeast Asian varieties, as it shows that the feature was more established in IndE already earlier.<sup>113</sup> This in turn lends support to the notion that the feature emerged in IndE first, from where it could have spread to other varieties over time; However, since the frequencies of use in PhiE are relatively low when compared with the other varieties, in addition to

---

<sup>113</sup> In the ICE-HKE data, for example, the use of clause-final *also* and *only* is restricted to the two youngest age groups, and for both groups the normalised frequencies are significantly lower when compared with those found in the ICE-IND (see Appendix 1 in A4). Moreover, this is the case even if the results of the HKE 26–35 age group are compared with the IndE 14–25 age group with which there might be some overlap regarding the years of birth of the informants. This shows that the use of the feature is indeed a fairly recent trend in the ICE-HK data.

the fact that a similar syntactic pattern can be found in the Philippine substrate, the notion of IndE extending its epicentric influence to PhiE is unlikely, and this conclusion is also supported by the discussion in the previous section (5.3.3–4). As a consequence, the rest of the present section will focus on discussing the two remaining Southeast Asian varieties, first describing the situation at the end of the 20<sup>th</sup> century when the data for both HKE and SinE was collected, and following this by a brief discussion on the role(s) IndE could have in the region in the future.

As the discussion in the present study and A4 shows, the results of HKE cannot be readily explained with substrate influence alone. Interestingly, the timeline for the emerging use of clause-final focus particles in HKE roughly matches the rising role of India and the growing movement of people in the region, which lends support to the argument that IndE might have contributed to the rise of the feature, even though direct connections between the varieties are not strong and further proof would be needed to support this argument. The same trends could also be argued to apply to SinE, but since the variety could not be included in the research reported in A4, it is impossible to say how established the use of clause-final *also* and *only* was in the variety at the time the data for ICE-SIN was collected. Therefore, it is also not possible to argue with certainty whether the use of the feature in SinE pre- or postdates that of IndE. Despite this, some conclusions can be drawn from the results presented in A1–A3: because the frequencies of SinE for all three syntactic features studied follow those of IndE the closest (when compared with the other Southeast Asian varieties), the results lend support to the argument that the influence of IndE extends to SinE. While the role of, for example, Indian teachers has been recognised to have influenced the development of SinE in the past (Platt et al. 1983: 8; Deterding 2007: 2; see also Mesthrie & Bhatt 2008: 19 and references therein), the recent trends in the inflow of educated IndE speakers entering the Singaporean work force means that speakers of SinE are now being exposed to IndE on a broader spectrum than ever before. Therefore, the different rates of exposure of Singaporeans and Hong Kongers have had to IndE could also explain the differences between the results of the two varieties.

Here it is worth reminding that varieties are more receptive to external influences during the early stages of variety formation, (see also section 2.1.3) and therefore, SinE, which is approaching the final stage of the Dynamic Model (Schneider 2003, 2007), is not likely to be as susceptible to the influence of IndE brought by the newest wave of Indian migration as it was when the variety was still in the early stages of its development. Despite this, the growing numbers of these new Indian migrants could be argued to function as proof of the continuous strong presence that Indian

culture and languages have traditionally had in Singapore (especially when compared with Hong Kong).<sup>114</sup> While this does not necessarily mean that the influence of IndE is as significant today as it was during the early years of SinE, some influence of IndE might still be transmitted by the second wave Indian immigrants.

In addition to the growing movement of people, another major factor contributing to IndE's growing influence in the Southeast Asian region is the rapid developments in the IT sector that have enabled global channels of communication to be established at an ever-growing pace. As a result, not only have the past few decades enabled ethnic Indians living in Southeast Asia to retain stronger connections with India (Rai 2006: 188), but also the occasions where other Southeast Asians are likely to interact with Indian nationals have diversified. An interesting example of the consequences of the growing interaction between speakers of IndE and other varieties is presented by Cameron (2012: 358), who suggests that the increased business email correspondence between speakers of IndE and AmE has resulted in the latter adopting some features from the former.<sup>115</sup> Indeed, if such intensified interaction can result in some IndE features being adopted by a particular occupational group speaking AmE, the "hub" of the Worlds System of Englishes" (Mair 2013), it is reasonable to argue that features from IndE could also have spread to some Southeast Asian varieties, as contacts between these groups have been more established for a longer period of time.

While the results of A1–A4 support the notion that IndE's epicentric influence could indeed extend to SinE and possibly, though less likely, to HKE, presenting estimations on the future of this trend can be only tentative. Because India has continued on the trajectory set by its reforms of the 1990s, it is unlikely that the country's importance in the region will significantly diminish in the near future. In addition, since Singapore's standing in the region also shows no signs of changing drastically, there does not seem to be any imminent reason why the numerous points of connection Singapore has with India and the social standing of Singaporean

---

<sup>114</sup> For example, as Mair's (2013) World System of Standard and Non-Standard Englishes shows, AmE dominates the global linguistic landscape to such a degree that the variety now extends its influence over its old superstrate, BrE, even though speakers of AmE are not a significant minority in Britain. While the presence of IndE in Singapore's cultural and linguistic landscape is of course not as strong as that of AmE in Britain, the dynamics of AmE and BrE show that a variety's potential to influence other varieties does not always require the presence of large numbers of speakers in the community, even in cases where the receiving variety is an established one.

<sup>115</sup> Cameron (2012: 358) mentions the use of the verb *revert*, which in IndE is used with the meaning 'replying to a message', having spread to American business English.

Indians would begin to wane in the foreseeable future. As a consequence, IndE's influence on SinE could be expected to, if not grow, then at least retain its current level.

The question of the future of IndE's future influence on HKE is more difficult to predict because of two reasons. Firstly, as the discussion presented here shows, IndE's influence over HKE could not be established with certainty, and therefore, it remains unclear what the relations between the two varieties is at the moment. The second problem is caused by the current political climate in Hong Kong, which makes it challenging to present any estimations of what the future might hold for Hong Kong and its local variety of English. While the PRC will undoubtedly continue its efforts to support the spread of Putonghua at the expense of Cantonese and English, the backlash among Hong Kongers against the tightening grip of the PRC might also result in a situation where (especially younger) people will connect their Hong Kong identity more strongly with both Cantonese and the local variety of English. Therefore, it is possible that HKE could turn into a symbol of the special status and history of the area, while also providing an important connection with the rest of the world. However, even if this were to happen, it is unlikely that IndE's direct influence on HKE would grow in the near future, as strong direct links between the varieties are missing at present.<sup>116</sup>

---

<sup>116</sup> Indeed, the only scenario where IndE's influence could be expected to grow more broadly among Southeast Asian varieties is that IndE would establish itself as a significant substrate for a regional standard, but as there are no clear indications of such a development taking place in the region at present, this issue will not be speculated on further here.

## 5 DISCUSSION

The following sections will take a step back from the results of the for articles and discuss broader issues related to the study on emerging epicentres. The issues discussed here range from the valuations on the applicability of the different models developed to describe World Englishes to providing suggestions on the practical implications of the results of the present study and suggesting new avenues for further research.

### 5.1 Theoretical implications

When the results of this study are examined against the three theoretical models introduced in section 2.1, it becomes apparent that some of them provide a more practical framework than others – at least when they are scrutinised from the perspective of emerging linguistic epicentres.

Even though Kachru's (1985) Three Circle Model mentions the potential influence that can be passed on from Outer Circle to the Expanding Circle varieties, it does not explicitly state whether this can also happen between Outer Circle varieties, though, as mentioned in 2.1.1, it is likely that the model would "allow" this to happen.<sup>117</sup> Another problem with the model is the absence of any mention of the term epicentre, which is why the model also does not provide any definitions that would help to identify varieties that might be on the path of acquiring this role in their respective regions. Furthermore, the model does not provide any criteria that would help in determining which varieties might be more susceptible to influences from other varieties within the same circle. Therefore, while it can be concluded that the results of the present study do broadly agree with Kachru's model, this connection is created on a more ideological level, and the details on how and when IndE could have influenced SinE (or possibly HKE) discussed here are not

---

<sup>117</sup> Kachru (1985: 28) mentions the export of teachers from South Asia to other countries located in the Outer Circle, which could be assumed to imply that influences can also be passed from one Outer Circle to another according to the model.

addressed in the model. This does not, of course, lessen the groundbreaking impact that the model has had in the field of World Englishes studies – it simply means that the model offers few practical tools that could be used in a study focusing on identifying emerging linguistic epicentres.

While Schneider's (2003, 2007) Dynamic Model does not include any mention of the term epicentre either, it provides a definition of the normative orientation of all the varieties included in the study. Because of this, it was possible to firstly establish that IndE has already passed the nativization phase and therefore has developed enough to function as a linguistic epicentre and secondly, to determine which of the other varieties included in the study were still more susceptible to external influences due to being in the exonormative phase. The latter issue is not, however, this simple when examined against the results of the present study: if it is assumed that the more advanced a variety is in the dynamic model, the less likely it is to adopt features from other varieties, this could not explain why SinE, a variety that has progressed furthest in the Dynamic Model, seems to be the likeliest recipient of IndE influences. One (partial) answer to this apparent conflict between the normative orientation of the varieties and the results of the present study is to remember that Schneider's (2003, 2007) model was developed to answer questions regarding the dynamics of interaction between different speaker populations within a country and how the variety begins to detach itself from the parent variety. Because of this, it does not adapt itself equally well to questions related to the dynamics of interaction between different speaker populations in different countries and the different normative models they might offer.

Of the three models used in this study, Mair's (2013) World System of Standard and Non-Standard Englishes appears to describe the situation implied by the results of the present study the closest, as IndE is located in the second highest category (Super-central Standard varieties) of the model. While this supports the notion that IndE indeed is an emerging epicentre in the region, the model unfortunately does not help answer the question of the reach of the variety's influence. In his discussion of IndE's role, Mair (2013: 263) mentions that it could function as a model for overseas Indian communities, in addition to providing an interesting example where an IndE feature has been taken up by a specific occupational group of AmE speakers, but this strain of influence is not elaborated upon more. Because of this, it remains somewhat unclear whether Mair considers IndE's role as a Super-central standard variety to extend only to overseas Indian communities or whether it should be considered to be capable of influencing other English varieties too, as suggested by previous research (see, for example, Heller et al. 2017; Gries & Bernaisch 2016,

Koch & Bernaisch 2013, and Hundt et al. 2012) and the results of the present study. Moreover, Mair (2013: 262) mentions that the “mediated and commodified uses” of African-American Vernacular English (AAVE) have spread the influence of the variety more widely around the globe and suggests that the same could be argued for IndE. Hence, even if the discussion presented in this study suggests that there are additional elements that have contributed to the spread of IndE features in Southeast Asia, based on Mair’s note it could be argued that the model does indeed allow paths of influence to be formed between IndE and other varieties and not just between IndE and Indian diasporas around the world.

Another question concerning Mair’s (2013) model that is difficult to answer concerns the placement of SinE, HKE and PhiE in the model. While it could be safely assumed that SinE would be located higher in the model when compared with HKE and PhiE, without a more detailed discussion of the criteria for the different categories, it is difficult to say whether SinE would qualify as a Super-central standard variety, or whether it should be placed among the Central standard varieties. This could in turn help provide some insight into the ways SinE could have contributed to the spread of IndE features to other Southeast Asian varieties such as HKE.

To conclude, it can be stated that for a study focusing on identifying emerging epicentres, Kachru’s (1985) Three Circles Model provides mostly ideological support, and while Schneider’s (2003, 2007) Dynamic Model is useful for identifying the normative orientations of different varieties, Mair’s (2013) World System of Standard and Non-standard Englishes provides the best theoretical model of the three. As Peters (2009: 122) argues, “[m]utual influence among emergent regional varieties should be factored into the evolutionary model for pluricentric languages”, which is exactly what Mair’s (2013) model does, even though there is a clear need for the exact criteria determining the placement of different varieties in the model to be expressed more clearly.

## 5.2 Practical implications

What the results of the present study show is that even though native speaker standards might still dominate in many countries, a growing number of new nativized standards are emerging around the world. One of the consequences of this is that there is also an ever-increasing number of non-native to non-native English interactions where the interlocutors opt to use local nativized patterns of the language in a way that reflects their own cultural backgrounds while also serving their



(intercultural) communicative needs. As these new norms become accepted more widely, the chances of their spreading to new areas also increases – the ramifications of this for English language teaching (for instance) are evident, as it highlights the growing need to diversify the type and number of varieties learners of English are exposed to. Although native speaker norms still hold power in the curricula of many countries, it is evident that the (present and) future users of the language will be operating in a linguistic landscape where variation is the new norm and hence, being a competent user of English requires the ability to deal with multiple linguistic norms and standards simultaneously. Furthermore, as Matsuda and Friedrich (2011: 338) note, “[t]he important thing is that students understand that diversity among varieties is not only a matter of different pronunciation features, but rather a much more encompassing manifestation of cultural, linguistic and other values”. Therefore, what the discussion above and the results of this study show is that the kind of English taught in classrooms around the world should acknowledge this already existing plurality in the language. In some countries, such as Finland, this new perspective has already taken root, as the new curriculum (POPS 2014) highlights the need for students to understand the role English plays as a global lingua franca, in addition to having important intranational roles in countries where it functions as an official second language. While the purpose of this study is not to take a stand on how such goals should be implemented in classroom teaching, based on the results, a suggestion can be made on the importance of including Asian Englishes among the varieties students are exposed to. After all, Asia is home to some of the largest English-speaking populations in the world, many of whom are already shaping the language to meet their communicative needs and thus, even if English might not have become an Asian language yet, it is clearly in the process of becoming one in the near future.

### 5.3 Recommendations for future research

While the results of the present study show that IndE extends a degree of influence over some English varieties spoken in Southeast Asia, to establish that it is indeed an emerging epicentre in the region would need to be supported by further corpus, sociolinguistic and attitudinal explorations into the use, role and social standing of IndE in the region (see also Hoffmann et al. 2011: 277; Hundt 2013: 203). Therefore, even though this study has mapped the use and potential origins of three syntactic innovations in IndE, SinE, HKE and PhiE, further studies on additional features are

needed to establish the extent to which features found in IndE are also used in Southeast Asian Englishes and whether the origins of these features can be traced back to IndE or somewhere else. Furthermore, while this study has shed some light into the user demographics of a particular feature, the use of clause-final *also* and *only* in IndE, HKE and PhiE, more detailed sociolinguistic and attitudinal studies<sup>118</sup> should be conducted on the users (and non-users) of the other features included in this study.

One of the challenges faced during this study was the limitations posed by the amount of data provided in the ICE-corpora (see also section 3.2 on Methods). A possible way to circumvent this problem in future studies would be to use more recently published corpora such as GloWbe (Davies 2013), which contains 1.9 billion words of online data representing 20 different varieties of English. Another way to tackle the problems would be to continue studying different nativized features in Asian varieties with the methods used by Hoffmann et al. (2011) and Hund et al. (2012), who focused on newspaper language, a genre which has elements from both more conservative (written) and innovative (spoken) language<sup>119</sup>, and for which samples can be collected from the Internet more easily. Since both of the aforementioned solutions would enable the examination of additional features, it should be noted here that the order of varieties (as presented in Table 4) might not of course be fully reproduced in future studies even if the features in question did meet the criteria set for the features included in this study. This should not, however, be assumed to automatically disprove the hypothesis presented here, since a number of different complementary (and sometimes contradictory) explanatory factors need to be taken into consideration before any more definitive arguments can be made based on the results.<sup>120</sup> For example, if there is a feature that is reported to exist in both IndE and SinE, but which upon closer examination is shown to be more prevalent in SinE, the results should not be assumed to necessarily disprove IndE's potential influence on SinE – it could be, for example, that the feature is merely

---

<sup>118</sup> According to Heller et al (2017: 137), “[c]ross- varietal attitudinal considerations should take center stage with epicentral investigations of features which a) are used in the epicentral, but not yet in rupture- zone varieties, b) rupture-zone variety users can clearly identify as an element of the epicentral variety and c) rupture-zone variety users have an active awareness of and can recognize in discourse.”

<sup>119</sup> See, for example, Hinrichs & Szmrecsanyi (2007: 441) and references therein.

<sup>120</sup> Some of the questions that need to be answered first for each variety include, for example, the number of substrates that also have the same feature, the prominence of the feature in these substrates, the relative “weight” of the substrate(s) in the feature pool and whether the feature existed in the settler varieties.

more prominent in the substrates of SinE than in those of IndE, in which case that particular feature simply is not well suited for the study of IndE's potential influence over SinE.

Both Hoffmann et al. (2011: 261) and Gries and Bernaisch (2016: 23) show that it is possible to identify a variety as an epicentre using large-scale synchronic data, but, as Gries and Bernaisch (2016: 23) note, “diachronic data are needed to study their seismic waves”. Therefore, once diachronic data becomes available for IndE, SinE, HKE and PhiE, it would be interesting to study when these different nativized features have first emerged and then become established in the varieties. In addition, if speaker data for ICE-SIN is one day made available for the academic community, SinE could be included in any future studies using the apparent-time method that aim to provide new insights into the old ICE data. The final suggestion for future research is also conditional on new type of data becoming available, since comparable corpora on the major substrates of the Asian varieties included in this study could help shed light on the usage patterns of parallel features in the substrates. These could then be compared with those found in the ICE-corpora and this could provide more definitive answers as to the extent to which substrate and/or IndE influences can be used to explain the usage patterns found in the Asian varieties.

## 6 CONCLUSION

This study set out to examine whether IndE could function as an emerging epicentre in Southeast Asia. As Hundt (2013: 196) notes, “the question of how we can distinguish between historical developments that result in parallel or convergent developments on the one hand and epicentric influence on the other hand will remain one of the most difficult challenges on empirical research on old, new and emerging epicentres” – a view the author of the present study fully agrees with. While there is growing evidence that IndE already functions as an epicentre in South Asia, these investigations have not yet been extended to Southeast Asia. Due to the limitations regarding the availability of data, three Southeast Asian Englishes were selected for the study: SinE, HKE and PhiE. SinE and HKE have stronger historical ties with IndE as they were all once part of the British empire. The connections between PhiE and IndE, in contrast, have been significantly weaker due to their greater geographical distance and the lack of a shared history under the same colonial power. The selection of the three features were based on the following criteria. Firstly, previous studies had noted these new features to be common in IndE (and possibly in SinE, HKE and PhiE). Secondly, to ensure their “innovativeness”, a link to the substrates was required (at least for IndE) and thirdly, retrieving the features needed to be possible even when using the untagged ICE-corpora.

While substrate influence can explain some of the results for all the varieties, for some aspects, further explanatory criteria were required, and these were considered following the order presented by Sharma (2012). What the results of this study show is that IndE most likely extends some of its epicentric influence over SinE, while the question for HKE remained undetermined, though this is considered unlikely based on extralinguistic factors. In addition, the results of PhiE showed no sign of having been influenced by IndE as expected. Although signs of IndE’s pre-epicentric influence in some parts of Southeast Asia can be detected already before the end of the 19<sup>th</sup> century, the variety’s rise to its current role likely began during the latter half of the 20<sup>th</sup> century – while further studies are needed to establish the extent to which IndE functions as an epicentre for Southeast Asian Englishes as a whole, there are strong indications that IndE has already extended its influence from South Asia to at least some of the varieties spoken in Southeast Asia also.

# REFERENCES

## Corpora

- The *Freiburg – Brown Corpus of American English*. 1999. Englisches Seminar, Albert-Ludwigs Universität, Freiburg.
- The *International Corpus of English: Great Britain*. 1998. The Survey of English Usage, University College London, London.
- The *International Corpus of English: Hong Kong*. 2006. Department of English, The University of Hong Kong and The Chinese University of Hong Kong, PRC.
- The *International Corpus of English: India*. 2002. Department of English, Shivaji University, Kolhapur and Freie Universität, Berlin.
- The *International Corpus of English: the Philippines*. 2004. Department of English and Applied Linguistics, De La Salle University, Manila.
- The *International Corpus of English: Singapore*. 2002. Department of English Language & Literature, The National University of Singapore, Singapore.
- The *International Corpus of English: USA*. 2012. University of Massachusetts-Boston, Boston, University of California, Los Angeles and the University of California, Santa Barbara.
- The *Santa Barbara Corpus of Spoken American English*. 2002–2005. Santa Barbara Center for the Study of Discourse, University of California, Santa Barbara.

## Bibliography

- Agnihotri, Rama Kant. 2007. *Hindi: An essential grammar*. London and New York: Routledge.
- Alsagoff, Lubna. 2012. The development of English in Singapore. In Ee Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language use*, 137–154. Amsterdam: John Benjamins.
- Asher, R. E. 2008. Language in historical context. In Braj B. Kachru, Yamuna Kachru & Shikaripur N. Sridhar (eds.), *Language in South Asia*, 31–48. Cambridge: Cambridge University Press.
- Aye, Khin K. 2005. *Bazaar Malay: History, grammar and contact*. Dissertation. National University of Singapore.
- Bailey, Richard W. 1991. *Images of English: A cultural history of the language*. Ann Arbor: University of Michigan Press.
- Balasubramanian, Chandrika. 2009. *Register Variation in Indian English*. Amsterdam: Benjamins.
- Bao, Zhiming. 2015. *The making of vernacular Singapore English: System, transfer, and filter*. Cambridge: Cambridge University Press.
- Bao, Zhiming & Huaqing Hong. 2006. Diglossia and register variation in Singapore English. *World Englishes* 25(1), 105–114.

- Baskaran, Loga. 2004. Malaysian English: Morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.), *A Handbook of Varieties of English, Vol. 2: Morphology and syntax*, 1073–1085. Berlin: Mouton de Gruyter.
- Bautista, Maria Lourdes S. & Kingsley Bolton. (eds.). 2008. *Philippine English: Linguistic and literary perspectives*. Hong Kong: Hong Kong University Press.
- Beebe, James & Maria Beebe. 1981. The Filipinos: A special case. In Charles A. Ferguson, Shirley B. Heath & David Hwang (eds.), *Language in the USA. CUP Archive, 1981*, 322–338. Cambridge: Cambridge University Press.
- Bernaish, Tobias & Christopher Koch. 2016. Attitudes towards Englishes in India. *World Englishes* 35(1), 118–32.
- Bernardo, Allan B. I. 2004. McKinley's questionable bequest: over 100 years of English in Philippine education. *World Englishes* 23(1), 17–31.
- Bhatt, Rakesh M. 2000. Optimal expressions in Indian English. *English Language and Linguistics* 4(1), 69–95.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, Edward Finegan & Randolph Quirk. 1999. *Longman Grammar of Spoken and Written English*. London, New York: Longman.
- Biewer, Carolin, Marianne Hundt & Lena Zipp. 2010. 'How' a Fiji corpus? Challenges in the compilation of an ESL ICE component. *ICAME* 34, 5–23.
- Bolton, Kingsley. 2002. The sociolinguistics of Hong Kong and the space for Hong Kong English. In Kingsley Bolton (ed.), *Hong Kong English: Autonomy and creativity*, 29–55. Hong Kong: Hong Kong University Press.
- Bolton, Kingsley. 2003. *Chinese Englishes: A sociolinguistic history*. Cambridge: Cambridge University Press.
- Bolton, Kingsley. 2012. Language policy and planning in Hong Kong. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use*, 221–240. Amsterdam: John Benjamins.
- Bolton, Kingsley & Ma Lourdes S. Bautista. 2004. Philippine English: Tensions and transitions. *World Englishes* 23(1), 1–5.
- Bolton, Kingsley & Ma Lourdes S. Bautista. 2008. Philippine English: Linguistic and literary perspectives. In Ma Lourdes S. Bautista & Kingsley Bolton (eds.), *Philippine English: Linguistic and literary perspectives*. 1–9. Hong Kong: Hong Kong University Press.
- Cameron, Deborah. 2012. The commodification of language: English as a global commodity. In Terttu Nevalainen & Elizabeth Traugott (eds.), *The Oxford Handbook of the History of English*, 352–364. Oxford: Oxford University Press.
- Carter, Ronald. & Michael McCarthy. 2006. *Cambridge Grammar of English: A comprehensive guide; Spoken and written English grammar and usage*. Cambridge: Cambridge University Press.
- Census and Statistics Department. 2011. *Hong Kong 2011 Population Census*. Population aged 5 and older by new town and usual language. Excel document. URL <https://www.census2011.gov.hk/en/main-table.html> (Accessed 3 November 2019).
- Census and Statistics Department. 2016a. *2016 Population By-census*. Population by ethnicity and year. Excel document. URL <https://www.bycensus2016.gov.hk/en/bc-mt.html> (Accessed 2 February 2020).
- Census and Statistics Department. 2016b. *2016 Population By-census, Thematic Report: Ethnic minorities*. Proportion of ethnic minorities aged 5 and over able to speak selected languages/ dialects by ethnicity. PDF document. URL

- <https://www.bycensus2016.gov.hk/data/16bc-ethnic-minorities.pdf> (Accessed 12 November 2019).
- Census and Statistics Department. 2018. *Women and Men in Hong Kong: Key statistics*. Population by ethnicity and sex. Excel document. URL <https://www.statistics.gov.hk/pub/B11303032018AN18B0100.pdf> (Accessed 12 February 2020).
- Chu, Cindy Yik-yi. 2005. Introduction. In Cindy Yik-yi Chu (ed.), *Foreign Communities in Hong Kong, 1840s–1950s*, 1–16. New York: Palgrave Macmillan.
- Chu, Cindy Yik-yi (ed.). 2005. *Foreign Communities in Hong Kong, 1840s–1950s*. New York: Palgrave Macmillan.
- Collins, Peter. 2005. The modals and quasi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide* 26(2), 249–273.
- Collins, Peter. 2009. Modals and quasi-modals in world Englishes. *World Englishes* 28(3), 281–92.
- Columbus, Georgie. 2009. A corpus-based analysis of invariant tags in five varieties of English. In Antoinette Renouf & Andrew Kenoe (eds.), *Corpus Linguistics: Refinements and reassessments*, 401–414. Amsterdam: Rodopi.
- Columbus, Georgie. 2010. A comparative analysis of invariant tags in three varieties of English. *English World-Wide* 31(3), 288–310.
- Crystal, David. 2003a. *The Cambridge Encyclopedia of the English language, 2nd edition*. Cambridge: Cambridge University Press.
- Crystal, David. 2003b. *English as a Global Language*. Cambridge: Cambridge University Press.
- Davies, Mark. 2013. *Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries* (GloWbE). Available online at [corpus.byu.edu/glowbe/](http://corpus.byu.edu/glowbe/).
- Davydova, Julia. 2012. Englishes in the Outer and Expanding Circles: A comparative study. *World Englishes* 31(3), 366–85.
- Dayag, Danilo T. 2012. Philippine English. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use*, 91–111. Amsterdam: John Benjamins.
- Deterding, David. 2007. *Singapore English*. Edinburgh: Edinburgh University Press.
- Dixon, L. Q. 2005. Bilingual education policy in Singapore: An analysis of its sociohistorical roots and current academic outcomes. *International Journal of Bilingual Education and Bilingualism* 8(1), 25–47.
- Dolan, Ronald E. 2003. Philippines: A country study. In Harry S. Calit (ed.), *The Philippines, Current Issues and Historical Background*, 25–227. New York: Nova Science Pub Incorporated.
- Dryer, Matthew S. 2007. Word order. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol. 1: Clause structure, 2nd edition*, 61–131. Cambridge: Cambridge University Press.
- Edwards, Alison. 2016. *English in the Netherlands: Functions, forms and attitudes*. Amsterdam / Philadelphia: John Benjamins.
- Evans, Stephen. 2008. The introduction of English-language education in early colonial Hong Kong. *History of Education* 37(3), 383–408.
- Evans, Stephen. 2009. The evolution of the English-language speech community in Hong Kong. *English World-Wide* 30(3), 278–301.
- Evans, Stephen. 2011. Hong Kong English and the professional world. *World Englishes* 30(3), 293–316.

- Evans, Stephen. 2014. The evolutionary dynamics of postcolonial Englishes: A Hong Kong case study. *Journal of Sociolinguistics* 18(5), 571–603.
- Finnish National Agency for Education. 2014. *Perusopetuksen opetussuunnitelman perusteet 2014* (POPS). PDF file. URL [https://www.oph.fi/sites/default/files/documents/perusopetuksen\\_opetussuunnitelman\\_perusteet\\_2014.pdf](https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf) (Accessed 18 November 2019).
- Fuchs, Robert. 2012. Focus marking and semantic transfer in Indian English: The case of also. *English World-Wide* 33(1), 27–53.
- Fuchs, Robert. 2015. You're not from around here, are you? – A dialect discrimination experiment with speakers of British and Indian English. In Elisabeth Delais-Roussarie, Mathieu Avanzi & Sophie Herment (eds.), *Prosody and Language in Contact: L2 acquisition, attrition and languages in multilingual situations*, 123–148. Berlin: Springer.
- Fuchs, Robert. 2016. *Speech Rhythm in Varieties of English: Evidence from educated Indian English and British English*. Berlin: Springer.
- Fuchs, Robert & Ulrike Gut. 2015. An apparent-time study of the progressive in Nigerian English. In Peter Collins (ed.), *Grammatical Change in English World-Wide*, 373–387. Amsterdam: John Benjamins.
- Fuchs, Robert, Ulrike Gut & Taiwo Soneye. 2013. 'We just don't even know': The usage of the pragmatic focus particles even and still in Nigerian English. *English World-Wide*, 34(2), 123–145.
- Gargesh, Ravinder. 2006. South Asian Englishes. In Braj Kachru, Yamuna Kachru & Cecil Nelson (eds.), *The Handbook of World Englishes*, 90–113. Oxford: Blackwell.
- Gisborne, Nikolas. 2009. Aspects of the morphosyntactic typology of Hong Kong English. *English World-Wide* 30(2), 149–69.
- Gonzalez, Andrew. 2004. The social dimensions of Philippine English. *World Englishes* 23(1), 7–16.
- Graddol, David. 2010. *English Next India: The future of English in India*. London: British Council.
- Greenbaum, Sidney. 1991. ICE: The international corpus of English. *English Today* 7(04), 3–7.
- Gries, Stefan T. & Joybrato Mukherjee. 2010. Lexical gravity across varieties of English: An ICE-based study of *n*-grams in Asian Englishes. *International Journal of Corpus Linguistics* 15(4), 520–548.
- Gries, Stefan T. & Tobias Bernaisch. 2016. Exploring epicentres empirically: Focus on South Asian Englishes. *English World-Wide* 37(1), 1–25.
- Görlach, Manfred. 1990. *Studies in the History of the English Language*. Heidelberg: Carl Winter.
- Götz, Sandra. 2017. Non-canonical syntax in South Asian Varieties of English: A corpus-based pilot study on fronting. *Zeitschrift für Anglistik und Amerikanistik* 65(3), 265–281.
- Hansen, Beke. 2017. The ICE metadata and the study of Hong Kong English. *World Englishes*. 36(3), 471–486.
- Hansen, Beke. 2018. *Corpus Linguistics and Sociolinguistics: A study of variation and change in the modal systems of World Englishes*. Leiden: Brill.
- Hassan, Norizah, Azirah Hashim & Adriana S. Phillip. 2012. Language use in the construction of interpersonal relationships: Electronic English in Malaysia. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use* 42, 323–341. Amsterdam: John Benjamins.



- Heller, Benedikt, Tobias Bernaisch & Stefan T. Gries. 2017. Empirical perspectives on two potential epicenters: The genitive alternation in Asian Englishes. *ICAME Journal* 41(1), 111–44.
- Hiramoto, Mie. 2015. Sentence-final adverbs in Singapore English and Hong Kong English. *World Englishes*, 34(4), 636–653.
- Ho, Debbie G. 2008. English in the Expanding Circle—A third diaspora? *Asian Englishes* 11(1), 36–50.
- Hoffmann, Sebastian, Marianne Hundt, & Joybrato Mukherjee. 2011. Indian English an emerging epicentre? A pilot study on light verbs in web-derived corpora of South Asian Englishes. *Anglia* 129(3–4), 258–280.
- Hoffmann, Sebastian, Anne-Katrin Blass & Joybrato Mukherjee. 2014. Canonical tag questions in Asian Englishes: Forms, functions, and frequencies in Hong Kong English, Indian English, and Singapore English. In Markku Filppula, Juhani Klemola & Devyani Sharma (eds.), *The Oxford Handbook of World Englishes*, 697–714. Oxford: Oxford University Press.
- Hogue, Cavan. 2001. The spread of Anglo-Indian words into South-East Asia. In Bruce Moore (ed.), *Who's centric now? The present state of post-colonial Englishes*, 165–197. South Melbourne, Victoria: Oxford University Press.
- Huber, Magnus. 2014. Stylistic and sociolinguistic variation in Schneider's nativization phase: T-affrication and relativization in Ghanaian English. In Sarah Buschfeld, Thomas Hoffmann, Magnus Huber & Alexander Kautzsch (eds.), *The Evolution of Englishes: The Dynamic Model and beyond*, 86–106. Amsterdam: John Benjamins.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hudson-Ettle, Diana & Josef Schmied. 1999. *Manual to accompany The East African Component of The International Corpus of English (ICE-EA)*. On CD-Rom.
- Hundt, Marianne, Andrea Sand & Paul Skandera. 1999. *Manual of Information to Accompany the Freiburg-Brown Corpus of American English (Frown)*. Freiburg: University of Freiburg. PDF file. URL <http://korpus.uib.no/icame/manuals/FROWN/INDEX.HTM> (Accessed 10 January 2020)
- Hundt, Marianne. 2013. The diversification of English: Old, new and emerging epicentres. In Daniel Schreier & Marianne Hundt (eds.), *English as a Contact Language*, 182–203. Cambridge: Cambridge University Press.
- Hundt, Marianne, Sebastian Hoffmann & Joybrato Mukherjee. 2012. The hypothetical subjunctive in South Asian Englishes: Local developments in the use of a global construction. *English World-Wide* 33(2), 147–64.
- Hung, Tony T. 2000. Towards a phonology of Hong Kong English. *World Englishes* 19(3), 337–56.
- Hung, Tony T. 2012. Hong Kong English. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use*, 113–133. Amsterdam: John Benjamins.
- ICE, *The International Corpus of English, homepage*. Website. URL <https://www.ice-corpora.uzh.ch/en.html> (Accessed 20 January 2020).
- ICE, *The International Corpus of English, design of the ICE corpora*. Website. URL <https://www.ice-corpora.uzh.ch/en/design.html> (Accessed 12 February 2020).
- Ingham, Michael. 2007. *Hong Kong: A cultural history*. Oxford: Oxford University Press.
- Jha, Murari K. 2009. The social world of Gujarati merchants and their Indian Ocean networks in the seventeenth century. In Rajesh Rai & Peter Reeves (eds.), *The South*

- Asian Diaspora: *Transnational Networks and Changing Identities*, 28–44. New York: Routledge.
- Junghare, Indira Y. 1988. Topic-prominence and zero NP-anaphora in Marathi and Hindi. In Mohammad Ali Jazayery & Werner Winter (eds.), *Languages and Cultures*, 309–328. Berlin: Mouton de Gruyter.
- Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In Randolph Quirk & Henry G. Widdowson (eds.), *English in the World: Teaching and learning the language and literatures*, 11–30. Cambridge: Cambridge University Press.
- Kachru, Braj B. 1992a. The second diaspora of English. In Tim W. Machan & Charles T. Scott (eds.), *English in Its Social Contexts: Essays in historical sociolinguistics*, 230–252. New York: Oxford University Press.
- Kachru, Braj B. 1992b. Teaching world Englishes. In Braj B. Kachru (ed.), *The Other Tongue*, 355–366. Urbana and Chicago: University Illinois Press.
- Kachru, Braj B. 1998. English as an Asian language. *Links & Letters*, 5: 89–108.
- Kachru, Braj B. 2005. *Asian Englishes: Beyond the canon*. Hong Kong: Hong Kong University Press.
- Kachru, Yamuna. 2006. *Hindi*. Amsterdam and Philadelphia: John Benjamins.
- Kachru, Yamuna & Cecil L. Nelson. 2006. *World Englishes in Asian Contexts*. Hong Kong: Hong Kong University Press.
- Kachru, Yamuna & Larry E. Smith. 2008. *Cultures, contexts, and world Englishes*. New York: Routledge.
- Keezhangatte, James J. 2008. Blue-collar Indians: Imperceptible Yet Important in Hong Kong. In *Rising India and Indian Communities in East Asia*, 207–225. Singapore: ISEAS Publishing.
- Kesavapany, K., A. Mani & Palanisamy Ramasamy (eds.). 2008. *Rising India and Indian Communities in East Asia*. Singapore: ISEAS Publishing.
- Kidwai, Ayesha. 2004. The topic interpretation in universal grammar: Evidence from Kashmiri (and German). In Veneeta Dayal & Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, 253–289. Boston: Kluwer Academic Publishing.
- King, Robert D. 2001. The poisonous potency of script: Hindi and Urdu. *International Journal of the Sociology of Language* 150, 43–59.
- Kiss, Katalin, É. 1998. Identificational focus versus information focus. *Language* 74, 245–274.
- Kittilä, Seppo. 2002. *Transitivity: Towards a comprehensive typology*. Turku: University of Turku.
- Koch, Christopher & Tobias Bernaisch. 2013. Verb complementation in South Asian English(es): The range and frequency of “new” ditransitives. In Gisle Andersen & Kristin Bech (eds.), *English Corpus Linguistics: Variation in time, space and genre – Selected papers from ICAME 32*, 69–89. Amsterdam: Brill Academic Publishers.
- Koul, Omkar N. 2008. *Modern Hindi Grammar*. Hyattsville: Dunwoody Press.
- Kulke, Hermann & Dietmar Rothermund. 1986. *A History of India*. London: Croom Helm.
- König, Ekkehard. 1991. *The Meaning of Focus Particles: A comparative perspective*. London: Routledge.
- Lal, Brij V., Peter Reeves & Rajesh Rai (eds.). 2006. *The Encyclopedia of the Indian Diaspora*. Honolulu: University of Hawai'i Press.
- Lamb, Alastair. 1975. Indian Influence in ancient South-East Asia. In Arthur L. Basham (ed.), *A Cultural History of India*, 442–454. Oxford: Clarendon Press.

- Lange, Claudia. 2007. Focus marking in Indian English. *English World-Wide* 28(1), 89–118.
- Lange, Claudia. 2012. *The Syntax of Spoken Indian English*. Amsterdam: John Benjamins.
- Latif, Asad-ul I. 2008. From mandalas to microchips: The Indian imprint on the construction of Singapore. In K. Kesavapany, A. Mani & P. Ramasamy (eds.), *Rising India and Indian Communities in East Asia*, 549–567. Singapore: ISEAS Publishing.
- Leech, Geoffrey & Jan Svartvik. 2002. *A Communicative Grammar of English*. London: Longman.
- Leitner, Gerhard. 1992. English as a pluricentric language. In Michael Clyne (ed.), *Pluricentric Languages: Differing norms in different nations*, 179–237. Berlin and New York: Mouton de Gruyter.
- Leppänen, Sirpa, Anne Pitkänen-Huhta, Tarja Nikula, Samu Kytölä, Timo Törmäkangas, Kari Nissinen, Leila Kääntä, Tiina Virkkula, Mikko Laitinen & Päivi Pahta. 2009. Kansallinen kyselytutkimus englannin kielestä Suomessa: Käyttö, merkitys ja asenteet. *Jyväskylän Studies in Humanities* (132).
- Lewis, M. P., Gary F. Simons & Fenning, Charles D (eds.). 2015, *Ethnologue: Languages of the World, 18<sup>th</sup> edition*. Dallas, Texas: SIL International. Website. <http://www.ethnologue.com>. (Accessed 1 May 2015).
- Lim, Lisa. 2007. Mergers and acquisitions: On the ages and origins of Singapore English particles. *World Englishes* 26(4), 446–73.
- Lim, Lisa & A. Foley Joseph. 2004. English in Singapore and Singapore English: Background and methodology. *Singapore English: A Grammatical Description*, 1–18.
- Lim, Lisa & Umberto Ansaldi. 2012. Contact in the Asian arena. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford Handbook of the History of English*, 560–571. Oxford: Oxford University Press.
- Lin, Hua. 2001. *A Grammar of Mandarin Chinese*. Muenchen: Lincom Europa.
- Low, Ee-Ling. 2012. Singapore English. In Ee Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy & language in use*, 35–54. Amsterdam: John Benjamins.
- Low, Ee-Ling & Azirah Hashim. (eds.). 2012. *English in Southeast Asia: Features, policy and language in use*. Amsterdam: John Benjamins.
- 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.
- Martin, Isabel P. 2012. Diffusion and directions: English language policy in the Philippines. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use*, 189–205. Amsterdam: John Benjamins.
- Matsuda, Aya & Patricia Friedrich. 2011. English as an international language: A curriculum blueprint. *World Englishes* 30(3), 332–44.
- Matthews, Stephen & Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.
- McArthur, Tom. 1987. The English languages? *English Today* 3(3), 9–13.
- McFarland, Curtis D. 2004. The Philippine language situation. *World Englishes* 23(1), 59–75.
- McLeod, John. 2002. *The History of India*. Westport, CT: Greenwood Publishing Group.
- McNeill, J. R. & William H. McNeill. 2003. *The Human Web: A bird's-eye view of world history*. New York: W.W. Norton.
- Mesthrie, Rajend & Rakesh M. Bhatt. 2008. *World Englishes: The study of new linguistic varieties*. Cambridge: Cambridge University Press.
- Miller, Stuart C. 1984. *Benevolent Assimilation: The American conquest of the Philippines, 1899-1903*. New Haven and London: Yale University Press.
- Milner, Anthony. 2008. *The Malays*. Chichester: Wiley-Blackwell.

- Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2), 339–84.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Mukherjee, Joybrato. 2007. Steady states in the evolution of new Englishes: Present-day Indian English as an equilibrium. *Journal of English Linguistics* 35(2), 157–87.
- Mukherjee, Joybrato, Marco Schilk, Tobias Bernaisch & Justus Liebig. 2010. Compiling the Sri Lankan component of ICE: Principles, problems, prospects. *ICAME Journal* 34, 64–77.
- Nevalainen, Terttu. 1991. *But, Only, Just: Focusing adverbial change in modern English 1500–1900*. Helsinki: Société néophilologique.
- Nevalainen Terttu & Elizabeth Closs Traugott (eds.). 2012. *The Oxford Handbook of the History of English*. Oxford: Oxford University Press.
- OED, *The Oxford English Dictionary*. Website. URL <https://www.oed.com/> (Accessed 3 June 2018).
- Parviainen, Hanna. 2012. Focus particles in Indian English and other varieties. *World Englishes* 31(2), 226–47.
- Parviainen, Hanna. 2016. The invariant tag isn't it in Asian Englishes. *World Englishes* 35(1), 98–117.
- Parviainen, Hanna. 2017. Omission of direct objects in New Englishes. In Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds.), *Changing English: Global and Local Perspectives* 92, 129–153. Berlin, Boston: Mouton de Gruyter.
- Parviainen, Hanna & Robert Fuchs. 2018. 'I don't get time only': An apparent-time investigation of clause-final focus particles in Asian Englishes. *Asian Englishes*, 1–20.
- Peters, Pam. 2009. Australian English as a regional epicenter. *World Englishes—problems, properties and prospects: Selected papers from the 13th LAWE conference* 40, 107–24.
- Philippine Statistics Authority. 2014. *Philippines in Figures*. PDF document. URL [https://psa.gov.ph/sites/default/files/2014%20PIF\\_0.pdf](https://psa.gov.ph/sites/default/files/2014%20PIF_0.pdf) (Accessed 5 May 2018).
- Platt, John T. 1982. English in Singapore, Malaysia and Hong Kong. In Richard R. Bailey & Manfred Görlach (eds.), *English as a World Language*, 384–414. Cambridge: Cambridge University Press.
- Platt, John, Heidi Weber & Mian Lian Ho. 1983. *Singapore and Malaysia*. Amsterdam: John Benjamins.
- Platt, John, Heidi Weber & Mian Lian Ho. 1984. *The New Englishes*. London: Routledge & Kegan Paul.
- Plüss, Caroline. 2005a. Constructing globalized ethnicity: Migrants from India in Hong Kong. *International Sociology* 20(2), 201–24.
- Plüss, Caroline. 2005b. Migrants from India and their relations with British and Chinese residents. In Cindy Yik-yi Chu (ed.), *Foreign Communities in Hong Kong, 1840s–1950s*, 155–170. New York: Palgrave Macmillan.
- Plüss, Caroline. 2006. Hong Kong. In Brij V. Lal, Peter Reeves & Rajesh Rai (eds.), *The Encyclopedia of the Indian Diaspora*, 206–209. Honolulu: University of Hawai'i Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rai, Rajesh. 2006. Singapore. In Brij V. Lal, Peter Reeves & Rajesh Rai (eds.), *The Encyclopedia of the Indian Diaspora*. 176–188. Honolulu: University of Hawai'i Press.

- Rai, Rajesh. 2008. 'Positioning' the Indian diaspora: The South-east Asian experience. In Parvati Raghuram, Ajaya K. Sahoo, Brij Maharaj & Dave Sangha (eds.), *Tracing an Indian Diaspora: Contexts, memories, representations*, 29–51. New Delhi: Sage.
- Ramos, Teresita V. 1971. *Tagalog Structures*. Honolulu: University of Hawaii Press.
- Ramos, Teresita V. & Resty M. Cena. 1990. *Modern Tagalog: Grammatical Explanations and Exercises for Non-native Speakers*. Honolulu: University of Hawaii Press.
- Rautionaho, Paula. 2014. *Variation in the Progressive: A corpus-based study into world Englishes*. University of Tampere dissertation.
- Reeves, Peter & Rajesh Rai. 2006. The Indian context. In Lal, Brij V., Peter Reeves & Rajesh Rai (eds.), *The Encyclopedia of the Indian Diaspora*, 18–31. Honolulu: University of Hawai'i Press.
- Registrar General and Census Commissioner. 2011. *2011 Census*. Language and mother tongue. PDF document. URL <http://censusindia.gov.in/2011Census/Language-2011/Statement-1.pdf> (Accessed 12 February 2020).
- Ross, Claudia & Jing-heng Sheng Ma. 2006. *Modern Mandarin Chinese Grammar: A practical guide*. London: Routledge.
- Rye, Ajit S. 1993. The Indian community in the Philippines. In K. S. Sandhu & A. Mani (eds.), *Indian Communities in Southeast Asia*, 709–773. Singapore: Institute of Southeast Asian Studies.
- Salazar, Lorraine C. 2008. The Indian community in metro Manila: Continuities, changes, and the effects of rising India. In K. Kesavapany, A. Mani & Palanisamy Ramasamy (eds.), *Rising India and Indian Communities in East Asia*, 499–548. Singapore: Institute of Southeast Asian Studies.
- SarDesai, D. R. 1997. *Southeast Asia: Past & present, 4th edition*. Boulder: Westview Press.
- Sato, Yosuke. 2011. Radical pro drop and fusional pronominal morphology in Colloquial Singapore English: Reply to Neeleman and Szendroi. *Linguistic Inquiry* 42, 356–365.
- SBCSAE, *The Santa Barbara Corpus of Spoken American English homepage*. Website. URL <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus> (Accessed 16 November 2019).
- Schneider, Edgar W. 2003. The Dynamics of New Englishes: from Identity Construction to Dialect Birth. *Language* 79(2), 233–281.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Schneider, Edgar W. 2014. New reflections on the evolutionary dynamics of world Englishes. *World Englishes* 33(1), 9–32.
- Scott, Henry W. 1994. *Barangay: Sixteenth-Century Philippine Culture and Society*. Quezon City: Ateneo De Manila University Press.
- Sedlatschek, Andreas. 2009. *Contemporary Indian English: Variation and change*. Amsterdam: John Benjamins.
- Setter, Jane, Cathy S. P. Wong & Brian H. S. Chan. 2010. *Hong Kong English*. Edinburgh: Edinburgh University Press.
- Sharma, Devyani. 2003. Discourse clitics and constructive morphology in Hindi. In Miriam Butt & Tracy Holloway King (eds.), *Nominals: Inside and Out*, 59–84. Stanford, CA: CSLI.
- Sharma, Devyani. 2009. Typological diversity in New Englishes. *English World-Wide* 30(2), 170–195.
- Sharma, Devyani. 2012. Shared features in New Englishes. In Raymond Hickey (ed.), *Areal Features of the Anglophone World* 80, 211–232. Berlin: Mouton de Gruyter.

- Shantakumar, G. & Pundarik Mukhopadhaya. 2008. Demographics, incomes and developmental issues amongst Indians in Singapore. In K. Kesavapany, A. Mani & Palanisamy Ramasamy (eds.), *Rising India and Indian Communities in East Asia*, 568–601. Singapore: Institute of South East Asian Studies.
- Sibayan, Bonifacio P. & Andrew Gonzalez. 1996. Post-imperial English in the Philippines. In Andrew W. Conrad & Alma Rubal-Lopez (eds.), *Post-imperial English: Status change in former British and American colonies, 1940–1990*, 139–172. Berlin: Walter de Gruyter.
- Singhvi, L. M., R. L. Bhatia, J. R. Hiremath, Baleshwar Agarwal & J. C. Sharma (eds.). 2001. *Executive Summary: Report of the High Level Committee on the Indian Diaspora*. New Delhi: Government of India, Ministry of External Affairs, Non-Resident Indian and Persons of Indian Origin Division.
- Singstat, Department of Statistics. 2010. *Singapore Census of Population 2010, Statistical release 1: Demographic characteristics, education, language and religion*. PDF document. URL [https://www.singstat.gov.sg/-/media/files/publications/cop2010/census\\_2010\\_release1/cop2010sr1.pdf](https://www.singstat.gov.sg/-/media/files/publications/cop2010/census_2010_release1/cop2010sr1.pdf) (Accessed 12 February).
- Singstat, Department of Statistics. 2017. *Population Trends 2017*. Singapore residents by age group, ethnic group and sex. PDF document. URL <https://www.singstat.gov.sg/-/media/files/publications/population/population2017.pdf> (Accessed 9 May 2019).
- Singstat, Department of Statistics. 2018. *Population Trends 2018*. Singapore residents by age group, ethnic group and sex. PDF document. URL <https://www.singstat.gov.sg/-/media/files/publications/population/population2018.pdf> (Accessed 12 February).
- Sivonen, Seppo. 2006. *Aasian kaupunkivaltioiden historia: Hong Kong, Macao ja Singapore*. Helsinki: Gaudeamus.
- Smith, Larry E. 1998. English is an Asian language. *Asian Englishes* 1(1), 172–4.
- Sridharan, Kripa & Terenjit S. Sevea. 2006. The Philippines. In Brij V. Lal, Peter Reeves & Rajesh Rai (eds.), *The Encyclopedia of the Indian Diaspora*, 198–199. Honolulu: University of Hawai'i Press.
- Stevens, Peter. 1980. *Teaching English as an International Language: From practice to principle*. Oxford: Pergamon Press.
- Subbārāo, Kārumūri V. 2012. *South Asian Languages: A syntactic typology*. Cambridge: Cambridge University Press.
- Tan, Ying-Ying. 2014. English as a 'mother tongue' in Singapore. *World Englishes* 33(3), 319–39.
- Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
- Thompson, Roger M. 2003. *Filipino English and Taglish: Language switching from multiple perspectives*. Amsterdam/Philadelphia: John Benjamins.
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Tsang, Steve. 2004. *A Modern History of Hong Kong: 1841–1997*. New York: IB Tauris.
- Turnbull, Constance M. 1977. *A History of Singapore 1979–1975*. Kuala Lumpur: Oxford University Press.
- Turnbull, Constance M. 2009. *A History of Modern Singapore, 1819–2005*. Singapore: NUS Press.

- Vennemann, Theo. 2002. Semitic > Celtic > English: The transitivity of language contact. In Markku Filppula, Juhani Klemola & Heli Pitkänen (eds.), *The Celtic Roots of English*, 295–330. Joensuu: Joensuu University Press.
- Wee, Lionel. 2004. Singapore English: Morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.), *A Handbook of Varieties of English, Vol 2: Morphology and Syntax*, 1058–1072. Berlin: Mouton de Gruyter.
- Wee, Lionel. 2013. Governing English in Singapore: Some challenges for Singapore's language policy. In Lionel Wee, Robbie Goh & Lisa Lim (eds.), *The politics of English: South Asia, Southeast Asia and the Asia Pacific*, 105–124. Amsterdam: John Benjamins.
- Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- Yong, Tan T. & Andrew J. Major. 1995. India and Indians in the Making of Singapore. In Mun C. Yong & VV B. Rao (eds.), *Singapore-India Relations: A primer*, 1–20. Singapore: Singapore University Press.





# APPENDIX A

Genre groups in SBCSAE and their corresponding groups in ICE (descriptions of file content obtained from <https://www.linguistics.ucsb.edu/research/santa-barbara-corpus#Contents>).

SBCSAE		
Corresponding ICE category	File number	Description
Private	1	This is a conversation recorded in rural Hardin, Montana. Mae Lynne is a student of equine science, and is the main speaker. She is telling Lenore (a visitor and near stranger) about her studies. Doris, Mae Lynne's mother, is doing housework, but joins the conversation near the end to discuss friends of their family.
Private	2	After-dinner conversation among four friends in San Francisco, California. Participants are in their late twenties or early thirties. Harold and Jamie are a married couple, Miles is a doctor, and Pete is a graduate student from Southern California.
Private	3	A conversation among three friends who are preparing dinner together, recorded in Southern California. Roy and Marilyn are a married couple, and Pete is a friend visiting from out of town. All participants are in their early thirties.
Private	4	Family conversation recorded in Santa Fe, New Mexico. The primary participants are three sisters all in their twenties.
Private	5	A conversation between a couple who are lying in bed, recorded in Santa Barbara, California.
Private	6	A very lively interaction between two female cousins in their mid-thirties, recorded in Los Angeles, California.
Private	7	Late-night conversation between two sisters, recorded in Montana.
Private	8	Task related interaction--an attorney preparing two witnesses to testify in a criminal trial. Recorded in San Francisco, California. Rebecca is a lawyer, June and Rickie are the witnesses, and Arnold is Rickie's husband.
Private	9	Task-related talk, a teenage couple recorded in Mobile, Alabama. Kathy is helping her boyfriend Nathan prepare for a math test.
Private	10	A business conversation recorded in New Mexico. Brad and Phil are board members of a local arts society. Phil wants to talk

business, while Brad keeps trying to leave to pick up his wife who's waiting for him at a bookstore.

Private	11	A conversation among three friends before lunch, recorded in Tucson, Arizona. All three participants are retired women; Samantha (Sam) is 72, Doris is 83, and Angela is 90.
Public	12	University lecture, recorded in Riverside, California. This is a Chicano Studies class; the professor is the primary participant, although it is a small, summer school class, and nine members of the class occasionally interact.
Private	13	This is a family conversation/birthday party, recorded in Fort Wayne, Indiana. The five participants are family members: Kendra (the birthday girl) and Kevin are siblings, Ken and Marci are their parents, and Wendy is Kevin's wife. This segment is highly interactional and contains a lot of overlap.
Private	14	Task related talk—this is a loan officers meeting, recorded in a bank in a small town in rural southern Illinois. Joe and Fred are loan officers working for the bank. Jim is the president of the bank, and Kurt is a board member.
Private	15	A conversation among three friends, recorded in Los Angeles, California. Ken and Joanne are a couple, and Lenore is a friend of theirs.
Private	16	A sales encounter, recorded in an audio store in Santa Barbara. Tammy is planning to buy a new tape deck. Brad, a salesman at the audio store, is discussing various tape decks which he is trying to sell her.
Private	17	A conversation between two male friends, recorded in Southern California.
Private	18	A task-related interaction recorded in a veterinarian office near Madison, Wisconsin. All five participants work in the office, some as secretaries and assistants and some as veterinarians.
Private	19	A family conversation, recorded in Michigan. Frank and Jan (a married couple) are talking with Ron—Jan's brother who is visiting from California. Brett and Melissa are Frank and Jan's junior-high-age children, who are doing homework and also taking part in the conversation.
Scripted	20	A segment from a sermon/lecture recorded at a small conference near Chicago, Illinois. The speaker is a pastor in his mid-seventies.
Scripted	21	A segment from a rather lively sermon recorded in Boston, Massachusetts.
Private	22	Task-related interaction, recorded in an air traffic control tower in Portland, Oregon. Lance is training to be an air traffic controller, and has just finished working a shift. Randy, an experienced controller, is giving Lance feedback/briefing on his performance on that shift.
Private	23	A segment from a book discussion group, recorded in Topeka, Kansas. The eleven participants are all women between the ages of 46 and 85

Private	24	This segment consists of game-playing and game-teaching on a computer, and was recorded near Cape Cod, Massachusetts. Jennifer and Dan are a couple in their early twenties.
Public	25	This is a segment from a lecture on the history and theology of Martin Luther, part of an evening class held at a church, recorded in Delaware.
Public	26	This is a city meeting, recorded in Chicago, Illinois. City officials interact with the public about a government grant which is being applied for, to fund community development. The city can only apply once, so are soliciting applications from various organizations and will submit the one they judge as best.
Public	27	An entertaining science lecture and demonstration, recorded at a large public science museum in Chicago, Illinois.
Private	28	A very intimate long-distance telephone conversation between a romantic couple in their early twenties, which took place between Pennsylvania and California.
Private	29	This is a business conversation recorded in Northern California between Seth and Larry, who are meeting for the first time. Seth works as an engineer who designs, installs, and sells heating and air conditioning units. Larry has invited him to his home to give him an estimate.
Scripted	30	A segment from a sermon, recorded at a large Baptist church in Chicago, Illinois.
Private	31	Face-to-face conversation recorded in a restaurant in Pullman, Washington. Sherry and Beth are sisters (in their late twenties), and Rosemary is their mother. The participants discuss what to order for lunch, interact with the waitress (Jamie) and engage in talk about family and friends while waiting for their food.
Private	32	A face-to-face conversation that takes place at an outdoor neighborhood 'block party' in Santa Fe, New Mexico. The three main participants are neighbors, age 60 and upward, all of whom happen to be named Tom. Discussion centers on life histories, World War II experiences, and neighborhood gossip. The three are briefly joined by Tucker (the daughter of Tom_1), and Elaine (the wife of Tom_3).
Private	33	A lively family argument/discussion recorded at a vacation home in Falmouth, Massachusetts. There are eight participants, all relatives or close friends. Discussion centers around a disagreement Jennifer (age 23) is having with her mother (Lisbeth).
Private	34	A late-night face-to-face conversation recorded in Northampton, Massachusetts. Participants are a married couple (Karen and Scott) in their early twenties. Karen has just returned home from work, and the two are talking while winding down for the evening.
Private	35	Lively family argument/discussion recorded in the kitchen of a family home in Pittsburgh, Pennsylvania.

Private	36	Face-to-face conversation recorded in Albuquerque, New Mexico. There are three participants and a baby. Lisa and Kevin are siblings, Marie (the baby's mother) is a friend of Lisa's. Much of the speech event focuses on interaction with, and talk about, the baby, as well as gossip about friends and co-workers.
Private	37	Informal, task-related (cooking) talk recorded in the kitchen of a family home in Corpus Christi, Texas. A family is making tamales. Main participants are Julia (an 80-year-old woman), her daughter (Dolores), and grandson (Shane). They are briefly joined by Kate (Shane's sister) who is watching TV in another room. The segment contains occasional codeswitching (English/Spanish).
Scripted/public	38	This segment is part of a tour of Hoover Dam, on the Nevada-Arizona border. The presentation is highly practiced. The main speaker also answers audience questions.
Private	39	Task-related talk, a training meeting recorded at an aquarium in Chicago, Illinois.
Scripted/public	40	Scripted tour of the Kentucky Horse Park / Museum. Presenter also addresses questions from the audience.
Private	41	Medical interaction recorded in Southern California. A patient (Paige) is consulting with her dietician (Kristen) regarding management of diabetes.
Private	42	Family argument and task-related talk, recorded in Pasco, Washington. The recording begins in a car, and moves to the kitchen of a family home. Main participants are three teenage sisters (Sabrina, Kendra, and Marlana), their mother (Kitty), and step-father (Curt). A friend of Sabrina's (Gemini) is also present. The dispute centers around Kitty's belief that Kendra stayed the night at a friend's house without permission, something which Kendra denies having done. Argument and shouting is interspersed with Saturday-morning housekeeping chores such as doing dishes and laundry.
Private	43	Face-to-face conversation recorded in the living room of a private home in Boise, Idaho, between Alice (a nurse, age 49) and her daughter Annette (a student and bank employee, age 24). Topics center mostly on their work day, as well as mutual acquaintances.
Private	44	Face-to-face conversation recorded in the living room of a private home in Milwaukee, Wisconsin. Two friends (Cam and Lajuan) are talking about their families and friends, and their own experiences as gay men.
Private	45	Face-to-face conversation recorded in the living room of an apartment in Milwaukee, Wisconsin. Two friends (Corinna and Patrick) are talking and watching TV. Topics are at times rather raunchy.
Private	46	Medical interaction, recorded in Shreveport, Louisiana. A patient (Darren) is consulting with his orthopedist (Reed) regarding a knee injury from a recent skiing accident.

Private	47	Face-to-face conversation between two cousins (Fred and Richard) in their early thirties, recorded in a private home in east Los Angeles, California. Topics include Richard's new job selling cars, Fred's frustration with factory work, and Richard's recent breakup with his girlfriend.
Private	48	Christmas morning traditions and gift-exchange among family members, recorded in Fresno, California. Tim and Lea are a couple in their late fifties, Judy is their daughter, and Dan is Judy's boyfriend.
Private	49	Face-to-face conversation recorded at an outdoor family birthday party near Boston, Massachusetts. There are ten speakers, all related. Four siblings in their mid thirties to mid forties: Dan, Al, Lucy, and Annette. Allen (Sr.), age 76, is their father. Al and Annette are twins. Linda is Al's wife, John is Annette's husband. Dave and Jane are Al and Linda's children. Glen is Lucy's son. Topics center primarily on recent renovations to Lucy's home.
Private	50	Face-to-face conversation among four roommates, recorded in a shared apartment in Burlington, Vermont. Speakers are all students at the University of Vermont, women ages 20-21. Speakers engage in small-talk, make plans for the evening, and discuss household matters.
Private	51	Conversation recorded before and during dinner, in a private home in Laguna Beach, California. There are four speakers, ranging in age from mid forties to early fifties. Sean and Bernard are a couple, Fran is a long-time friend visiting from New York. Alice is also a friend of Sean and Bernard, but had never met Fran. Discussion focuses on travels, and reminiscing about New York City.
Private	52	Phone conversation between family members at Christmas. Andrew and Cindy, a couple in their mid forties in Albuquerque, NM, are calling Andrew's sisters in San Antonio, Texas. Discussion centers primarily on Christmas and Christmas gifts, and topics prompted by recent television news shows.
Public	53	Task-related talk recorded in a small claims court in Santa Barbara, California. This segment consists of a judge pro tem hearing and deciding two cases.
Public	54	Public storytelling event recorded after a church potluck in Chicago, Illinois. The speaker, a professional storyteller in her mid-forties, tells several stories and interacts with the audience.
Public	55	Public lecture/forum in Santa Barbara, California. Noted artist and ceramist Beatrice Wood gives a public lecture at the Santa Barbara Museum of Art, shortly after her 101st birthday. Wood talks about her life and answers audience questions.
Private	56	Face-to-face conversation recorded on a ranch near Colorado Springs, Colorado. Julie has recently bought a pony from Gary's

		wife, and is giving him a bill-of-sale. She then gives a brief tour of her property and barn.
Public	57	Task-related talk, a recording of a judo class in Shreveport, Louisiana. The five students and their instructor are males between the ages of 22 and 37. The instructor is demonstrating and coaching the Hane-Makikomi throw, which students are practicing with varying degrees of success.
Private	58	Face-to-face conversation recorded in a private home in Boise, Idaho. Sheri, a single mom in her mid thirties, and her son Steven (age 11) talk while Sheri prepares dinner.
Private	59	Face-to-face conversation, recorded in a family home near Beloit, Wisconsin on Christmas Eve. Cam and Fred are a couple in their early thirties. Jo and Wess are Cam's parents. Topics include talk about family and friends, a football game which Wess and Fred had just finished watching, and holiday baking.
Private	60	Face-to-face casual conversation recorded in an office in Shreveport, Louisiana. The two speakers, Jon (age 72) and Alan (age 66) are friends/co-workers taking a break from work. Alan is primarily telling Jon about his travel adventures and interests.

---

## APPENDIX B

Genre groups in Frown and their corresponding groups in ICE (Data for Frown obtained from <http://clu.uni.no/icame/manuals/FROWN/INDEX.HTM>).

ICE	Frown			
Genre group	Genre group	Category	Content of category	No. of texts
Press news report	Press	A	Reportage	44
Persuasive writing		B	Editorial	27
-		C	Review	17
-	General prose	D	Religion	17
Instructional writing		E	Skills, trades and hobbies	36
Popular writing		F	Popular lore	48
-		G	Belles lettres, biographies, essays	75
-		H	Miscellaneous	30
Academic writing	Learned	J	Science	80
Creative writing	Fiction	K	General fiction	29
Creative writing		L	Mystery and detective fiction	24
Creative writing		M	Science fiction	6
Creative writing		N	Adventure and western	29
Creative writing		P	Romance and love story	29
Creative writing		R	Humor	9
<b>Total</b>	<b>500</b>			





# PUBLICATIONS

- Publication I Parviainen, Hanna. 2012. Focus particles in Indian English and in other varieties. *World Englishes* 31(2), 226–247.
- Publication II Parviainen Hanna. 2016. The invariant tag *isn't it* in Asian Englishes. *World Englishes* 35(1), 98–117.
- Publication III Parviainen, Hanna. 2017. Omission of direct objects in New Englishes. 2016. In Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds.), *Changing English: Global and Local Perspectives* 92, 129–153. Berlin, Boston: Mouton de Gruyter.
- Publication IV Parviainen, Hanna & Robert Fuchs. 2018. 'I don't get time only': An apparent-time investigation of clause-final focus particles in Asian Englishes. *Asian Englishes* 21(3), 285–304.



# PUBLICATION I

**Focus particles in Indian English and other varieties**

Hanna Parviainen

World Englishes 31(2), 226–247

<https://doi.org/10.1111/j.1467-971X.2012.01752.x>

**Publication reprinted with the permission of the copyright holders.**



## Focus particles in Indian English and other varieties

HANNA PARVIAINEN\*

**ABSTRACT:** This paper studies the Indian English use of clause-final focus particles *also* and *only* in sentences such as ‘He doesn’t listen only’ (ICE-IND) in order to see if this use has spread to Singapore, Philippine or Hong Kong English. The data for these Asian varieties was obtained from the International Corpus of English and the British and American varieties were added for point of reference. Although the main focus of this quantitative study is on the syntactic analysis of the data, some semantic aspects are also discussed. The results suggest that the Indian English way of placing focus particles in a clause-final position could be a local innovation which has today spread to Singapore, Philippine and possibly to Hong Kong English. In addition, the data was analysed for all subcategories of ICE, which reveals that the frequencies for this use correlate positively with the level of informality of the speech situation in all four Asian varieties, whereas the results for British English suggest the opposite. Thus, the results indicate that clause-final focus particles have acquired some additional uses, especially in spoken Indian English, and that these features have now spread to other English varieties spoken in Southeast Asia. The paper suggests that the cause for this can be found from the centuries-old role that India has had as a cultural force in the Southeast Asian region.

### INTRODUCTION

Today, some of the largest populations using English can be found in Asia. In fact, this trend is now growing so rapidly that it has made some scholars predict that ‘by 2020, Asia will be the CENTRE of gravity of the English language’ (Dalby, cited in Gupta 2001: 149). Despite all the controversy surrounding the English language and its colonial baggage, it has remained as an official language in India, Singapore and Hong Kong, which were once colonies of Britain, and the Philippines, which was colonized by the US. One explanation for the popularity of English in these multilingual nations is the fact that it provides an ‘interethnically neutral link language’ (Schneider 2007: 167), while in many countries it is also believed that English will help them to create stronger economic connections with the rest of the world (Crystal 2003: 106; Dolan 2003: 82; Dixon 2005: 27).

Although Agnihotri (1999: 193) has argued that ‘there is no syntactic feature that may be said to be uniquely associated with Indian English’, researchers such as Bhatt (2000) and Lange (2007) have proven otherwise. Their studies on the use of focus markers in Indian English (IndE) indicate that the speakers of this variety have created a potentially innovative way of using focus particles such as *only*, which Lange (2007: 115) suggests ‘to be the first candidate for a nativized syntactic pattern’ that has been documented in the variety. India has had a significant impact on the development of the cultures in Southeast Asia (Lamb 1975: 442; SarDesai 1997: 16; Rai 2008: 29–30), and since English has been spoken in India longer than in any other country in Asia, the local IndE variety could be expected to have influenced some of the other English varieties spoken in the Southeast Asian region. This idea is supported by Hogue (2001), who has studied the use of Anglo-Indian<sup>1</sup>

\*School of Language, Translation and Literary studies; University of Tampere; Kanslerinrinne 1. 33014; Tampere, Finland. E-mail: hanna.e.parviainen@uta.fi

words in Singapore English (SinE), Malaysian English (MyE), Hong Kong English (HkE), Philippine English (PhiE) and “International” English. Hogue’s (2001) results indicate that Indian influences did in fact spread in the area even during colonial times; since the early contacts between the British and the local Indian population, words from Indian languages were integrated into the lexicon of the English spoken in India, from where they spread to other emerging varieties of English in the region. This argument is further supported by the fact that the British often sent English teachers from India to their other colonies in the area (Platt, Weber, and Ho 1983: 8; Hogue 2001: 169; Rai 2007: 178), and because India has continued to export its English teachers even to this day (Graddol 2010: 94), the Indian variety may still be influential in some parts of Asia. Today, India’s economic rise has increased the country’s importance in the area, which is why many Southeast Asian countries have begun to strengthen their connections with India (Rai 2007: 188; Sridharan and Sevea 2007: 199). In addition, the last decades have witnessed a new wave of professional Indian workforce migrating to, for example, Singapore (Rai 2007: 176), which could also have increased the influence of IndE in the country. Dalby has argued that because the role of Asian Englishes will grow in the future, IndE will also become more influential, since it has the highest number of speakers in Asia (Dalby, cited in Gupta 2001: 149) and bearing all this in mind, his prediction might well become true one day.

Therefore, the purpose of this study is to examine the current level of influence that the Indian variety has on some Southeast Asian varieties of English. For this paper, I conducted a quantitative corpus-based analysis on the potentially innovative use of clause-final focus particles *also* and *only* in IndE in order to see if this use has spread to SinE, PhiE or HkE. The data for all Asian varieties were obtained from the *International Corpus of English* (ICE) (2002a; 2002b, 2004; 2006) and the British (BrE) (ICE 1998) and American (AmE) (*The Freiburg-Brown Corpus of American English* 1999; *The Santa Barbara Corpus of Spoken American English* 2002–2005) varieties were added for point of reference. Although the focus of this study is on the syntactic analysis of the data, some semantic aspects of this phenomenon will also be discussed.<sup>2</sup>

## FOCUS PARTICLES

Focus particles have been studied by various researchers over the years and the following section will present a brief survey of the uses that *also* and *only* have in the English language,<sup>3</sup> which will later be compared with the ways focus particles are used in some Indian languages.

According to Leech and Svartvik (2002: 206), the most common location of the focus is at the end of a phrase, where the nucleus will be on the last ‘major-class word’ of the tone unit. However, the position of focus adverbs, that is, the adverbs referring to the focused element in the clause, can vary greatly. According to Jacobson (cited in Nevalainen 1991: 39), the adverb can immediately precede or follow its centre of focus, or it may be further away from it in anteposition or postponement. The following example shows all the possible locations of focus adverbs in a sentence:

- (1) (Only) they (only) fed (only) the cats (only). (Nevalainen 1991: 39)

Despite all these different possibilities, there are a number of constraints which regulate the place a focus adverb can actually take.<sup>4</sup> This study will examine the adverbs *also*

and *only* which follow the verb, the complement or the object of a phrase and thus the section under observation in the previous example would be '(only) the cats (only)'. The majority of theories only discuss the semantic qualities of adverbs when they precede their focus, which seems to suggest that this is also their most common position. This view is supported by Brown (1992: 89) who argues that 'in GenE, [*only*] is normally placed with the verb [or] ... it can be placed immediately before the element it focuses on'. Lange (2007: 93) agrees with this argument, but adds that *only*, for instance, can also follow its focus as the phrase 'ladies only' shows and Brown (1992: 89) refines Lange's statement further by noting that 'the effect in GenE of placing *only* after the element is to sound rather authoritarian, formal or declarative'. As the question of location is central in this study, the issue will be discussed in greater detail in the Methodology section.

#### *Additive adverbs: also*

Focus particles can be divided into *additive* and *restrictive* groups depending on the influence they have on the semantic meaning of a sentence (König 1991: 32–3). The *additive*, or, 'inclusive', as they are sometimes called, includes adverbs such as *too* and *also*. They are used to indicate that an item is added to another item, which can happen either on a clausal (2) or phrasal (3) level:

- (2) Oh, my dad was a great guy, too. (conv) (Biber, Johansson, Leech, Conrad, and Finegan 1999: 556)
- (3) I can hear the hatred, but also the need. (fict) (Biber et al. 1999: 556)

When operating on a clausal level, as in example (2), these adverbs often point towards a specific part of the phrase's meaning, which it indicates to be 'additional to something else' (Biber et al. 1999: 556). However, this can sometimes be hard to detect when the phrase is taken out of context. For example, it is difficult to say whether sentence (2) refers to 'my dad was a great guy (like his brother)' or that 'my dad was a great guy (in addition to being a great father)'. This problem of interpreting the referent correctly has been discussed by many researchers and therefore it will be examined further in the Methodology section of this study.

The category of additive particles can be further divided into two sub-groups, *scalar additive particles* and a group that is used to express 'simple inclusion' (Lange 2007: 91; König 1991: 62). This latter group is important for this study as it includes adverbs such as *also*, *too*, *as well*, *similarly* and *either*, when they do not appear to express any evaluation of the given information (Lange 2007: 91). As an example Lange presents the following:

- (4) She is *also* an accomplished playwright and a dramatist. (ICE-IND:S1B-045#1:1, cited in Lange 2007: 91)

Although this example certainly provides additive information about the woman in question, there is no indication that being a playwright is valued over being a dramatist or vice versa.

#### *Restrictive adverbs: only*

Focus particles belonging to the second group are called *restrictive particles*: they include adverbs such as *only* and *especially*. Sometimes they are also called *exclusive*

*particles* due to their ‘semantic property of excluding other focus alternatives’ (Nevalainen 1991: 31), and indeed their function appears to be to focus on the importance of one element in the clause by ‘restricting the truth value of the proposition either primarily or exclusively to that part’ (Biber et al. 1999: 556). For the purposes of this study, the particle *only* is of great interest, and its various meanings can be analysed further with the following examples:

- (5) *Only* they will know how much they have right (ICE-SIN:S1A-017#239:1:A)
- (6) She was *only* a shopkeeper’s daughter (Lange 2007: 93)

The first example contains an instance where *only* is used to create exclusivity whereas in the second example the particle could be argued to have a scalar meaning instead. What both of these sentences have in common is that the focus particle is in its most common location which is, according to Nevalainen (1991: 33), before its focus.

### *Focus particles in Indian languages*

Having described the common meanings and uses of focus particles *also* and *only* in the previous section, I will now examine their use in IndE. However, in order to fully understand why speakers of IndE could have begun to use *also* and *only* in a new way, the influence of local languages should also be discussed. Therefore, the following section will briefly describe the use of focus particles in some Indian languages such as Hindi.<sup>5</sup>

According to Koul (1990: 22), the main particles in Hindi which are used to denote contrast or emphasis are *hii*, *bhii* (also sometimes transliterated to as *hī* and *b<sup>h</sup>ī*), *to*, *tak*, *bhar* and *maatra* of which the first two are of greatest interest concerning this study. These words have been called by many names, such as ‘particles’ (Koul 2008: 137), ‘emphatic elements’ (Verma 1971: 85), ‘emphasis markers’ (Kachru 2006: 108), ‘emphatic enclitics’ (McGregor 1995: 30) and ‘discourse markers’ (Sharma 2003: 60), but here I shall use the term *focus particles* in order to keep my discussion coherent with the previous sections. Sharma (2003: 60–2) explains that *hii* is often used to mark exclusive focus, similar to *only*, whereas *bhii* is used to mark inclusive focus, similar to *also*. Since these Hindi particles ‘only take scope over the constituents to their left’ (Sharma 2003: 137; see also Koul 2008: 137), it could be argued that one of the differences between the English and the Hindi focus particles is that in English, focus particles precede their referents very frequently, whereas in Hindi they are placed after their referent. As an example of the way these are linked to their centre of focus, Sharma presents the following examples:

- (7) rād<sup>h</sup>ā=nē=hī      bacchō=kō      kahānī      sunāyī  
Radha=ERG=EXCL FOC children=ACC story –F.NOM hear-caus-PERF.F.SG  
‘It was (only) Radha who told the children a story’ (Sharma 2003: 61)
- (8) rād<sup>h</sup>ā=nē=b<sup>h</sup>ī      bacchō=kō      kahānī      sunāyī  
Radha=ERG=INC FOC children=ACC story-F.NOM hear-caus-PERF.F.SG  
‘Radha (also) told the children a story’ (Sharma 2003: 62)

From (7) and (8) it can be seen that the Hindi *hii* and *bhii* are indeed placed after their referent, although in the translated version of (7), the focus particle *only* is placed before its referent, which is its most common location in English. It is worth noting that the word *also* in translation (8), without the original sentence in Hindi above it, could also be interpreted as referring to something after it. For example, the sentence could mean that ‘Radha (first



put the children to bed, and then she) also told them a story'. However, this reading of the sentence becomes impossible when the original Hindi sentence is examined. In both (7) and (8), the focus particles *hii* and *bhii* follow the name Radha, which signals that it is the centre of focus for both particles. As proof of this, Sharma presents three examples that highlight the different ways in which *hii* and *only* are connected to their referents:

- (9) \*māyā=nē hī anū kō KITĀB dī  
 Maya Anu book give  
 'Maya only gave Anu A BOOK'. (Sharma 2003: 64)
- (10) \*māyā=nē hī ANŪ=KŌ kitāb dī  
 Maya Anu book give  
 'Maya only gave ANU a book'. (Sharma 2003: 64)
- (11) MĀYĀ=NĒ hī anū kō kitāb dī  
 Maya Anu book give  
 'Only Maya gave Anu a book'. (Sharma 2003: 64)

From the examples above the difference between the scope of *hii* and *only* become apparent. In examples (9) and (10), when trying to keep the word order of the translated version as close to the original Hindi sentence as possible, the possibility of a false reading is great. As was mentioned earlier, the most common location for a focus particle in English is before its referent and thus the interpretations of the referent of *only* in (9) and (10) seem only natural. However, when the original Hindi sentence is taken into consideration, these readings turn out to be incorrect. In order to highlight that the focus of *only* is indeed *Maya*, Sharma (2003) has placed it before its referent in (11), which is its usual location in English and thus creates a translation that carries the same meaning with the original Hindi sentence.

In addition to the exclusive meaning that the word *hii* has, it can, according to Verma (1971), also take on an emphatic meaning:

- (12) laRkee hii aa rahee hāī  
 Only the boys are coming. (exclusive) (Verma 1971: 91)
- (13) laRkee hii aa rahee hāī  
 It's the boys who are coming. (emphatic) (Verma 1971: 91)
- (14) māī hii jaaūūga  
 I myself will go. (Verma 1971: 91)

In examples (12) and (13), the two possible meanings of the same particle are revealed through their differing translations. As another example of the emphatic reading, Verma presents the example (14) where the non-exclusive nature of *hii* is made even more apparent as it is translated as *myself*.

Interestingly, there are some other Indian languages that also could have contributed to the creation of this IndE innovation. For example, Tamil, which is also an official language in Singapore, contains two emphatic particles *-ee* and *taan* which are placed directly after their referent (Asher 1985: 94) and thus their location in a sentence resembles the syntactic rule found from Hindi as the following examples show:

- (15) avan kaaleylleyee vantaan  
 he morning-loc-emph come-past-3sm  
 'He came in the *morning* (and not at some other time of day)'. (Asher, 1985: 94)

- (16) aamaam, raaman taan cepnaan  
 yes, Raman emph do-past-3sm  
 'Yes. It was Raman who did (it)'. (Asher 1985: 94)

As (15) and (16) illustrate, *-ee* is used as a marker of exclusivity whereas *taan* seems to be a general marker of emphasis and the translation provided for the latter particle resembles greatly that presented in example (13) for the emphatic use of Hindi *hii*. However, Asher (1985: 94) mentions that there is 'considerable overlap between the use of these two emphatic markers', but despite this, *-ee* is more commonly used to mark contrastive emphasis whereas *taan* is more often used to mark noncontrastive emphasis. An example of this difference in meaning is given below:

- (17) aantaan neettee vantaaru;  
 Anand yesterday-emph come-past-3sh  
 Muttu innikki taan vantaaru  
 Muthu today emph come-past-3sh  
 'Anand came *yesterday*; Muthu came only today'. (Asher 1985: 95)

In addition to Hindi and Tamil, a similar focus particle construction is used in Marathi where, according to Pandharipande (1997: 239–41), the emphatic inclusive particle *hī* and the emphatic exclusive particle *ts* are placed after their referents X where they convey the meanings 'X hī "X also", X ts "X only"'. Although the task of examining all of the major Indian languages and how they use focus particles exceeds the scope of this paper, this brief examination of a few of the commonly spoken Indo-European and Dravidian languages supports Lange's (2007: 113) argument that 'many Indian languages have enclitic focus particles ... [and in] the case of *only* ... in [IndE], substrate influence is indisputable'.<sup>6</sup>

## METHODOLOGY

### Data

All four Asian varieties and the British variety were studied by using *The International Corpus of English* (ICE), but since the ICE corpus of American English is not yet available, the variety was studied by using two different corpora: the 250,000 word Santa Barbara corpus of spoken American English and the Freiburg-Brown corpus of American English (Frown), which contains one million words of written language.<sup>7</sup> All the ICE corpora consist of approximately one million words of spoken and written English (see Table A1 in the Appendix) and they all have an identical structure which facilitates the comparison between different varieties and different genres. Since the two corpora used for AmE do not correspond to ICE in their structure, I have had to settle for the overall frequencies of only two genres: spoken and written.

### Methodology

All the corpus searches for the two particles were done with the WordSmith 5.0 program, which allows the user to search for individual words throughout different sections of the corpora. It also provides a link to the text unit where the found words belong to, thus providing additional information for the semantic analysis of the word and the context where it appears.

Although previous studies by Bhatt (2000) and Lange (2007) are based on the semantic analysis of the particles regardless of their location in a sentence, I have decided not to follow their methods for a simple reason: since the usual position of a focus particle is before its focus (Nevalainen 1991), but it is commonly placed after its focus in IndE (Bhatt 2000; Lange 2007), there is no way of knowing when a speaker of, for example, SinE is using the focus particle to refer to a word preceding or following that particle. A good example of this problem is presented by Lange (2007: 108), who argues that when a sentence such as (18) is ‘taken out of its [IndE] context, [it] does not necessarily have to be interpreted as representing an innovative use of *only*.’

- (18) Yeah – yesterday **only** I came here.  
(ICE-IND:S1A-075#128:1:A, cited in Lange 2007: 107)

Although Lange’s solution to this problem is to observe the discussion preceding and following the sentence containing the particle *only*, this was considered to be too unreliable for this study. Because the speakers in all six varieties cannot be assumed to give similar meanings to sentence structures if those structures can be interpreted in multiple ways, I have decided to focus solely on syntactic criteria. Therefore, only those cases will be examined where focus particles *also* and *only* are in a clause-final position, which disposes of the problem of interpreting the referent correctly.

The collection of data was conducted in two stages. First, all the clause-final instances of *also* and *only* were selected manually. In practice, this meant that several methods had to be used in order to assure that these particles were indeed in clause-final position. For spoken data, all clauses containing *also* and *only* were subjected to a syntactic analysis where the elements of the clauses were identified. Then, those clauses where *also* and *only* were after the verb (19), or the complements (20) or the objects<sup>8</sup> (21) of the clause were selected from the data as the following examples show:

- (19) Just running *only*. (ICE-SIN:S1A-063#97:1:B)  
(20) I thought Thomas was a an Asian *also*. (ICE-PHI:S1A-028#15:1:B)  
(21) Then why why don’t you go to uh America *also*. (ICE-HK:S1A-017#85:1:A)

I then continued by examining the context where these clauses appeared in order to make sure that they were indeed in clause-final position. Contextual clues of this were, for example, when the clause was followed first by a longer pause or a co-ordinating conjunction and then by another independent clause, or when the speaker ended his/her speech turn with *also* or *only* and was followed by another speaker as the example below indicates:

- (22) A: You know it’s Raymond’s birthday *also*. (ICE-PHI:S1A-051#93:1:A)  
B: Yes I heard. (ICE-PHI:S1A-051#94:1:B)

In addition, subjects or predicates are sometimes omitted, especially in spoken English, which had to be taken into consideration and thus such clauses were selected individually after careful examination of their context. An example of this is given below:

- (23) A: Oh no for food *only*. (ICE-SIN:S1A-017#180:1:A)  
C: Ya this is *only* for food. (ICE-SIN:S1A-017#181:1:C)

The task of finding all clauses where *also* and *only* were in final position was less complex for the written section of the data, as punctuation provided strong indications where one clause ended and another began. Finally, the absolute and relative frequencies (per one million words) were calculated and the comparison between AmE, BrE, IndE, SinE, PhiE and HkE was possible. Although the use of clause-final *also* and *only* is rare in BrE and AmE, previous studies by Bhatt (2000), Lange (2007) and Balasubramanian (2009) have shown that this usage is common in spoken IndE. Interestingly, also SinE has been noted to use *only* (Brown 1992: 88–90) and *also* in final position (Brown 1992: 6; Bao and Hong 2006), but possible substrate influences have been suggested only for the use of clause-final *also* (Bao and Hong 2006: 109–12). However, researchers such as Platt et al. (1983: 8) have noted that because of the high number of Indian English teachers that worked in Singapore during colonial times, there are now ‘similarities in lexical and syntactic usage between [SinE and IndE]’ and thus it is also possible that the use of *also* in SinE is in fact an example of the way IndE has influenced SinE. In addition, it should be noted that previous studies of PhiE and HkE have not found the use of clause-final focus particles to be a common feature in these varieties. Therefore, because, *also* and *only* at the end of a sentence have been argued to be rare in BrE and AmE and common in IndE, my hypothesis is that the influence of IndE on SinE, PhiE and HkE could be detected in the elevated numbers of cases where *also* and *only* are in a clause-final position. Because the possible influence of other substrate languages cannot be ignored, these will be examined in greater detail with the results of this study in the Discussion section.

## CORPUS FINDINGS

### *Also*

#### Frequency of *also*

This section discusses the frequency of clause-final *also* in the Santa Barbara corpus, Frown, ICE-GB, ICE-IND, ICE-SIN, ICE-PHI and ICE-HK. The figures obtained for this section are of special interest, because no systematic statistical research across several Asian varieties of English has been done on this question until now. Table 1 presents the frequencies of clause-final *also* in the six varieties under examination (normalized per one million words).

Figure 1 reveals a striking difference between the varieties. The frequency for clause-final *also* in BrE is the lowest with only 2.8 hits per million words, followed by AmE with a slightly higher frequency of 23.1 for clause-final *also*, whereas ICE-IND shows a very high frequency of 417.1 for the entire corpus. One possible explanation for the IndE results could be that the variety has been influenced by substrate languages, which is a theory

Table 1. Frequency of *also* per one million words, absolute figures in parentheses

<i>Also</i>	Santa Barbara / Frown	ICE-GB	ICE-IND	ICE-SIN	ICE-PHI	ICE-HK
Spoken	63.9 (16)	1.6 (1)	622.6 (425)	233.9 (141)	172.4 (117)	84.4 (62)
Written	12.9 (13)	4.7 (2)	92.5 (40)	7.5 (3)	15.7 (7)	16.2 (8)
All	23.1 (29)	2.8 (3)	417.1 (465)	144.0 (144)	110.3 (124)	57.0 (70)

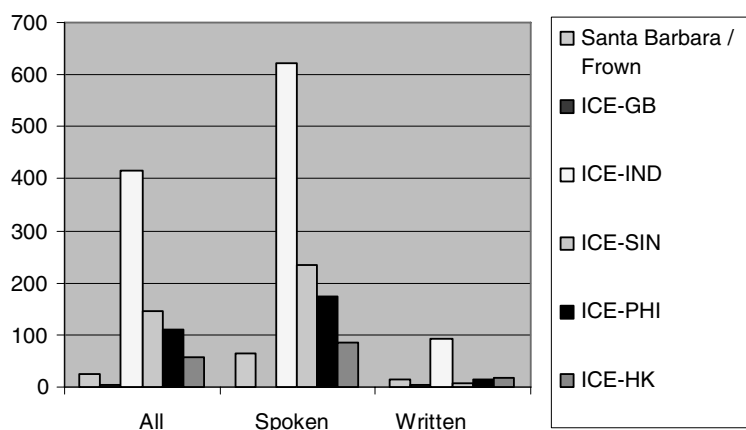


Figure 1. Frequency of *also* per 1 million words

supported, for example, by Lange (2007). Interestingly, the highest frequencies can all be found from the Asian varieties as the results from IndE are followed by SinE, PhiE, HkE, AmE and BrE in descending order (417.1 / 144.0 / 110.3 / 57.0 / 23.1 / 2.8). When looking at the frequencies from the spoken sections of the corpora, the order of the varieties stays the same, although the difference between the Asian and Western varieties becomes even more apparent (622.6 / 233.9 / 172.4 / 84.4 / 63.9 / 1.6). On the basis of these figures, it could be argued that SinE has adopted the IndE way of using focus particles, whereas the lower frequencies of clause-final *also* in PhiE tell of a similar, but weaker trend in the use of the word. Although spoken AmE shows surprisingly high results, the frequency for clause-final *also* is not high enough to solely explain the results found from the spoken section of ICE-PHI (63.9 / 172.4). In addition, HkE has slightly higher frequencies for the use of clause-final *also* when compared with BrE and AmE, but as the frequencies for spoken HkE are quite low when compared with the frequencies of other Asian varieties, nothing certain can be said about the variety except that it might have been influenced by IndE. This issue will be discussed further in the following sections. The spread of the use of clause-final *also* from IndE to SinE, PhiE and (possibly) HkE is represented in the examples (24) to (27), whereas (28) and (29) show the way *also* is used in a clause-final position in AmE and BrE:

- (24) But at present there is no time and no energy *also*. (ICE-IND:S1A-030#113:1:B)
- (25) So next time you must study until uh university *also*.  
(ICE-SIN:S1A-006#188:1:B)
- (26) The other leg is just following so when you come down it's just bending *also*.  
(ICE-PHI:S2A-053#15:1:A)
- (27) They went a little quick *also*. (ICE-HK:S2A-002#98:1:A)
- (28) Did he use a newspaper that time *also*? (Santa Barbara, 1322.46 1324.46)
- (29) That being what the rest of the work force have done *also*.  
(ICE-GB:S2A-067 #124:1:A)

Here it must be noted that all four Asian corpora contained three types of cases, that is, cases where focus particles had been used in the ways often associated with BrE and AmE, in the IndE way and in ways where this distinction was impossible to make.<sup>9</sup>

When looking at the results from ICE-IND, another issue becomes clear: the written section has a noticeably lower frequency than the spoken one (92.5 / 622.6). This supports the arguments presented by Bhatt (2000) and Lange (2007), who claim that the presentational use of focus particles appears mostly in spoken language. The following examples are from the written sections of the Asian corpora:

- (30) So they started to develop social forestry *also*. (ICE-IND: W1A-001#105:8)
- (31) Similarly, when his social actions drive up business costs, which are then passed on to consumers through higher prices, he is spending their money *also*. (ICE-SIN:W2A-017#40:1)
- (32) Ask them to make arrangements for Eurail *also*. (ICE-PHI:W1B-005#137:2)
- (33) Alongside with hose reels are fire alarms and break glass call points which can be installed as recessed fittings *also*. (ICE-HK:W2A-035#76:1)

What can be gathered from sentences (30) to (33) is that in the written varieties of Asian English, the word *also* is used as an additive particle. For example, (30) could be rephrased as 'They also started to develop social forestry' which would indicate that the cases where *also* is placed in a clause-final position are due to the influence of the focus particle use that is common in spoken IndE and possibly other languages spoken in the area. It is noteworthy that although all four Asian varieties show higher frequencies in the spoken section, in ICE-GB the trend is the opposite, as its written section has a higher frequency when compared to the results of the spoken section (4.7 / 1.6). Therefore the results of AmE are rather surprising, as the spoken genre has higher frequencies for clause-final *also* than the written genre (63.9 / 12.9).

The case of the Philippine variety is interesting, because unlike SinE and HkE, it does not seem to have any historical connections with IndE. One possibility is that the use of clause-final focus particles is not tied to any specific ethnicity or nationality. Instead of entering the PhiE variety directly, as, for example, in the form of ethnic minorities, it is possible that this potentially innovative IndE use of focus particle *also* could have filtered into the PhiE variety through other Southeast Asian varieties, which according to Lamb (1975: 442) is also how Indian influences have usually entered the Philippine islands in the past. Because focus particles used restrictively/additively in clause-final position are rare in BrE and AmE, a possible explanation for the high frequencies found from SinE, PhiE and HkE could be that they have been influenced by IndE, as it favours the use of this feature the most.

### Distribution of *also*

The identical structure of the ICE-corpora allows the results presented in the previous section to be broken down into smaller subsections; by calculating the frequency of clause-final *also* for each genre separately, it was possible to differentiate between the registers that favour the use of clause-final *also* from those that did not. This analysis (see Table A2 in the Appendix) reveals an interesting pattern for the British variety when compared with the Asian varieties: In BrE, there is a positive correlation between the frequencies

for clause-final *also* and the formality of the speech situation<sup>10</sup> as the only instance of this use occurs in the genre ‘Unscripted speeches’. In contrast, the number of hits for clause-final *also* in Asian varieties correlates negatively with the formality of the situation where the speech act takes place, as the highest frequencies can be found from ‘Private dialogues’ and the lowest from ‘Scripted monologues’. It is also worth noting that the order established already in the previous section remains almost the same throughout the different speech categories so that the highest frequencies for *also* can be found from IndE, followed by SinE, PhiE and HkE. The only exceptions to this can be found from the ICE-PHI categories for ‘Spoken dialogues’ and ‘Unscripted monologues’, which exceed those found from SinE.

The results for the written sections of the corpora are rather different, as they are scattered more evenly across categories, although the majority of the corpora have relatively high proportions in the categories of ‘Non-professional writing’ and ‘Correspondence’. The observation that these categories appear to favour the use of clause-final *also* could, again, be explained with the notion of formality. Since these two example groups are less formal in nature and are relatively close to spoken language, it is fair to expect that also their syntax resembles that of spoken language to some extent as, for example, Balasubramanian (2009: 225) has argued. As mentioned earlier in this paper, the clause-final focus particle *also* is not commonly used in British or American English and the low numbers from ICE-GB, Santa Barbara corpus of spoken American English and Frown support this view.

## Only

### Frequency of *only*

This section discusses the frequency of clause-final *only* in the six varieties of English examined for this study. Although previous studies have shown that IndE has developed a new use for the particle *only*, no systematic research enabling the comparison between different Asian English varieties across different genres has been done until now. Figure 2 and Table 2 present the frequencies of clause-final *only*, normalized per one million words, in the seven corpora examined in this study:

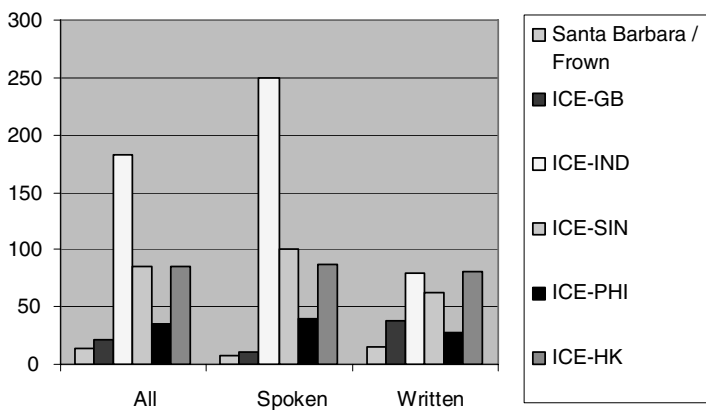


Figure 2. Frequency of *only* per 1 million words

Table 2. Frequency of *only* per 1 million words, absolute figures in parentheses

<i>Only</i>	Santa Barbara / Frown	ICE-GB	ICE-IND	ICE-SIN	ICE-PHI	ICE-HK
Spoken	8.0 (2)	11.0 (7)	249.1 (170)	101.2 (61)	39.8 (27)	87.1 (64)
Written	15.9 (16)	37.8 (16)	78.6 (34)	62.9 (25)	26.9 (12)	81.1 (40)
All	14.3 (18)	21.7 (23)	183.0 (204)	86.0 (86)	34.7 (39)	84.7 (104)

Again, the lowest overall frequencies for the use of clause-final *only* are found in the Western varieties, but this time AmE has the lower frequencies when compared with BrE (14.3 / 21.7). Interestingly, the written sections of both BrE and AmE contained more hits for the clause-final *only* than their spoken sections. This deviation from the pattern found from the Asian varieties is highly interesting, since it gives reason to suspect that in AmE and BrE, the way of placing *only* at the end of a clause belongs to the literary (34), or more official spoken genre (35) with possible difference in meaning too when compared with the Asian varieties:

- (34) These safety precautions are for information purposes *only*. (Frown, J77 24)
- (35) Concessionary fees apply to Camden residents *only*.  
(ICE-GB:S2B-044#121:2:A)
- (36) Extensive literature is available on these techniques and on certain occasions when no other methods are available, one has to depend on these methods *only*.  
(ICE-IND:W2A-038#96:1)
- (37) Not too bad *only*. (ICE-SIN:S1A-002#101:1:C)
- (38) Benjamin have a girlfriend *only*. (ICE-HK:S1A-050#375:1:B)

As can be seen from sentences (34) and (35), clause-final *only* is used as a restrictive focus particle in both written AmE and spoken BrE, whereas both the restrictive and the IndE use of the particle can be found from the Asian varieties in (36), (37) and (38). When looking at the results from the other corpora, it becomes evident that the order established already with *also* is almost the same, the only exception being the PhiE variety: in the spoken sections of the corpora, the overall frequencies of IndE are followed by SinE, HkE, PhiE, BrE and AmE in descending order (249.1 / 101.2 / 87.1 / 39.8 / 11.0 / 8.0) and again, the results for PhiE cannot be explained by the influence of AmE alone.

The results from ICE-HK are very interesting as they do not follow the trend that has been set by the other Asian varieties. When the frequencies for clause-final *also* (Table 1) and *only* (Table 2) in the spoken section of ICE-HK are compared with each other, it becomes clear that they are almost equally high (84.4 / 87.1). In fact, the use of *only* actually exceeds that of *also*, which would seem to indicate that the possible IndE influence on the HkE use of *only* is not as distinct, because in the other Asian corpora, the use of *also* is much more frequent when compared with the use of *only*. In addition, since the frequencies for clause-final *only* are almost equally high in the spoken and written sections of ICE-HK (87.1 / 81.1), there must be something else that is causing these results. This issue will be examined in greater detail in the Discussion section.

Interestingly, although the frequencies for *only* in the spoken parts in ICE-IND, ICE-SIN, ICE-PHI and ICE-HK are relatively low when compared to the results for *also*, the figures for the written parts have not decreased in the same proportion. Both Bhatt (2000) and Lange (2007) have argued that in IndE, the tendency of placing



presentational focus particles after their referent appears only in spoken language and thus a possible explanation for the difference between the results of clause-final *also* and *only* in the written sections of the corpora could be found from the way these particles are used in BrE and AmE. According to Brown (1992: 6) '*also* normally occurs before the main verb (or after a main verb *be*), regardless of the element in the sentence which it qualifies' whereas *only*, when placed after its referent, often sets a formal, authoritarian or declarative tone to the sentence (Brown 1992). Thus it could be argued that the reason why clause-final *only* is used more frequently in the written sections of the corpora is because it is relatively common even in BrE and AmE.

### Distribution of *only*

This section will examine the distribution of the frequencies across the different subsections that can be found from the ICE corpora (see Table A3 in the Appendix). This enables a more detailed comparison between the different varieties and thus helps to identify those genres that favour the use of the clause-final focus particle *only* from those that do not. As mentioned in the previous section, the spoken part of ICE-GB contained lower frequencies of clause-final *only* than its written part, which can also be seen in the results for AmE. For example in ICE-GB, the section 'Scripted monologues' contains 6.3 instances of clause-final *only* per million words, which is almost 13 per cent of all the cases where *only* is used in a clause-final position in the entire corpus. In fact, in ICE-GB there appears to be a positive correlation between the formality of the speech situation and the frequency of the clause-final *only*. Since this is the exact opposite from what can be detected in the four Asian corpora, it could be argued that they carry different meanings and that the tendency of placing *only* at the end of a clause is connected to more formal and literary genres in BrE, where it is used as a restrictive particle. The next highest frequencies can be found from the genre 'Academic writing', which only strengthens the idea that in BrE, clause-final *only* belongs to formal registers, where the rules of argumentation and deduction are followed more closely. In addition to these two, the trait is most prominent in instructional writing, as could be expected, and all corpora contain many instances of the instructional use.

The trend of using *only* in IndE appears to be the reverse of BrE, since the results from ICE-IND show that almost 56 per cent of the cases where *only* is used in a clause-final position occur in 'Private dialogues' (183.1 / 0 in ICE-GB), followed by 'Public dialogues' (ICE-IND 35.2 / 1.6 ICE-GB). The likely explanation is that although clause-final *only* is used sometimes in formal BrE, its use is avoided in similar contexts in IndE, because it has been stigmatized as 'bad English' and therefore it is not used as frequently in formal contexts such as public speeches. The written part of ICE-IND contained only 24 per cent of all the cases of clause-final *only* in the entire corpus and although it still had higher frequencies when compared with the results in the written part of ICE-GB (78.6 / 37.8), it is clear that these two varieties favour the use of *only* differently.

The results for the different spoken categories for the other three Asian corpora are highly interesting; overall, they all follow the same pattern that can be seen in ICE-IND as the highest frequencies for clause-final *only* can be found from 'Private dialogues' and the lowest from 'Scripted monologues'. However, a close analysis of the figures (see Table A3 in the Appendix) reveals some noteworthy details about the varieties. The results for SinE indicate that clause-final *only* is more often used in spoken language, especially

in ‘Private dialogues’, which could be seen as an indicator of the influence of IndE on the variety.

Although the HkE variety produced higher frequencies for the use of *only* when compared with the PhiE variety, the more detailed observation of the frequencies reveals some features that should be discussed. When looking at the percentages for the use of *only* in the spoken categories of the Asian corpora, the order established already by the results on *also* reappears: in spoken IndE, the clause-final *only* is used 76 per cent of the time and the corresponding figures in SinE are 61.7 per cent and 59.7 per cent in PhiE, whereas in spoken HkE the clause-final *only* is used only 51.8 per cent of the time. In addition, since the frequencies for *only* in the written section of ICE-HK are highest out of all the other corpora (81.1), it becomes clear that the results of *only* in HkE are influenced by some other factor besides IndE. One possible explanation for the relatively even distribution of the use of *only* across different categories is that it could be used more as a restrictive particle – this will be discussed in greater detail in the following section.

When the results for *also* are set against the results for *only* (see Tables A2 and A3 in the Appendix) some interesting details appear that are worth noting. As was mentioned earlier, the frequencies for these two particles are almost equally high in the spoken section of ICE-HK (84.4 / 87.1). However, the comparison between the frequencies of these particles in the category ‘Private dialogues’ shows a clear difference between them: the frequency for clause-final *also* is much higher than for clause-final *only* (70.8 / 47.6). These results reinstate *also* as the more frequently used particle in colloquial HkE and thus the variety again follows the trend set already by IndE, SinE and PhiE.

## DISCUSSION

One of the greatest challenges in this study was to enable the comparison between six varieties where the same constructions might be interpreted in different ways. Since the earlier studies on this topic focus solely on IndE, they could rely on the knowledge of IndE stress patterns when analysing the data semantically. This, however, was not deemed to be a valid method for reasons discussed above. As a solution to the problem, I decided to base my statistical analysis of *also* and *only* on syntax. Since the use of clause-final focus particles in BrE and AmE is rare, a decision was made to study solely those cases where these particles were in a clause-final position so that the differences between focus particle uses in IndE, BrE and AmE could be revealed. The aim was to show that by concentrating on the frequencies of clause-final focus particles, the diffusion of the IndE influence into SinE, PhiE and HkE would become visible through the elevated numbers of clause-final focus particles. This method also provided a more reliable way of comparing the results from the different varieties with each other.

One of the major problems in this kind of a comparative study is the linguistic diversity of the Southeast Asian area. For example, Singapore has four official languages: English, Mandarin Chinese, Tamil and Malay, which, in fact, contains particles that follow their referent (see, for example, Goddard 2001). Interestingly, a ‘Malay-lexified Pidgin’ called Bazaar Malay was spoken widely in the Singaporean area as a lingua franca up to the twentieth century, when it began to be replaced by English (Bao and Aye 2010: 155–6). Today, only 13.4 per cent of the Singaporean population are ethnic Malays (Singapore Department of Statistics 2010a) and Lim (2007: 456) has noted that the language is used only rarely outside the Malay speech community. However, Ansaldo (2009: 62),

using Mufwene's linguistic adaptation of the population genetics concept of the *founder principle*, suggests that this former lingua franca of the area could indeed have influenced the SinE variety more strongly. Bazaar Malay has three particles that are relevant for this study: the first two are *pun*, which follows its referent, and *juga*, which is placed in a sentence-final position and they both can be translated as *also* (Aye 2005: 70). The third relevant particle is *saja*, which is also placed in a sentence-final position and can be translated as *only* (Aye 2005: 70). Here it should be noted that the word *saja* itself is of Sanskrit origin (Jones 2011), and this illustrates well the complex history of the Malay language: according to Milner (2009: 24, 38) the Malay language has been heavily influenced by various languages such as Sanskrit during its long history. However, despite the actual origins of the Malay and Bazaar Malay words, they seem to contain particles that could have influenced the results of SinE to some extent.

Both Mandarin Chinese, which has the highest percentage of native speakers among Singaporeans,<sup>11</sup> and Cantonese contain particles that could increase the use of clause-final *only* (see, for example, Matthews and Yip 1994; Lin 2001; Ross and Ma 2006) and *also* (Bao and Hong 2006: 110) in SinE and thus the current influence of these languages on the SinE variety could be expected to be higher than, for example, Malay. In addition, Wee (2004: 1068) notes that speakers of SinE frequently use particles such as *lah* and *lor*,<sup>12</sup> which perform various discourse-pragmatic functions and are placed in a clause-final position. This SinE feature could in itself have reinforced the use of clause-final *also* and *only* in the variety, although it is also possible that the tendency of using these Chinese particles could have caused the SinE variety to adopt the use of clause-final *also* and *only* from IndE more thoroughly.

Despite the fact that over '100 distinct languages' (McFarland 2004: 59), are spoken in the Philippines, the country has only one other official language in addition to English, the Tagalog-based Filipino. Although Tagalog contains particles that could be translated as *also* and *only*,<sup>13</sup> it is unlikely that they have caused the speakers of the PhiE variety to use clause-final focus particles; Ramos and Cena (1990: 98) explain that these particles 'normally occur after the first full word of the sentence', which, according to Schachter (1990: 211), is usually a predicate. Therefore the tendency of using clause-final focus particles in PhiE is not likely to be due to Tagalog influence on the variety. Despite this, the PhiE results do follow the pattern established by IndE although with lower frequencies. Since the use of *also* and *only* in clause-final position has become relatively common in IndE and SinE, it is possible that the growing economic connections that the Philippines has created with these countries has contributed to the spreading of this linguistic feature. After all, English is the only language that is spoken in all three countries and thus increased interaction between the Asian varieties could be the reason why this usage can now also be found from PhiE.

One possible explanation for the difference between the results for HkE and the other three Asian varieties could be found from the homogenous language situation of Hong Kong, as approximately 90 per cent of its population speaks Cantonese (Hong Kong Census and Statistics Department, 2006 Population By-census). According to Goddard (2005: 144–5), Cantonese is rich in clause-final particles<sup>14</sup> and thus the language could have influenced the HkE variety so that it tends to use clause-final particles quite frequently in general, which again would explain the equally high results for *also* and *only* in the spoken section of ICE-HK. Interestingly, Cantonese contains clause-final particles which could be translated as the 'restrictive' *only* (see, for example, Matthews and Yip 1994: 340) and

this could explain why HkE shows a relatively even distribution on the use of clause-final *only* across spoken and written genres; restrictive *only* is used in clause-final position even in British and American English, albeit in more formal contexts, and thus the use of this feature is less constrained than the use of clause-final *also* in HkE. However, this is not the case and since, for example, the frequency of *also* actually exceeds that of *only* in the most colloquial speech genres, one possible explanation for these results could be that a slight IndE influence can be detected in the HkE variety. Despite this, nothing conclusive can be said about this issue due to the low frequencies in the results and the possible influence of Cantonese, which is spoken by the majority of Hong Kongers. Despite the fact that both SinE and HkE have substrate languages that could have contributed to the emergence of this focus particle use, speakers of SinE tend to use clause-final *also* and *only* much more commonly than speakers of HkE. One explanation for this could be that this feature is strengthened in SinE because of the country's strong connections with India and the Indian minority that lives in Singapore.

Although the paragraphs above discuss the influence that other local languages could have had on the results to some extent, the complexity of the issue remains: when dealing with several languages from different language families, it is difficult to determine which words correspond to the English focus particles the most. Conducting a semantic analysis on all the particles in all varieties and all substrate languages would have exceeded the scope of this paper and the results would have been allusive at best. Goddard (2001: 27) describes the problem well when he says that 'it is misguided to attempt to explain the meaning of a discourse particle in one language by reference to discourse particles in a second language . . . because particles tend to have rather language-specific meanings, so that equivalents in other languages tend to be approximate at best'. Since no extensive study on the use of focus particles in SinE, PhiE and HkE has been conducted, the ratio between the influence of IndE and local languages is difficult to estimate. Despite this, the results provide a surprisingly homogenous picture, which supports the view that there is a correlation between the level of interaction between India and other countries and the level in which the Indian use of clause-final *also* and *only* has been adopted in the English varieties spoken in these countries. The results from the Asian corpora show that clause-final focus particles are most frequently used in spoken language and that their frequency decreases as the formality of the speech situation increases. Since this type of detailed analysis based on different text types across many varieties of Asian English has not been done before, the results proved to be of great interest.

The only varieties where the IndE pattern of using focus particles did not emerge were AmE and BrE. The explanation for this might be that in AmE and BrE, focus particles in clause-final position, especially in the case of *only*, belong to more formal language, where their use is constrained by a different set of grammatical rules. Although earlier studies have proven that this potentially innovative way of using focus particles is common in spoken IndE and to some extent also in SinE, this study shows that the trend has also spread to other varieties of Asian English. Because the use of clause-final *also* and *only* has not been described to be a common feature in PhiE or HkE, it is possible that this is a fairly recent development in these two varieties. It should also be noted that the habit of placing focus particles in a clause-final position has not been found to be a feature that is typical for L2 varieties (see for example Kortmann and Szmrecsanyi 2004) and therefore it could be argued to be a local Asian phenomenon and IndE seems to have taken the lead in its use.

In pre-colonial times, Indian culture, religion and the Sanskrit language filtered through Southeast Asian kingdoms, all the way to the islands of the present-day Philippines. As a consequence, India's influence in the Southeast Asian region can be detected in the writing systems and vocabularies of the local languages. Interestingly, a similar pattern could be used to explain the SinE and PhiE results presented in this study: today, as the increased commercial interests between India and Singapore have resulted in a new wave of Indian professionals immigrating to Singapore, the interaction between IndE and SinE has increased. Therefore, the tendency of using clause-final focus particles in SinE may have been influenced by its close contacts with IndE and it is possible that PhiE has, at least partly, adopted these IndE influences indirectly via SinE. Thus, the results of this study show that the influence that India has on the region has not disappeared and today this can be seen in the way it spreads new ideas and linguistics features to the Southeast Asian area.

### CONCLUSION

The aim of this study was to discover if IndE has developed a tendency of placing focus particles *also* and *only* in a clause-final position and whether this use has spread to SinE, HkE and PhiE. Since the English language arrived to India, Singapore and Hong Kong from Britain and to the Philippines from the US, both British and American English were added to the study for point of reference. The results indicate that this syntactic pattern is not common in AmE or BrE, whereas it is clearly used most frequently in IndE, from where it appears to have spread to SinE. A likely explanation for this is the continuous connection that India and Singapore have had for centuries. The case of Hong Kong is slightly different, as the use of English has spread among the population only fairly recently when compared to the other British colonies in the area. This and the relatively homogenous language situation among the population could also explain why IndE influences have not been found from the English variety that is spoken in the Hong Kong area today. The case of the Philippines is more problematic, since the country has not had any clear connections with India since it was colonized by Spain. Despite this, PhiE does show tendencies of adopting the IndE way of using clause-final focus particles, but in lower frequencies and therefore, the spread of IndE influences to PhiE suggests that India's role as a major cultural force in Southeast Asia has survived the test of time.

In order to discover whether this influence is growing or declining, further research on this topic needs to be conducted. For example, a diachronic study on the use of focus particles in these varieties of English could provide the answer to the questions left unanswered here. The scope of this study could also be extended, for example, to other neighbouring countries such as Pakistan, which shares a long common history with India, but whose relations with its neighbour have been strained for decades. It would also be interesting to compare the results on clause-final focus particle use between SinE and MyE because of their long shared history; the results of such a study might reveal some interesting new perspectives on the development of clause-final focus particle use in the region. Another topic for further research could be to examine the semantic qualities of these clause-final focus particles and compare them with the semantic qualities of similar particles used in local languages. Although this would be an exceedingly difficult task, it would also solve some of the questions that have remained unanswered here due to the limited scope of the study.

The extremely low frequencies for both particles in AmE and BrE, suggest that the usage examined in this study could have originated in India and has since spread across Southeast Asia. This, again, gives rise to the question whether the older varieties of BrE and AmE should be used as the standard against which new grammatical features are evaluated, which is something that users of different varieties have done up to this day. After all, it appears that popular linguistic features are emerging around the world and clause-final focus particles such as *also* and *only* are not, in fact, used anymore in Indian English only.

## NOTES

1. Hogue (2001: 167) states that 'the old Anglo-Indian English ... was the language used by British expatriates and Indians during colonial times' and adds that although the current IndE and the old Anglo-Indian English 'are not the same, there is clearly a close relationship'.
2. The research reported here was funded by the Academy of Finland (grant no. 133432).
3. The term 'English' is used here because many of the sources used for this section of the study (for example, König 1991, Nevalainen 1991, Brown 1992) only use the term 'English' or 'general English' without differentiating between different varieties of English.
4. According to Nevalainen (1991: 39), these constraints 'include the focus and semantic scope of the Adverbial ... and its context of use'.
5. It should also be noted that although Indians living outside India speak various other languages the increasingly popular Bollywood films often contain Hindi-Urdu vernacular (Dudrah 2007: 101). In addition Rai (2007: 187) notes that in Singapore, the number of Indian students studying Hindi has increased rapidly in recent years, partly because of the 'dominant position' it has in India.
6. In addition, Hogue (2001: 167) has argued that during the colonial era, most of the words of Indian origin that entered the English that was spoken in the region came from Hindustani and Tamil and thus the influence of these languages on IndE could be stronger.
7. Although the structures of Santa Barbara and Frown differ quite much from the corpora of the ICE family, other researchers such as Collins (2005) have also used these two corpora on AmE in order to compensate for the current unavailability of ICE-US.
8. These were often either NPs or PPs.
9. Although no definite figures on the semantic aspects can be presented here due to the highly interpretative nature of the data, the frequency for the use of the presentational type seems to correlate positively with the frequency of clause-final *also* so that the highest numbers of this use can be found from IndE and the lowest from HkE.
10. Since the structure of the Santa Barbara corpus differs from the ICE corpora, it could not be included in this analysis on different speech genres that favour the use of clause-final *also*.
11. Approximately 36 per cent (Singapore Department of Statistics 2010b).
12. Lim (2007: 469) states that most of these particles are of Cantonese origin.
13. According to Ramos and Cena (1990: 99) the Tagalog words *din* and *rin* can be translated as *too* and thus their meaning is close to the English *also*, whereas the words *lang* and *lamang* can be translated as *only*.
14. *Tim*, according to Kwok (1984: 44), is a Cantonese sentence-final focus particle that can be translated as 'in addition to' and thus its meaning is close to the English word *also*. Matthews and Yip (1994: 340) have also noted that the sentence-final particle *je* can often be translated as *only* and therefore these two particles could have caused the HkE variety to use clause-final *also* and *only* in increased frequencies.

## APPENDIX

Table A1. Number of words in each corpora (using WordSmith5.0)

Word counts	Total	Spoken	Written
ICE-GB	1,061,263	637,682	423,581
ICE-IND	1,114,930	682,572	432,358
ICE-SIN	1,000,127	602,703	397,424
ICE-PHI	1,124,328	678,678	445,650
ICE-HK	1,228,085	734,705	493,380
Santa Barbara (AmE)	1,256,695	250,414	N/A
FROWN (AmE)	1,256,695	N/A	1,006,281

Table A2. The frequency of *also* in each text category per one million words, absolute figures in parentheses and percentages in the right hand columns

Frequency of <i>also</i> / 1 mio words (abs)	Santa Barbara/Frown		ICER-GB		ICE-IND		ICE-SIN		ICE-PHI		ICE-HK	
	freq/ mio (abs)	%	freq/ mio (abs)	%	freq/ mio (abs)	%	freq/ mio (abs)	%	freq/ mio (abs)	%	freq/ mio (abs)	%
SPOKEN												
(S1) Dialogue												
Private	N/A	N/A	0.0 (0)	0	401.4 (247)	56.1	195.8 (118)	81.1	117.9 (80)	62.7	70.8 (52)	70.4
Public	N/A	N/A	0.0 (0)	0	98.2 (67)	13.7	21.6 (13)	8.9	28.0 (19)	14.9	9.5 (7)	9.4
(S2) Monologue												
Unscripted	N/A	N/A	1.6 (1)	25.4	79.1 (54)	11.1	14.9 (9)	6.2	26.5 (18)	14.1	2.7 (2)	2.7
Scripted	N/A	N/A	0.0 (0)	0	42.5 (29)	5.9	1.7 (1)	0.7	0.0 (0)	0	1.4 (1)	1.4
Total (spoken)	63.9 (16)	83.2	1.6 (1)	25.4	622.6 (425)	87.1	233.9 (141)	96.9	172.4 (117)	91.7	84.4 (62)	83.9
WRITTEN												
Non-professional writing	N/A	N/A	2.4 (1)	38.1	41.6 (18)	5.8	0.0 (0)	0	2.3 (1)	1.3	4.1 (2)	4.1
Correspondence	N/A	N/A	2.4 (1)	38.1	9.3 (4)	1.3	2.5 (1)	1.0	11.2 (5)	6.0	8.1 (4)	8.1
(W2) Printed												
Academic	N/A	N/A	0.0 (0)	0	18.5 (8)	2.6	5.0 (2)	2.1	0.0 (0)	0	2.0 (1)	2.0
Non-academic writing	N/A	N/A	0.0 (0)	0	11.6 (5)	1.6	0.0 (0)	0	2.3 (1)	1.3	0.0 (0)	0
Reportage	N/A	N/A	0.0 (0)	0	2.3 (1)	0.3	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0
Instructiona	N/A	N/A	0.0 (0)	0	4.6 (2)	0.6	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0
Persuasive	N/A	N/A	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0
Creative	N/A	N/A	0.0 (0)	0	4.6 (2)	0.3	0.0 (0)	0	0.0 (0)	0	2.0 (1)	2.0
Total (written)	12.9 (13)	16.8	4.7 (2)	74.6	92.5 (40)	12.9	7.5 (3)	3.1	15.7 (7)	8.3	16.2 (8)	16.1
Total (all)	23.1 (29)	100	2.8 (3)	100	417.1 (465)	100	144.0 (144)	100	110.3 (124)	100	57.0 (70)	100

Table A3. The frequency of *only* in each text category per 1 million words, absolute figures in parentheses and percentages in the right hand column

Frequency of <i>only</i> / 1 mio words (abs)	Santa Barbara/Frown		ICER-GB		ICE-IND		ICE-SIN		ICE-PHI		ICE-HK	
	freq/ mio (abs)	%	freq/ mio (abs)	%	freq/mio (absolute)	%	freq/ mio (abs)	%	freq/ mio (abs)	%	freq/ mio (abs)	%
<b>SPOKEN</b>												
(S1) Dialogue												
Private	N/A	N/A	0.0 (0)	0	183.1 (125)	55.9	76.3 (46)	46.5	20.6 (14)	30.9	47.6 (35)	28.3
Public	N/A	N/A	1.6 (1)	3.3	35.2 (24)	10.7	14.9 (9)	9.1	11.8 (8)	17.7	23.1 (17)	13.7
(S2) Monologue												
Unscripted	N/A	N/A	3.1 (2)	6.4	23.4 (16)	7.1	8.3 (5)	5.1	5.9 (4)	8.8	12.2 (9)	7.3
Scripted	N/A	N/A	6.3 (4)	12.9	7.3 (5)	2.2	1.7 (1)	1.0	1.5 (1)	2.2	4.1 (3)	2.4
Total (spoken)	8.0 (2)	33.5	11.0 (7)	22.5	249.1 (170)	76.0	101.2 (61)	61.7	39.8 (27)	59.7	87.1 (64)	51.8
<b>WRITTEN</b>												
(W1) Non-Printed												
Non-professional writing	N/A	N/A	4.7 (2)	9.6	13.9 (6)	4.2	5.0 (2)	3.0	2.2 (1)	3.3	8.1 (4)	4.8
Correspondence	N/A	N/A	4.7 (2)	9.6	18.5 (8)	5.6	10.1 (4)	6.2	6.7 (3)	10.0	14.2 (7)	8.4
(W2) Printed												
Academic	N/A	N/A	7.1 (3)	14.5	16.2 (7)	4.9	7.5 (3)	4.6	0.0 (0)	0	14.2 (7)	8.4
Non-academic writing	N/A	N/A	2.3 (1)	4.7	6.9 (3)	2.1	12.6 (5)	7.7	2.2 (1)	3.3	10.1 (5)	6.0
Reportage	N/A	N/A	2.3 (1)	4.7	2.3 (1)	0.7	5.0 (2)	3.0	2.2 (1)	3.3	2.0 (1)	1.2
Instructional	N/A	N/A	11.8 (5)	24.2	16.2 (7)	4.9	22.6 (9)	13.8	11.2 (5)	16.8	30.4 (15)	18.1
Persuasive	N/A	N/A	2.3 (1)	4.7	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0	0.0 (0)	0
Creative	N/A	N/A	2.3 (1)	4.7	4.6 (2)	1.4	0.0 (0)	0	2.2 (1)	3.3	2.0 (1)	1.2
Total (written)	15.9 (16)	66.5	37.8 (16)	77.5	78.6 (34)	24.0	62.9 (25)	38.3	26.9 (12)	40.3	81.1 (40)	48.2
Total (all)	14.3 (18)	100	21.7 (23)	100	183.0 (204)	100	86.0 (86)	100	34.7 (39)	100	84.7 (104)	100



## REFERENCES

- Ansald, Umberto (2009) *Contact Languages: Ecology and Evolution in Asia*. Cambridge: Cambridge University Press.
- Agnihotri, R.K. (1999) Reports: India: English in India. In Rajendra Singh (ed.), *The Yearbook of South Asian Languages and Linguistics* (pp. 184–197). New Delhi: Sage.
- Asher, R.E. (1985) *Tamil*. Amsterdam: Routledge.
- Aye, Khin Khin Daw (2005) Bazaar Malay: History, grammar and contact. PhD, National University of Singapore.
- Balasubramanian, Chandrika (2009) *Register Variation in Indian English*. Amsterdam: Benjamins.
- Bao, Zhiming and Aye, Khin Khin (2010) Bazaar Malay topics. *Journal of Pidgin and Creole Languages* 25, 155–171.
- Bao, Zhiming and Hong, Huaqing (2006) Diglossia and register variation in Singapore English. *World Englishes* 25, 105–114.
- Bhatt, Rakesh M. (2000) Optimal expressions in Indian English. *English Language and Linguistics* 4, 69–95.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan, and Finegan, Edward (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Brown, Adam (1992) *Making Sense of Singapore English*. Singapore: Federal Publications.
- Collins, Peter (2005) The modals and quasi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide* 26, 249–273.
- Crystal, David (2003) *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.
- Dixon, L. Quentin (2005) Bilingual education policy in Singapore: An analysis of its sociohistorical roots and current academic outcomes. *International Journal of Bilingual Education and Bilingualism* 8, 25–47.
- Dolan, Ronald E (2003) Philippines: A country study. In Harry S. Calit (ed.), *The Philippines, Current Issues and Historical Background* (pp. 25–227). New York: Nova Science.
- Dudrah, Rajinder (2007) Popular culture. In Brij V. Lal, Peter Reeves and Rajesh Rai (eds.), *The Encyclopedia of the Indian Diaspora* (pp. 101–107). Honolulu, HI: University of Hawai'i Press.
- Goddard, Cliff (2001) The polyfunctional Malay focus particle *pun*. *Multilingua* 21, 27–59.
- Goddard, Cliff (2005) *Languages of East and Southeast Asia: An Introduction*. Oxford: Oxford University Press.
- Graddol, David (2010) *English Next: India*. London: British Council.
- Gupta, R.S. (2001) English in post-colonial India: An appraisal. In Bruce Moore (ed.), *Who's Centric Now? The Present State of Post-colonial Englishes* (pp. 148–164). Melbourne: Oxford University Press.
- Hogue, Cavan (2001) The spread of Anglo-Indian words into South-East Asia. In Bruce Moore (ed.), *Who's Centric Now? The Present State of Post-Colonial Englishes* (pp. 165–197). Melbourne: Oxford University Press.
- Hong Kong Census and Statistics Department (2006) Population by-census, population aged 5 and over by usual language, 1996, 2001 and 2006, Retrieved 2 June 2010 from [http://www.byccensus2006.gov.hk/en/data/data3/statistical\\_tables/index.htm#](http://www.byccensus2006.gov.hk/en/data/data3/statistical_tables/index.htm#)
- Jones, Russell (ed.) (2011) SEAlang Library (2011) *Loan-words in Indonesian and Malay*. Compiled by the Indonesian etymological project. Retrieved 1 June, 2011 from <http://sealang.net/indonesia/lwim/>
- Kachru, Yamuna (2006) *Hindi*. Amsterdam: John Benjamins.
- Kortmann, Bernd and Szmrecsanyi, Benedict (2004) Global synopsis: Morphological and syntactic variation in English. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A Handbook of Varieties of English, Vol 2: Morphology and Syntax* (pp. 1142–1202). Berlin: Mouton de Gruyter.
- Koul, Omkar N. (1990) The use of particles in Hindi. *International Journal of Dravidian Linguistics* 19, 22–36.
- Koul, Omkar N. (2008) *Modern Hindi Grammar*. Hyattsville, MD: Dunwoody Press.
- Kwok, Helen (1984) *Sentence Particles in Cantonese*. Hong Kong: Centre of Asian Studies, University of Hong Kong.
- König, Ekkehard (1991) *The Meaning of Focus Particles: A Comparative Perspective*. London: Routledge.
- Lamb, Alastair (1975) Indian Influence in Ancient South-East Asia. In A.L. Basham (ed.), *A Cultural History of India* (pp. 442–454). Delhi: Oxford University Press.
- Lange, Claudia (2007) Focus marking in Indian English. *English World-Wide* 28, 89–118.
- Leech, Geoffrey and Svartvik, Jan (2002) *A Communicative Grammar of English*. London: Longman.
- Lim, Lisa (2007) Mergers and acquisitions: On the ages and origins of Singapore English particles. *World Englishes* 26, 446–473.
- Lin, Hua (2001) *A Grammar of Mandarin Chinese*. Munich: LINCOM Europa.
- Matthews, Stephen, and Yip, Virginia (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.
- McFarland, Curtis (2004) The Philippine language situation. *World Englishes* 23, 59–75.
- McGregor, R. Stuart (1995). *Outline of Hindi Grammar*. Oxford: Oxford University Press.
- Milner, Anthony (2009) *The Malays*. Hoboken, NJ: Wiley-Blackwell.
- Nevalainen, Terttu (1991) *But, Only, Just: Focusing Adverbial Change in Modern English 1500–1900*. Helsinki: Société Néophilologique.
- Pandharipande, Rajeshwari V. (1997) *Marathi*. London: Routledge.
- Platt, John, Weber, Heidi, and Ho, Mian Lian (1983) *Singapore and Malaysia*. Amsterdam: John Benjamins.
- Rai, Rajesh (2007) Singapore. In Brij V. Lal, Peter Reeves and Rajesh Rai (eds.) *The Encyclopedia of the Indian Diaspora* (pp. 176–188). Honolulu, HI: University of Hawai'i Press.

- Rai, Rajesh (2008) 'Positioning' the Indian diaspora: The South-east Asian experience. In Parvati Raghuram, Ajaya Kumar Sahoo, Brij Maharaj, Dave Sangha (eds.), *Tracing an Indian Diaspora: Contexts, Memories, Representations* (pp. 29–51). New Delhi: Sage.
- Ramos, Theresita V., and Cena, Resty M (1990) *Modern Tagalog: Grammatical Explanations and Exercises for Non-native Speakers*. Honolulu, HI: University of Hawaii Press.
- Ross, Claudia and Ma, Jing-heng Sheng (2006) *Modern Mandarin Chinese Grammar: A Practical Guide*. London: Routledge.
- SarDesai, D R. (1997) *Southeast Asia: Past and Present*, 4th edn. Boulder, CO: Westview Press.
- Schachter, Paul (1990) Tagalog. In Bernard Comrie (ed.), *The Major Languages of East and South-East Asia* (pp. 208–230). London: Routledge.
- Schneider, Edgar W. (2007) *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.
- Sharma, Devyani (2003) Discourse clitics and constructive morphology in Hindi. In Miriam Butt and Tracy Holloway King (eds.), *Nominals: Inside and Out* (pp. 59–84). Stanford, CA: CSLI.
- Singapore Department of Statistics. Census of population (2010a) Table 1: Resident Population by Age Group, Ethnic Group, Sex and Residential Status, Retrieved 14 February, 2012 from <http://www.singstat.gov.sg/pubn/popn/c2010sr1/t1-11.pdf>
- Singapore Department of Statistics, Census of population (2010b), Table 50: Resident Population Aged 5 Years and Over by Age Group and Language Most Frequently Spoken at Home, Residential Status and Sex, Retrieved 14 February, 2012 from <http://www.singstat.gov.sg/pubn/popn/c2010sr1/t47-57.pdf>
- Sridharan, Kripa and Sevea, Terenjit S. (2007) The Philippines. In Brij V. Lal, Peter Reeves, Rajesh Rai (eds.) *The Encyclopedia of the Indian Diaspora* (pp. 198–199). Honolulu, HI: University of Hawai'i Press.
- The Freiburg – Brown Corpus of American English (1999) *Englisches Seminar, Albert-Ludwigs-Universität, Freiburg*.
- The International Corpus of English: Great Britain (1998) *The Survey of English Usage, University College London, London*.
- The International Corpus of English: Hong Kong (2006) Department of English, The University of Hong Kong and The Chinese University of Hong Kong. PRC.
- The International Corpus of English: India (2002a) Department of English, Shivaji University, Kolhapur and Freie Universität, Berlin.
- The International Corpus of English: the Philippines Corpus (2004) Department of English and Applied Linguistics, De La Salle University, Manila.
- The International Corpus of English: Singapore Corpus (2002b) Department of English Language & Literature, The National University of Singapore, Singapore.
- The Santa Barbara Corpus of Spoken American English (2002–2005) Santa Barbara Center for the Study of Discourse, University of California, Santa Barbara.
- Verma, Manindar K. (1971) *The Structure of the Noun Phrase in English and Hindi*. New Delhi: Motilal Banarsidass.
- Wee, Lionel (2004) Singapore English: Morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A Handbook of Varieties of English, Vol 2: Morphology and Syntax* (pp. 1058–1072). Berlin: Mouton de Gruyter.

(Received 28 September 2011)

# PUBLICATION II

**The invariant *isn't it* in Asian Englishes**

Hanna Parviainen

World Englishes 35(1), 98–117

<https://doi.org/10.1111/weng.12168>

**Publication reprinted with the permission of the copyright holders.**



## The invariant tag *isn't it* in Asian Englishes

HANNA PARVIAINEN\*

**ABSTRACT:** The aim of this paper is to conduct a quantitative corpus-based analysis in order to detect how frequently the invariant tag *isn't it* is used in the English varieties of India, Singapore, Hong Kong and the Philippines. American and British English were added to the study for point of reference. The data were obtained from the International Corpus of English for all other varieties except for spoken American English, for which the Santa Barbara Corpus of Spoken American English was used. The results indicate that the use of *isn't it* as an invariant tag appears almost solely in spoken language and the highest frequencies can be found in Indian English, followed by Singaporean English, Hong Kong English and Philippine English.

### INTRODUCTION

Recent years have witnessed a growing interest in world Englishes and in the various innovative features speakers around the globe have introduced into their local varieties of English. Asia is home to some of the largest English speaking populations in the world today and therefore, also the Asian varieties of English have received growing scholarly attention. The increasing importance of English as a lingua franca, the accelerating rate of mobilisation and the ever-thickening communication networks have all contributed to the current situation where speakers of different varieties of English are interacting with each other at an increasing pace. Asia has been at the forefront of this development for years and as a consequence, some features are used more frequently among Asian Englishes than in the English varieties spoken in other regions. The invariant tag *isn't it* is a good candidate for this type of development; Mesthrie and Bhatt (2008: 86) have speculated that the use of the invariant tag *isn't it* could have originated from Indian English (IndE) and hence contributed to the increased use of this feature in other Asian Englishes.

Trudgill and Hannah (1985: 111; see also Nihalani et al. 1979: 104) have noted that speakers of IndE use the tag *isn't it* as an invariant question marker, while Balasubramanian (2009), Columbus (2010) and Lange (2012) have also conducted quantitative analyses of the use of this feature in the variety. It is not, however, a feature that is unique to IndE: Kortmann and Szmrecsanyi (2004: 1147, 1192) have noted that invariant tags such as *isn't it* are used frequently in many L2 varieties of English. This view is supported by Columbus (2010), who has examined the use of invariant tags in the English varieties spoken in Britain, New Zealand, India, Singapore and Hong Kong. Despite the existing literature on the use of the invariant tag *isn't it*, the possible origins of the feature in Asian Englishes (such as the possibility of IndE influence put forward by Mesthrie and Bhatt (2008: 86)) has not yet been examined in greater detail.

---

\*School of Language, Translation and Literary Studies, University of Tampere, Kanslerinrinne 1, 33014 Tampere, Finland. E-mail: hanna.e.parviainen@uta.fi

This paper presents a quantitative corpus-based analysis on the use of the invariant tag *isn't it* in IndE, Singapore English (SinE), Hong Kong English (HKE) and Philippine English (PhiE) in order to discuss the spread, frequency and possible origins of the feature in different Asian Englishes. In addition, the American (AmE) and British (BrE) Englishes were added for point of reference.

### TAG QUESTIONS IN ENGLISH

Leech and Svartvik (2002: 132, 382–3) describe tag questions as ‘shortened *yes-no* questions’ which are placed at the end of statements, where their function is to ask the listener to confirm the truth-value of the statement. The form of the tag question is determined by its main clause; if the main clause contains a lexical verb *be*, a modal verb or an auxiliary *be*, *do* or *have*, this verb is repeated in the tag question, but in cases where the main clause does not contain any of the verbs mentioned above, an appropriate form of the verb *do* is used instead (Carter and McCarthy 2006: 547–8). When the subject of the main clause is a pronoun, this pronoun is repeated in the tag question (Huddleston and Pullum 2002: 893), but if the subject of the main clause is not a pronoun, an appropriate personal pronoun must be chosen for the tag question so that it agrees with the subject of the main clause in person, number and gender (Quirk et al. 1985: 810).

Carter and McCarthy (2006: 547) present two different tag question categories which are relevant for this paper:<sup>1</sup> exclamation tags and question tags; examples of the first category are provided in (1) and (2):

- (1) What a beautiful painting it is, *isn't it*? (Carter and McCarthy 2006: 551)
- (2) How odd, *isn't it*? (Quirk et al. 1985: 813)

The structure of exclamation tags is simple; they have interrogative word order and they follow exclamative *wh*-clauses (Carter and McCarthy 2006: 551). Quirk et al. (1985: 813) also add that exclamation tags can be ‘appended to abbreviated verbless exclamations’ as in (2).

The second category, question tags, can be further divided into two categories which Huddleston and Pullum (2002: 892–3) call reversed polarity tags and constant polarity tags; examples of these are provided in (3) to (6):

- (3) She is a teacher, *isn't she*? (Carter and McCarthy 2006: 547)
- (4) I haven't seen you before, *have I*? (Quirk et al. 1985: 810)
- (5) He is ill, *is he*? (Huddleston and Pullum 2002: 892)
- (6) So you call that hard work, *do you*? (Leech and Svartvik 2002: 133)
- (7) \*So he doesn't like his job, *doesn't he*? (Quirk et al. 1985: 913)

The first two, (3) and (4), are examples of reversed polarity tags, which are the most commonly used tag forms in the English language (Huddleston and Pullum 2002: 892; Quirk et al. 1985: 810). In contrast, (5) and (6) provide examples of the more rarely used constant polarity tags, which can take on slightly different meanings when compared with reversed polarity tags. For instance, example (5) shows how constant polarity tags can be used in cases where the speaker has been provided with some new information which s/he agrees with (Huddleston and Pullum 2002: 895). To this Quirk et al. (1985: 812) add that constant polarity tags are ‘characteristically preceded by *oh* or *so*, indicating the speaker's arrival at a conclusion by inference, or by recalling what has already been said’. When

the main clause is preceded by these words, the tag can also take on an emotive meaning of reproach, belligerence, disapproval or even irony (Huddleston and Pullum 2002: 895; Leech and Svartvik 2002: 133), which is illustrated in (6). Interestingly, Quirk et al. (1985: 813) argue that a combination where a negative main clause is followed by a negative tag question as in example (7) 'has not been clearly attested in actual use', although Huddleston and Pullum (2002: 787) note that some speakers do in fact allow its use. However, since Huddleston and Pullum (2002) do not elaborate on this further, it could indeed be argued that the use of the negative form of the constant polarity tag is a rare feature in the language.

## DATA AND METHODS

### Data

The data analysed for the majority of varieties come from the International Corpus of English (ICE 1998, 2002a, 2002b, 2004, 2006), a family of corpora, which contains approximately one million words per corpus (600,000 words spoken, 400,000 words written). Apart from a few exceptions, which are not included in the present study, all the corpora share an identical structure and contain both spoken and written language, which enables a detailed comparison between the different varieties. Because the spoken section of the ICE-US (2012) has not yet been released, the Santa Barbara Corpus of Spoken American English (2002–2005) is also included in the study. The overall word counts for all corpora are presented in Appendix Table A1.

### Methods

The searches for the use of the invariant tag *isn't it* were conducted by using the Wordsmith 5.0 program for all corpora. The program allows the user to go back to the original file of the hit, so that the context of the sentence can be examined when necessary. The searches for the invariant *isn't it* were conducted in three stages.<sup>2</sup> First, all the cases where *isn't it* had been used were collected from the corpora. Second, all instances where *isn't it* functioned as a tag question were identified. Third, all those cases where the subject or verb of the tag *isn't it* did not agree with the subject or verb of the main clause were marked as examples of the invariant use of *isn't it*. Instances where *isn't it* had been used as a constant polarity tag were also calculated as instances of the invariant use, since this combination is not commonly used in the English language. In addition, the data contained sentences where the speaker followed his/her statement with *isn't it*, but this had been marked as the beginning of a new sentence or formed a sentence by itself as in (8):

- (8) D: Who wants to insure Michael Jackson then  
 D: Everybody is so frighten of insuring his concerts then  
 D: *Isn't it* (ICE-SIN:S1A-068#262–4:2:D)  
 A: Ya I mean who wants to go now (ICE-SIN:S1A-068#265:2:A)

After careful examination of the surrounding context of these cases, some were included in the final results, while unclear instances were discarded from the data. Thus, in the third phase, only those sentences where *isn't it* had clearly been used as an invariant tag remained. The number of hits for each corpus was then calculated and the frequencies were normalised per 100,000 words.

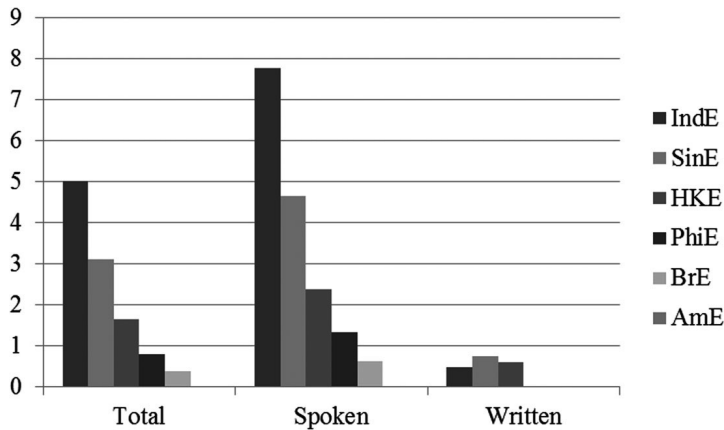


Figure 1. Use of the invariant tag *isn't it*, frequencies normalised per 100,000 words

Table 1. Use of the invariant tag *isn't it*, frequencies normalised per 100,000 words (absolute figures in parentheses)

Invariant <i>isn't it</i>	ICE-IND	ICE-SIN	ICE-HK	ICE-PHI	ICE-GB	SBC / ICE-US
Spoken	7.76 (52)	4.64 (28)	2.38 (17)	1.34 (9)	0.63 (4)	0.0 (0)
Written	0.49 (2)	0.76 (3)	0.61 (3)	0.0 (0)	0.0 (0)	0.0 (0)
All	5.01 (54)	3.10 (31)	1.66 (20)	0.81 (9)	0.38 (4)	0.0 (0)

## RESULTS

### *Frequency of the invariant tag isn't it.*

This section discusses the frequency of the use of the invariant tag *isn't it* in the seven corpora used for this study. Figure 1 and Table 1 present the frequencies for the use of the invariant tag *isn't it* in IndE, SinE, HKE, PhiE, BrE and AmE. As mentioned in the introduction of this paper, there are studies which have examined the use of this tag in different English varieties (see, for example, Tottie and Hoffmann 2006; Columbus 2009, 2010; Balasubramanian 2009), but, the scope of these studies has ranged from one (Balasubramanian 2009) to five (Columbus 2009) varieties of English. Thus, no comparative quantitative study that covers both spoken and written registers of a similar range of Englishes has been conducted before.

The results presented in Table 1 indicate that the invariant tag *isn't it* is used most frequently in IndE, followed by SinE, HKE and PhiE in descending order. Interestingly, some instances can even be found in BrE, whereas AmE did not contain a single instance of this feature. The difference between BrE and AmE is not surprising: according to Tottie and Hoffmann (2006: 287), speakers of BrE use tag questions slightly more frequently than speakers of AmE and thus, finding more invariant tags from the BrE data could be expected. When the spoken and written sections of the corpora are examined separately (see Figure 1 and Table 1), it becomes clear that the use of the invariant tag *isn't it* is more common in spoken language. In fact, in all corpora that contain instances of invariant *isn't it*, the feature is used most frequently in more informal speech situations and the frequencies decrease as the formality of the speech situation increases (see Appendix Table A2).



This could be expected, as the most informal register in the ICE-corpora contains private dialogues, which are frequently characterised by more non-standard patterns of language use.

As was noted earlier, Columbus (2009) has studied the use of invariant tags *isn't it* and others in the category 'Private conversations' of ICE-GB, ICE-NZ, ICE-IND, ICE-SIN and ICE-HK. Interestingly, the absolute figures for BrE, IndE, SinE and HKE in this category (see Appendix Table A2) differ somewhat from the results presented by Columbus (2009: 408). This disparity between the results is likely caused by the difference in the methods of the two studies: while Columbus (2009) focuses solely on those instances of *isn't it* that are in clause-final position, the present study adopted a wider perspective when locating the tags in the data (see the section on methods, above). When the frequencies for the invariant *isn't it* across the spoken sections of the corpora are examined closer, it becomes apparent that the order of varieties stays exactly the same when compared with the overall frequencies of the corpora: the highest frequencies for the use of invariant *isn't it* can be found from IndE, which is again followed by SinE, HKE, PhiE and BrE. Examples from each variety are provided in (9) to (13):

- (9) He is already scale three *isn't it*? (ICE-IND:S1A-058#160:1:B)
- (10) It is not on cats *isn't it*? (ICE-SIN:S1A-061#31:1:A)
- (11) You have kids *isn't it*? (ICE-HK:S1A-050#282:1:A)
- (12) I think the properties are very important *isn't it*? (ICE-PHI:S1A-089#28:1:A)
- (13) It'd be about a half-hour journey *isn't it*? (ICE-GB:S1A-019#243:1:C)

The examples above illustrate the different ways in which the invariant *isn't it* can be employed: for example, the invariant tag does not agree with its subject in number (12), person (11) or gender (9), the verb of the tag does not agree with the verb of the main clause in (13), while (10) is an example of a negative constant polarity tag, which is considered ungrammatical by many speakers of English (Huddleston and Pullum 2002: 892). Although speakers of IndE use the invariant *isn't it* most frequently when compared with the other varieties included in the present study, they rarely use it as a constant polarity tag; in fact, only three such instances were found from the ICE-IND data. Interestingly, SinE and HKE also contained only four instances of negative constant polarity tags and the results thus suggest that even in Asian Englishes, invariant *isn't it* is used more frequently as a reversed polarity tag. Although the British variety did contain a few instances of the invariant *isn't it*, the significantly higher frequencies of ICE-IND, ICE-SIN and ICE-HK imply that the influence of BrE cannot be the sole reason why speakers of these varieties favour the use of this feature.<sup>3</sup> In addition, PhiE has been noted to share some features with AmE due to their historical connections. However, as the results of this paper indicate, AmE influence cannot explain the results of PhiE, as the use of invariant *isn't it* appears to be extremely rare in AmE.

In contrast to the spoken sections, the frequencies for the written sections of the corpora are much lower. The highest frequencies can be found from SinE, followed by HKE and IndE, but as the absolute figures are low, ranging from two to three hits across the three corpora, the differences between the varieties are not notable. Examples from the written sections of the corpora are provided in (14) to (16):

- (14) The name of the place sounds like it is from some part of the African Continent *isn't it*? (ICE-IND:W1B-014#36:1)

- (15) Must be trying to finish up your thesis, *isn't it*? (ICE-SIN:W1B-002#6:1)  
 (16) It sounds good, *isn't it*? (ICE-HK:W1B-010#60:1)

The written sections were included in the study in order to determine if the use of the feature has expanded to more formal contexts in any of the varieties where the feature has been reported to exist: Balasubramanian (2009: 225), who has studied a number of grammatical innovations in IndE, states that the use of innovative features is not always connected to the informality of the genre alone and that some are also found in spoken and written academic IndE. The explanation for this, according to Balasubramanian (2009: 225), is that 'younger users of Indian English, those who didn't necessarily grow up with the same external norm (British English) that older speakers did, are responsible for the spread of Indianisms [...] to the more formal written registers'. Furthermore, Sahgal and Agnihotri (cited in Sedlatschek 2009: 29–30), who have conducted studies on the use of invariant *isn't it* and *no* in IndE in the mid-1980s, conclude that the use of *isn't it* was considered the more acceptable alternative of the two in informal IndE. Therefore, the feature was, as Sedlatschek notes (2009: 30), 'possibly on its way toward further integration'. What is noteworthy about the results of the present paper, however low they may be, is that the majority of hits occur in the category 'Correspondence' (see Appendix Table A2). This is not surprising, as 'Correspondence' is an informal genre and hence contains many grammatical features that are associated with spoken language and face-to-face conversations. In contrast, only one corpus, ICE-IND, contained any hits (the only hit, in fact) in the category 'Non-professional writing', which consist of student essays and examination scripts. Since the feature is only very marginally found in the written genres produced by younger users of English in the data, the results indicate that the use of the invariant *isn't it* is still strongly tied to informal spoken language in all Asian Englishes included in the present study. This observation is also supported by Lange (2012: 199) who states that 'invariant *isn't it* is a non-standard feature that has no place in the register of written/printed IndE'.

#### *Proportion of the invariant tag isn't it*

To ensure that the heightened frequencies of the invariant *isn't it* in Asian Englishes are not caused by a generic elevated use of canonical reversed polarity tags in the data, the *proportion* of the invariant tag out of all canonical reversed polarity tags must be established. In order to do this, all instances where an auxiliary or any form of the verbs *be*, *do* and *have* (including the negated forms) were followed by the pronoun *I*, *you*, *she*, *he*, *it*, *we* and *they* were searched from the corpora and all reverse polarity tags were extracted from the data. This search produced a large number of hits and therefore, the searches were restricted to the most informal category, 'Private conversations', as this section of the data contained the greatest number of hits for invariant *isn't it* in the majority of corpora. Since AmE did not contain any instances of the invariant tag *isn't it*, the variety was excluded from the analysis. Figure 2 and Table 2 present the proportion that the invariant tag *isn't it* has of all reversed polarity tags in the corpora.

The percentages in Figure 2 and Table 2 show that there are clear differences in the proportion that the invariant tag *isn't it* represents out of all reversed polarity tags in the five corpora. The highest percentages can be found in the IndE data, where the speakers favour the invariant tag in almost 50 per cent of the cases when they use reversed polarity

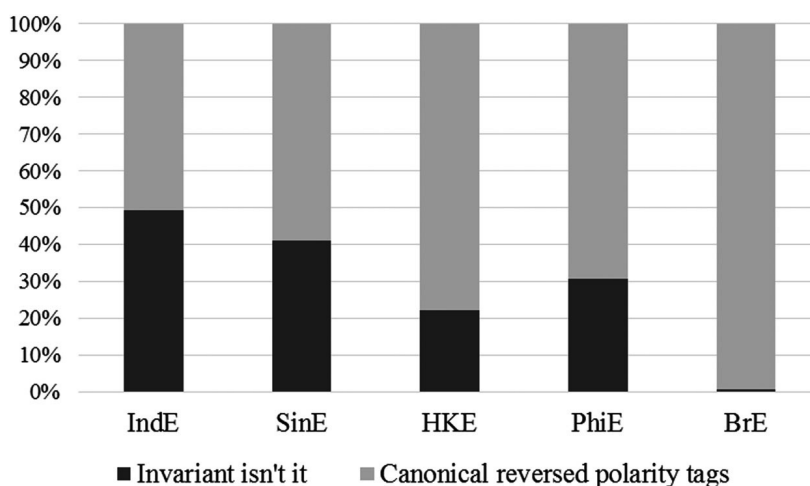


Figure 2. Proportion of the invariant tag *isn't it* from the total number of reversed polarity tags in the category 'Private dialogues'

Table 2. Proportion of the invariant tag *isn't it* from the total number of reversed polarity tags

	ICE-IND		ICE-SIN		ICE-HK		ICE-PHI		ICE-GB	
	absolute figures	%	absolute figures	%	absolute figures	%	absolute figures	%	absolute figures	%
Canonical reversed polarity tags	35	51	26	59	21	78	9	69	395	99
Invariant <i>isn't it</i>	34	49	18	41	6	22	4	31	3	1
All	69	100	44	100	27	100	13	100	398	100

tags. Again, the Indian variety is followed by SinE, where the speakers use invariant *isn't it* in 41 per cent of the cases. Interestingly, the order of varieties presented in Table 2 does not follow the order presented in Table 1 completely, as the percentage of invariant *isn't it* in the PhiE data is actually higher than in the HKE data (31 per cent for PhiE and 22 per cent for HKE). As Table 2 indicates, the absolute numbers of invariant *isn't it* used in the HKE and PhiE data differ only by two hits, but the choice of using the invariant form is almost 10 per cent more likely in PhiE. Therefore, it can be argued that when reversed polarity tags are used in PhiE, the speakers of the variety favour the invariant form more strongly. In contrast to the Asian varieties, the proportion of invariant *isn't it* in the BrE data is negligible. In fact, although BrE and, for example, PhiE differ only by one hit as regards to the number of invariant *isn't it* found in the category 'Private conversations', the absolute number of reversed polarity tags in BrE is actually higher than in all the Asian data combined.

Table 3. The structure of the English invariant tag *isn't it* and the different structures of invariant tag questions (with examples) used in Hindi, (Bazaar) Malay, Mandarin, Hokkien, Cantonese, Tagalog and Cebuano

Tag questions	English	Hindi	(Bazaar) Malay	Mandarin	Hokkien	Cantonese	Tagalog	Cebuano
verb – NEG – PRO	isn't it							
verb – NEG		hai na						
NEG		na	tak					
NEG – Q							hindi ba	di(li) ba
POS – Q				duì ma				
verb – Q				shì ma				
Q				ba		le	ano	ba
verb – NEG – verb				shì bu shì		haih-mhaih		
POS – NEG – POS				hǎo bu hǎo	hoù â m̃ hoù	hóu-mhóu		
POS – or – NEG			ya 'tak					
or – NEG			atau tidak					

### TAG QUESTIONS IN SOUTH AND SOUTHEAST ASIAN LANGUAGES

As the results of the previous section indicate, the superstrate varieties BrE and AmE are unlikely sources for the use of the invariant tag *isn't it* in the Asian varieties examined in this paper. Therefore, other potential origins for the feature in the varieties should be explored. One possible explanation could be the influence of substrate languages spoken in the region, which is why the question formation patterns of the major local languages should be discussed in greater detail. In fact, there are many Asian languages which contain both invariant tags and invariant question markers that are placed after declarative sentences and thus, could have independently caused the tendency of using invariant *isn't it* to emerge in some local varieties of English.

König and Siemund (2007: 296) have noted that the difference between interrogative tags and interrogative particles can be observed in the way the former 'contribute a certain bias by raising expectations toward either a positive or a negative answer'. However, in some languages, invariant tags and invariant question markers can both be used to indicate this bias and in such cases, both markers should be included in the present analysis. Therefore, this section will discuss not only invariant tags but also those invariant question particles which indicate that the speaker expects the listener to provide a positive or a negative answer. Table 3 presents the basic structures of both the invariant tag *isn't it* in English and some of the most common invariant tag questions and question particles that are used in the major languages of India, Singapore, Hong Kong and the Philippines. The explanations for the abbreviations used in this section are provided in Appendix Table A3.

#### India

Hindi is an official language of India and it also has the largest speaker population in the country. According to Koul (2008: 225), tag questions (which Koul refers to as 'leading questions') in Hindi can be formed in three ways: when the expected answer to the question is positive, the tag is formed by repeating the verb of the clause, which is usually in clause-final position, and by adding the word *na* ('no') at the end of the sentence as in (17). In

cases where the speaker expects a negative answer, there are two alternatives: either only the verb is repeated as in (18), or a negative statement is placed before the tag as in (19):<sup>4</sup>

- (17) a:j            garmi: hε,                            hε                            na:?  
       today        hot    be.3SG.PRES        be.3SG.PRES        NEG.Q  
       It is hot today, *isn't it?* (Koul 2008: 225)
- (18) vah        patr        nahī:        parh-e-ga:,                    parh--e-ga:?  
       3SG        letter        NEG        read-3SG-FUT                read-3SG-FUT.Q  
       He won't read a letter, will he? (Koul 2008: 225)
- (19) a:j            garmi: nahī:                    hε,                            na:?  
       today        hot    NEG        be.3SG.PRES                NEG.Q  
       It isn't hot today, is it? (Koul 2008: 225)

Agnihotri (2007: 31) argues that tag questions in Hindi can also be formed by merely adding the phrase *hai na* ('is not') or the particle *na* ('no') to the end of the sentence as (20) and (21):

- (20) vah    kal            baazaar    gay-aa                    thaa,                    hai                    na  
       3SG    yesterday    market    go-3SG.M.PERF    be.SG.M.PERF    be.3SG.PRES    NEG.Q  
       'He went to the market yesterday, didn't he' (Agnihotri 2007: 30)
- (21) tum        kal                    karnaal        jaa-o-ge                    na  
       2PL        tomorrow        karnaal        go-2PL-FUT                NEG.Q  
       'You will go to Karnal tomorrow, won't you?' (Agnihotri 2007: 31)

The invariant use of the tag *hai na* is demonstrated well in (20), where the tense of the verb (*hai*) in the tag differs from the tense of the verb in the main clause (*thaa*). Interestingly, Agnihotri (2007: 31) suggests that due to the tendency of Hindi speakers to use tags such as *hai na* and *na*, 'many speakers of English in India tend to replace the whole range of Standard English tag questions by either "isn't it?" or just "no?"'. Since the structure of the invariant Hindi tag *hai na* ('is not') is close to the structure of the invariant English tag *isn't it* ('is-not it'), substrate influence is a likely explanation for the frequent use of invariant *isn't it* in IndE.

### Singapore

The four official languages of Singapore are English, Mandarin Chinese, Malay and Tamil.<sup>5</sup> According to Hung-nin et al. (1994: 113), all tag questions in Mandarin Chinese are invariant. For the purposes of this study, the most relevant commonly used tag questions are *shì ma* ('is – question particle') and *duì ma* ('right – question particle'), which can both be placed after positive or negative statements. In addition, the question particle *ba* should be included in the present discussion; Po-Ching and Rimmington (2004: 349) have argued that when *ba* is placed at the end of a sentence, it 'conveys a surmise with the speaker presuming that what is stated in the question must or must not be the case'.<sup>6</sup> Furthermore, König and Siemund (2007: 298) note that the construction A-not-A, when placed at the end of a declarative sentence, functions as an interrogative tag. However, König and Siemund (2007: 298) add that the number of these constructions is 'extremely limited and highly lexicalised' and includes such phrases as *shì bu shì* ('is not is'), *hǎo bu hǎo* ('good not

good') and *xíng bu xíng* ('ok not ok'). Thus, it could be argued that there are three question structures that are relevant to the present study. Examples of each are provided in (22) to (24):

- (22) Nǐ            shì            chīsù de            ba  
 2SG        be            vegetarian        PART Q  
 'You are a vegetarian, aren't you?' (Po-Ching and Rimmington 2004: 349)
- (23) Nǐ            hěn            xiǎng jiā,            shì            ma?  
 2SG        very            homesick            be            Q  
 'You are homesick, aren't you?' (Hung-nin et al. 1994: 112)
- (24) Zhè            tiáo            qúnzi            tài            dà,            shì            bu            shì?  
 this        item            skirt            too            big            be        NEG        be  
 'This skirt is too big, isn't it?' (Hung-nin, et al. 1994: 112)

It is possible that the tag *shì bu shì* ('is not is') could have influenced the tendency of Singaporeans to use *isn't it* as an invariant tag, as the first two components of the three in the tag correspond to the first two components of *isn't it*. The question whether such tags as *shì ma* ('is – question particle') correspond equally well to the structure of *isn't it* is more problematic, as finding equivalent English translations for question particles is difficult.

Although Hokkien has never been an official language of Singapore, it functioned as a lingua franca among the Chinese population of Singapore up until the 1970s (Lim 2007: 454) and thus, the language has had a significant influence on the development of SinE. According to Bodman (1955: 22), the A-not-A construction is a commonly used strategy for forming questions in Hokkien. Bodman (1955: 22) calls this 'the sandwich type', where the 'question will consist of the presentation of a choice or alternative'. The question forming patterns most relevant for this study can be formed by adding the following two constructions after a statement: *â boú* ('or not') and *hoù â m̃ hoù* ('good or not good') (Bodman 1955: 22) as (25) and (26) show:

- (25) ciâ +        ũ            lû-tiam            â            boú  
 here        there        be hotel            or            NEG  
 'Are there any hotels here?' (Bodman 1955: 22)
- (26) cīt            kiēng            chài-kuàn            hoù            â            m̃            hoù  
 this        building        restaurant            good            or            NEG        good  
 'Is this a good restaurant?' (Bodman 1955: 22)

Both of these tags have only one structural item in common with *isn't it*, and thus they are unlikely sources for the invariant *isn't it* that is used in SinE.

Baskaran (2004: 1079) presents two ways of forming questions in Malay: by adding the phrase *ya tak* ('yes or-not') or *atau tidak* ('or not') to the end of a sentence as in (27) and (28). Also Bazaar Malay, a form of Malay which was used as a lingua franca in the Singaporean area<sup>7</sup> for centuries (Gupta 1992: 327) before it was replaced by English, has been argued to be a significant input language in the Singaporean area. Khin Khin Aye (2005: 139) argues that Bazaar Malay tag questions are formed by placing one of two negators after a statement clause; in sentences where the VP is the predicate, the speaker will add *tak* ('no') at the end of the clause (29), and when the NP is the predicate,

the negator *bukan* is used. Interestingly, unlike English, Bazaar Malay allows the use of negative tags for negative sentences (Khin Khin Aye 2005: 140):

- (27) Dia makan ya 'tak?  
 3SG eat yes or-NEG  
 'He ate didn't he?' (Baskaran 2004: 1079)
- (28) Dia makan atau tidak?  
 3SG eat or NEG  
 'Did he eat?' (Baskaran 2004: 1079)
- (29) Lu cakap Melayu boleh, tak?  
 2SG speak Malay can, NEG  
 'You can speak Malay, can't you?' (Khin Khin Aye 2005: 139)

As these three tag question structures ('yes or-not', 'or not, 'not') indicate, none of the forms used in these two dialects of Malay resemble the form of the invariant tag *isn't it* closely and thus, it is unlikely that they could have caused the emergence of invariant *isn't it* in the variety.

### *Hong Kong*

The majority of Hong Kong residents speak Cantonese (Census and Statistics Department 2011), where invariant tags are placed at the end of statements which turn them into questions (Matthews and Yip 1994: 317). Kwok (1984: 92) mentions that the question particle *le* 'is used to suggest that the speaker is correct about his assumption about something'; an example of this is presented in (30). In addition, other commonly used tags include many A-not-A constructions such as *haih-mhaih* ('is not-is') and *hóu-mhóu* ('good not-good'), which are exemplified in (31) (Matthews and Yip 1994: 317–8). Kwok (1984: 74–5) notes that these A-not-A constructions can sometimes be followed by additional particles as in (31):

- (30) kau sin go go je da din wa nei le  
 just now that one thing call telephone 2SG Q  
 'That person just called you, didn't he?' (Kwok 1984: 92)
- (31) tej dou di dzai seŋ, haih m haih a  
 hear those children's voice, be NEG be PART  
 I can hear your children's voices. Am I right? (Kwok 1984: 74)

Like Mandarin Chinese, Cantonese also utilises the *haih-mhaih* ('is-not-is') construction, where the two first components of the tag correspond to the English invariant tag *isn't it* closely and hence could have given rise to the tendency to use the invariant *isn't it* in HKE.

### *The Philippines*

Although over a hundred languages are spoken in the Philippine archipelago (McFarland 2004: 59), the country has only two official languages, English and the Tagalog-based Filipino, which has one of the largest speaker populations in the country. According to Ramos and Cena (1990: 84–5), the negative tag in Tagalog is *hindi ba* ('no – question particle', see (32)) which can also be reduced to *di ba* in rapid speech. In addition, Ramos

and Cena (1990: 86–7) note that there are no affirmative tag questions in Tagalog and that a negative statement can be turned into a question by adding the question word *ano* ('what') at the end of a statement as in (33):

- (32) Artista siya, hindi ba?  
 actor 3SG NEG Q  
 'He's an actor, isn't he?' (Ramos and Cena 1990: 85)
- (33) Hindi siya p-um-unta, ano?  
 NEG 3SG go-INF/PAST-go what  
 'He didn't go, did he?' (Ramos and Cena 1990: 86)

Again, the problem of finding suitable English translations for question particles such as *ba* in the construction *hindi ba* arises, which is why it must be concluded that it is unlikely that Tagalog is the source of the invariant tag *isn't it* in PhiE.

Another major language in the Philippines is Cebuano, a language which has only one question marker, *ba* (Marking 2005). According to Marking (2005: 98), this question particle (like other particles used in the language) is 'typically placed after the verb in statements'. When *ba* is preceded by the negative particle *di(li)*, their function is to ask the listener to confirm what the speaker has stated. (34) is an example of this:

- (34) Estudyante siya, dili ba?  
 student 3SG NEG Q  
 He's a student, *isn't he?* (Marking 2005: 97)

The problem with the Cebuano question particle is the same as with the Tagalog question particle above and therefore, it is unlikely that the language could have influenced the patterns in which invariant tag questions are formed in PhiE.

As this section shows, the closest structural equivalence to the English tag *isn't it* can be found from Hindi, although structural resemblances of lesser degree can also be found from Mandarin and Cantonese. A detailed description of the various structures of the tags discussed in this section and how they correspond to the structure of the invariant *isn't it* is provided in Appendix Table A4.

## DISCUSSION

Sharma (2012) presents a useful theoretical framework that states the order in which different explanatory factors should be considered when examining features that are common in new Englishes. Furthermore, Sharma (2012: 219) notes that the framework should only be used when *genuine* similarity in the use of a feature in different varieties has been established. Since the present paper focuses on a single tag form, the invariant *isn't it*, which the results indicate to be used more frequently in Asian Englishes when compared with BrE and AmE, it can be argued that the condition of genuine similarity is met. Explanations for shared features in L2 Englishes can be roughly divided into four main categories, which Sharma (2012: 221) labels as 'Properties shared with the superstrate' (Type A), 'Properties shared with the substrate' (Type B), 'Acquisitional universals' (Type C) and 'General universals' (Type D). Importantly, Sharma (2012: 221) adds that it is 'logical to *first* examine the languages in contact (Types A and B) for proximate causes of shared



features before appealing to more general motivations' of the C and D type; this order will also be followed here.

The three explanations that Sharma (2012: 222) lists under Type A include 'Founder effects', 'Exogenous prestige-driven mimicry' and 'Diffusion from one variety to another'. As the results of this study indicate, the use of the invariant tag *isn't it* is clearly more common in the Asian varieties than in BrE and AmE and therefore, the 'Founder effect' is an unlikely explanation for the higher frequencies present in the Asian varieties; the same argument can be made against 'Exogenous prestige-driven mimicry', since neither the older prestige variety (BrE) nor the new prestige variety (AmE) have this feature. However, 'Diffusion from one variety to another' is a possibility that should be explored further.

The influence of India in the Southeast Asian region is not a novel idea as the country's impact on the cultures of Southeast Asia can be traced back for hundreds of years (Lamb 1975: 442; SarDesai 1997: 16; Rai 2008: 29–30). Before the era of colonisation, the number of Indians living in Southeast Asia remained low, but the situation changed dramatically after the arrival of the British, when many Indians moved to the newly established British colonies in the Southeast Asian area to work as indentured servants, governmental clerks and even English teachers. Today, the results of this large-scale migration can be seen in the demographics of the former colonies, as many countries, especially Singapore and Malaysia, still have large Indian minorities (see, for example, Kaur 2007; Rai 2008). Some scholars such as Platt et al. (1983: 8) have already noted that the presence of English speaking Indians in such respected professions as clerks and teachers could have influenced the development of the variety of English that was emerging in Singapore and Malaysia. As the results of the present paper show, the order of varieties favouring the use of invariant *isn't it* declines as one moves from IndE to SinE, HKE and PhiE. Interestingly, this order of varieties resembles the results of an earlier study to some degree: Hogue's (2001) examination of the spread of Anglo-Indian words in Southeast Asia during colonial times shows that the highest number of loan words is found from Malaysian English followed by SinE, HKE and PhiE, while the lowest number of loanwords can be found from International English. It is also worth noting that Parviainen (2012), who examined the IndE use of clause-final focus particles *also* and *only*, argued that SinE and possibly also PhiE have adopted the use of these particles from IndE. Therefore, IndE influence on the Southeast Asian varieties of English is a possibility; this is also acknowledged by Mesthrie and Bhatt, who note the following:

It is not clear whether the Welsh English invariant form *isn't it* served as a model or this tag, in some new Englishes at least; or whether the Hindi-Urdu invariant tag *na* played a role. It is also possible that once adopted, the Indian English form *isn't it* played a subsequent influential role in other [world Englishes] (Mesthrie and Bhatt 2008: 86).

However, it is difficult to say whether the invariant *isn't it* was already used by the ethnic Indians living in Southeast Asia during colonial times, or whether the feature emerged later in the variety. In order to provide any definitive answers to the *when*-question, one would need diachronic data, which unfortunately is not currently available.<sup>8</sup>

Sharma (2012: 23) divides Type B, 'Properties shared with the substrate', into two subtypes, 'Accidental resemblance' and 'Areal convergence'. As the discussion in the previous section shows, all of the substrate languages spoken in India and Southeast Asia

examined for this paper contain invariant tags, although they belong to three different language families (Hindi to Indo-European, Mandarin, Cantonese and Hokkien to Sino-Tibetan and Malay, Tagalog and Cebuano to the Austronesian language family). Interestingly, the Southeast Asian area has for centuries been a 'melting pot' of different cultures and languages and, as a consequence, areal convergence can be observed in the languages spoken in the region. Therefore, 'Areal convergence' is the more likely explanation of the two, although more detailed investigations into the evolution of invariant question tags and particles in the local languages are still needed. Interestingly, Dryer (2007: 93) argues that '[m]any languages employ particles that occur in *leading questions*, in which the speaker makes an assumption as to what the answer will be, with a function analogous to the tag in English questions like *Mary is here, isn't she?*'. Furthermore, Dryer (2007: 93) notes that 'such markers of leading questions appear to exhibit a tendency to occur at the end of sentences, regardless of the order of object and verb', which is something that can also be observed in the substrates discussed in this paper. Therefore, if there is attested widespread use of invariant tag questions in languages spoken around the world, as Dryer (2007: 93) implies, it might be tempting to argue that invariant *isn't it* might, after all, be feature that qualifies as a universal of new Englishes. However, Sharma advises against jumping into such conclusions:

In some cases, substrates may share a common trait because it is typologically unmarked. The resurgence of this trait in the offspring Englishes nevertheless constitutes substrate transfer and not clear intervention of a universal. True emergence of universals would only be certain if the substrates do *not* have a particular unmarked feature, and yet the feature arises in offspring varieties of English (Sharma 2012: 223).

As was already mentioned, the closest equivalent to the structure of the invariant *isn't it* can be found in Hindi, but some structural resemblance is also found in the tags of Mandarin Chinese and Cantonese. Interestingly, SinE favours the use of invariant *isn't it* more strongly when compared with HKE, although the same structure ('is not is') is used in the substrates of both varieties. Some studies (see, for example, Mukherjee and Gries 2009; Hoffmann et al. 2014) have shown that the level of nativisation of a linguistic feature in a postcolonial variety could in some cases be linked to the developmental stage it has reached in Schneider's (2007) dynamic model.<sup>9</sup> Since SinE has advanced further in the dynamic model when compared to Hong Kong (SinE has reached stage 4, the endonormative phase while HKE is still in stage 3, the nativisation phase), it might explain the differences between the two varieties. However, Gries and Mukherjee (2010: 542) state that 'a neat alignment of evolutionary stages on the one hand and linguistic features [...] on the other can only be found at the level of rather abstract linguistic configurations based on a wide range of linguistic forms'. As the focus of the present study is on a single feature, the invariant tag *isn't it*, the evolutionary stages of SinE and HKE provide an unlikely explanation for the difference between the varieties. A more plausible explanation for the differences between SinE and HKE is that Singapore's proximity (not only geographically, but also culturally) to India has led to the situation where SinE speakers are in closer interaction with speakers of IndE; hence, some linguistic patterns that are used frequently in IndE become more strongly favoured also in SinE. This in turn leads to the conclusion that both Type A and B are at work in the case of SinE. Sharma (2012: 223) states that 'in situations of overlap, where we find evidence of both Type A and B effects, further

varieties may need to be added to the comparison, or reinforcement may be arising from both sources'. For this study, the presence of two additional Asian varieties (PhiE and HKE) and the lower, but still significant uses they show for the feature strengthen the argument that in the case of SinE, the interaction of Type A and B is the more likely explanation.

It is difficult to say what might explain the results of HKE and PhiE, the former having higher frequencies of use, while the proportion of invariant *isn't it* out of all reversed polarity tags was greater for the latter. Unlike PhiE, HKE had some instances of the feature also in the written category 'Correspondence' and thus it could be argued that the use of the feature has spread further in HKE. Both varieties have substrates where invariant tags are used and thus, the mere existence of these invariant tags in the substratum can give rise to the tendency to use invariant tags also in the local varieties of English. Because *isn't it* is not the only invariant tag used in English varieties (see, for example, Columbus 2009), it may be that the speakers of new Englishes tend to favour those invariant tags in their local English variety which are structurally closer to the tags of substrate languages. The differences between PhiE and HKE can thus be explained as Type B, influence of the substrate: the main substrate of HKE, Cantonese, has tags that are structurally closer to *isn't it* than the tags of PhiE substrates and therefore, speakers of HKE use *isn't it* more frequently.

Sharma notes (2012: 221, 224) that 'in the absence of either Type A or B explanations, it is reasonable to next consider the possibility of SLA effects (Type C) before assuming the wider operation of genuine universals (Type D)'. Since Type A and B could together explain the use of the invariant tag *isn't it* in Asian Englishes, the other two explanations (Type C and D) will not be elaborated here. Here it should be acknowledged that the use of invariant *isn't it* has been argued to exist also in other Englishes spoken in, for example, Africa. African varieties were excluded from the present paper due to the unavailability of extensive comparable data, as many of the ICE-corpora on African Englishes are still in the process of being compiled. Should a future study indicate that invariant *isn't it* is used frequently in African Englishes and that substrate influence cannot explain this development, only then should the possibility of second language acquisition and genuine language universals be reconsidered as possible explaining factors also in the case of Asian Englishes.

## CONCLUSION

This study examined the use of the invariant tag *isn't it* in six varieties of English spoken in India, Singapore, Hong Kong and the Philippines while BrE and AmE were added to the study for point of reference. Since the invariant tag *isn't it* is used most frequently in spoken IndE and an invariant tag with a similar structure can be found from the Indian substratum, it is likely that this feature has developed independently in the variety. As for the other Asian Englishes studied here, the possible origins of the feature are more varied. The findings of this paper support the conclusion that IndE is, if not the initiator of this development, a contributor to the popular usage of the feature in Asia. This possibility is also supported by the results of two earlier studies conducted on similar topics. Although the results of the present study show that the use of the invariant tag *isn't it* is common in many Asian Englishes and that the roots of the feature are located in Asia, this issue

should be examined further by, for example, doing a similar study with the major L2 varieties of English in Africa. Another interesting topic would be to conduct a diachronic study, as this could shed more light into the development of IndE's influence in the region; although Hogue (2001) has dated the spread of Anglo-Indian words to colonial times, estimations of the time when this particular feature evolved in the region are difficult to make.

### ACKNOWLEDGEMENTS

The research reported here was funded by the Academy of Finland (Change Tampere, grant no. 269114). I would also like to thank the anonymous referees for their valuable comments and suggestions. All mistakes are naturally my own.

### NOTES

1. Carter and McCarthy (2006: 547) also introduce two other groups: statement tags such as 'I'm hungry, I am', and directive tags such as 'Open the door, will you', but since the tag in the former category is not a question and because the tag *isn't it* cannot be used in the latter category, these will not be discussed further in this paper (for more information, see Quirk et al. 1985: 813).
2. According to the current recommendations on the ICE homepage, the following sections were excluded from the data: 'extra corpus' text (marked <X>...</X>), 'editorial comments' (marked <&>...</&>) and 'untranscribed' text (marked <O>...</O>) (<http://ice-corpora.net/ice/importantnote.htm>). In addition, some of the data especially in ICE-IND contained misspelled or mispronounced words which were followed by the transcriber's corrections (marked <+>...</+>). As these sections were not produced by the original speaker/writer, they were also excluded from the searches.
3. Interestingly, Columbus (2009: 410) has noted that the IndE pattern of invariant clause-final tag use seems to be an extension of the BrE pattern.
4. The markings of this and all subsequent examples presented in this section can differ from the originals as the sentences have been harmonised for this paper by the author.
5. Lim (2007: 456) argues that Tamil 'plays no real role in the contact dynamics [of Singapore]' and since Tamil is spoken only by approximately 3 per cent of the population (Census of Population 2010, Singapore Department of Statistics), Tamil has been left out of the present discussion.
6. Po-Ching and Rimmington (2004: 350) also note that in cases where the question particle *ma* is placed at the end of a negative sentence, the particle expresses 'a degree of positive surmise.'
7. According to Lim (2007: 456) Bazaar Malay is still being used by the older generations and possibly by people belonging to lower social class, but overall, the use of the language is declining.
8. Interestingly, there are some indications that the IndE tendency to use invariant *isn't it* could also have influenced one type of English spoken outside Asia: Andersen (2001: 100), who has studied the use of another invariant tag, *innit*, says that 'there are good reasons for claiming that the use of *innit* ... as invariant tags and follow-ups in London teenage English is largely due to the ethnic minority speakers in this area' (see also Torgersen et al. 2011: 112). Jamaican Creole has been argued to have the strongest influence (see, for example, Hewitt 1986, cited in Andersen 2001: 113), although, quite interestingly, Patrick (2004: 419) has noted that the feature occurs only in BrC but not in JaC. Andersen (2001: 113) states that the biggest ethnic minorities in London are 'Indian, Black-Caribbean, Black-African, Pakistani and Bangladeshi groups' and thus, it could be argued that the large Indian minorities may also have contributed to the growing use of invariant *innit* in the London area.
9. The five phases are called the foundation phase, exonormative phase, nativisation phase, endonormative phase and differentiation phase (Schneider 2007).

## APPENDIX

Table A1. Word counts for all corpora

Word counts	Total	Spoken	Written
ICE-IND	1,077,618	670,064	407,552
ICE-SIN	998,599	603,175	395,424
ICE-HK	1,204,636	713,757	490,879
ICE-PHI	1,115,501	670,630	444,871
ICE-GB	1,061,263	637,682	423,581
SBC (spoken AmE)	–	250,414	N/A
ICE-US (written)	675,514	N/A	425,100

Table A2. Frequency of invariant isn't it across different context categories of ICE normalised per 100,000 words (absolute figures in parenthesis)

		ICE-IND	ICE-SIN	ICE-HK	ICE-PHI	ICE-GB	SBC/ ICE-US
Spoken	Private	5.07 (34)	2.98 (18)	0.84 (6)	0.60 (4)	0.47 (3)	0
	Public	2.09 (14)	1.49 (9)	1.54 (11)	0.60 (4)	0.16 (1)	0
	Non-scripted	0.60 (4)	0.17 (1)	0	0.15 (1)	0	0
	Scripted	0	0	0	0	0	0
Written	Non-Professional	0.25 (1)	0	0	0	0	0
	Correspondence	0.25 (1)	0.76 (3)	0.76 (3)	0	0	0
	Academic	0	0	0	0	0	0
	Non-Academic	0	0	0	0	0	0
	Instructional	0	0	0	0	0	0
	Persuasive	0	0	0	0	0	0
	Fiction	0	0	0	0	0	0

Table A3. Abbreviations

FUT	Future tense	PL	Plural
INF	Infinitive	POS	Positive attribute (usu. adjective)
M	Masculine	PRES	Present tense
NEG	Negation	PRO	Pronoun
PAST	Past tense	Q	Question particle
PART	Particle (other than Q)	SG	Singular
PERF	Perfective		

Table A4. Summary of all invariant tags. Components of substrate tags corresponding to the components of *isn't it* are in bold

Language	Tag	Structure	Components of invariant tags					PRON	VERB	Q	POS
			POS	CONJ	VERB	NEG					
English	isn't it	is – NEG – it			be.3SG.PRES	NEG	it.3SG				
Hindi	hai na*	is – NEG			be.3SG.PRES	NEG					
Mandarin	shì bu shì	is – NEG – is			be	NEG		be			
Cantonese	haih-mhaih	is – NEG – is			be	NEG		be			
Mandarin	shì ma	is – Q			be					ma	
Tagalog	(hin)di ba	NEG – Q				NEG				ba	
Cebuano	dí(lí) ba	NEG – Q				NEG				ba	
Hindi	na	NEG				NEG					
Bazaar Malay	tak	NEG				NEG					
Bazaar Malay	bukan	NEG				NEG					
Malay	atau tidak	or – NEG		or		NEG					
Hokkien	â boú	or – NEG		or		NEG					
Malay	ya 'tak	yes – or – NEG	yes	or		NEG					good
Hokkien	hoù â m'hoù	good – or – NEG – good	good	or		NEG					good
Mandarin	hǎo bu hǎo	good – NEG – good	good			NEG					good
Cantonese	hóu-mhóu	good – NEG – good	good			NEG					good
Mandarin	xíng bu xíng	ok – NEG – ok	ok			NEG					ok
Mandarin	duì ma	right – Q	right							ma	
Mandarin	ba	Q								ba	
Cantonese	le	Q								le	
Tagalog	ano	Q								what	
Cebuano	ba	Q								ba	

\*Both the invariant form hai na and the form where the verb of the main clause is repeated in the tag are used.  
 k – NEG – okO in Hindi ('bove, the closest resemblance to the structure of the invariant

## REFERENCES

- Agnihotri, Rama Kant. 2007. *Hindi: An essential grammar*. London and New York: Routledge.
- Andersen, Gisle. 2001. *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Philadelphia, PA: John Benjamins.
- Balasubramanian, Chandrika. 2009. *Register variation in Indian English*. Amsterdam: John Benjamins.
- Baskaran, Loga. 2004. Malaysian English: Morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A handbook of varieties of English, Vol. 2: Morphology and syntax*, 1073–1085. Berlin: Mouton de Gruyter.
- Bodman, Nicholas. 1955. *Spoken Amoy Hokkien*. Kuala Lumpur: Government of the Federation of Malaya.
- Carter, Ronald & Michael McCarthy. 2006. *Cambridge grammar of English*. Cambridge: Cambridge University Press.
- Census and Statistics Department. Hong Kong Statistics. 2011. 2011 Population census, Population aged 5 and over by usual language, 2001, 2006 and 2011. <http://www.census2011.gov.hk/en/index.html>. (8 January, 2012.)
- Columbus, Georgie. 2009. A corpus-based analysis of invariant tags in five varieties of English. In Antoinette Renouf and Andrew Kenoe(eds.), *Corpus linguistics: Refinements and reassessments*, 401–414. Amsterdam: Rodopi.
- Columbus, Georgie. 2010. A comparative analysis of invariant tags in three varieties of English. *English World-Wide* 31(3). 288–310.
- Dryer, Matthew S. 2007. Word order. In Timothy Shopen (ed.), *Language typology and syntactic description, Vol. 1: Clause structure*. 2nd edn. 61–131. Cambridge: Cambridge University Press.
- Gries, Stefan T. & Joybrato Mukherjee. 2010. Lexical gravity across varieties of English: An ICE-based study of *n*-grams in Asian Englishes. *International Journal of Corpus Linguistics* 15(4). 520–548.
- Gupta, Anthea Fraser. 1992. Contact features of Singapore colloquial English. In Kingsley Bolton and Helen Kwok (eds.), *Sociolinguistics today: International perspectives*, 323–345. London: Routledge.
- Hoffmann, Sebastian, Anne-Katrin Blass & Joybrato Mukherjee. 2014. Canonical tag questions in Asian Englishes: forms, functions and frequencies in Hong Kong English, Indian English and Singapore English. In Markku Filppula, Juhani Klemola and Devyani Sharma (eds.), *Oxford handbook of world Englishes*. Oxford and New York: Oxford University Press.
- Hogue, Cavan. 2001. The spread of Anglo-Indian words into South-East Asia. In Bruce Moore (ed.), *Who's centric now? The present state of postcolonial Englishes*, 165–197. Oxford: Oxford University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hung-nin Samuel Cheung, Sze-yun Liu & Li-lin Shih. 1994. *A practical Chinese grammar*. Hong Kong: The Chinese University Press.
- Kaur, Amarjit. 2007. Malaysia. In Brij V. Lal, Peter Reeves & Rajesh Rai (eds.), *The encyclopedia of the Indian diaspora*, 156–167. Honolulu: University of Hawai'i Press.
- Khin Khin Aye. 2005. *Bazaar Malay: History grammar and contact*. National University of Singapore PhD Thesis.
- Kortmann, Bernd & Benedict Szmrecsanyi. 2004. Global synopsis: Morphological and syntactic variation in English. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A handbook of varieties of English, Vol. 2: Morphology and syntax*, 1142–1202. Berlin: Mouton de Gruyter.
- Koul, Omkar N. 2008. *Modern Hindi grammar*. Hyattsville, MD: Dunwoody Press.
- Kwok, Helen. 1984. *Sentence particles in Cantonese*. Hong Kong: Center of Asian Studies, University of Hong Kong.
- König, Ekkehard & Peter Siemund. 2007. Speech act distinctions in grammar. In Timothy Shopen (ed.), *Language typology, a syntactic description*, Vol. 1, 2nd edn., 276–324. Cambridge: Cambridge University Press.
- Lamb, Alastair. 1975. Indian Influence in Ancient South-East Asia. In A.L. Basham (ed.), *A cultural history of India*, 442–454. Oxford: Oxford University Press.
- Lange, Claudia. 2012. *The syntax of spoken Indian English*. Amsterdam: John Benjamins.
- Leech, Geoffrey & Jan Svartvik. 2002. *A communicative grammar of English*. London: Longman.
- Lim, Lisa. 2007. Mergers and acquisitions: On the ages and origins of Singapore English particles. *World Englishes* 26(4). 446–473.
- Marking, Tom. 2005. *Cebuano study notes*. <http://www.tomandcathymarking.com/cebuano/CebuanoStudyNotes.pdf>. (5 July, 2013.)
- Matthews, Stephen & Virginia Yip. 1994. *Cantonese: A comprehensive grammar*. London: Routledge.
- McFarland, Curtis. 2004. The Philippine language situation. *World Englishes* 23(1). 59–75.
- Mesthrie, Rajend & Rakesh M. Bhatt. 2008. *World Englishes*. Cambridge: Cambridge University Press.
- Mukherjee, Joybrato & Stefan T. Gries. 2009. Collostructional nativisation in New Englishes: Verb-construction associations in the International Corpus of English. *English World-Wide* 30(1). 27–51.
- Nihalani, Paroo, R.K. Tongue & Priya Hosali. 1979. *Indian and British English: A handbook of usage and pronunciation*. New Delhi: Oxford University Press.
- Patrick, Peter L. 2004. Jamaican Creole: Morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A handbook of varieties of English, Vol. 2: Morphology and syntax*, 407–437. Berlin: Mouton de Gruyter.
- Parviainen, Hanna. 2012. Focus particles in Indian English and in other varieties. *World Englishes* 31(2). 226–247.

- Platt, John, Heidi Weber & Mian Lian Ho. 1983. *Singapore and Malaysia*. Amsterdam: John Benjamins.
- Po-Ching, Yip & Don Rimmington. 2004. *Chinese: A comprehensive grammar*. London: Routledge.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Rai, Rajesh. 2008. 'Positioning' the Indian diaspora: The South-east Asian experience. In Parvati Raghuram, Ajaya Kumar Sahoo, Brij Maharaj & Dave Sangha (eds.), *Tracing and Indian diaspora: Contexts, memories, representations*, 29–51. New Delhi: Sage.
- Ramos, Teresita V. & Resty M. Cena. 1990. *Modern Tagalog*. Honolulu, HI: University of Hawaii Press.
- SarDesai, D R. 1997 *Southeast Asia: Past & present*, 4th edn. Boulder: Westview Press.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Sedlatschek, Andreas. 2009. *Contemporary Indian English: Variation and change*. Amsterdam: John Benjamins.
- Sharma, Devyani. 2012. Shared features in new Englishes. In Raymond Hickey (ed.), *Areal features of the Anglophone world*, 211–232. Berlin: De Gruyter Mouton.
- Singapore Department of Statistics. 2010. *Census of population 2010, Table 48: Resident population aged 5 years and over by language most frequently spoken at home, ethnic group and sex*. [http://www.singstat.gov.sg/publications/publications\\_and\\_papers/cop2010/census10\\_stat\\_release2.html](http://www.singstat.gov.sg/publications/publications_and_papers/cop2010/census10_stat_release2.html). (3 September, 2011.)
- The International Corpus of English. <http://ice-corpora.net/ice/importantnote.htm>. (24 November, 2011.)
- The International Corpus of English: Great Britain. 1998. *The Survey of English Usage*. University College London, London.
- The International Corpus of English: Hong Kong. 2006. Department of English, The University of Hong Kong and The Chinese University of Hong Kong. PRC.
- The International Corpus of English: India. 2002. Department of English, Shivaji University, Kolhapur and Freie Universität, Berlin.
- The International Corpus of English: the Philippines. 2004. Department of English and Applied Linguistics, De La Salle University, Manila.
- The International Corpus of English: Singapore. 2002. Department of English Language & Literature, The National University of Singapore, Singapore.
- The International Corpus of English: USA. 2012. University of Massachusetts-Boston, Boston, University of California, Los Angeles and the University of California, Santa Barbara.
- The Santa Barbara Corpus of Spoken American English. 2002–2005. Santa Barbara Center for the Study of Discourse, University of California, Santa Barbara.
- Torgersen, Eivind, Costas Gabrielatos, Sebastian Hoffman & Susan Fox. 2011. A corpus-based study on the pragmatic markers in London English. *Corpus Linguistics and Linguistic Theory* 7(1), 93–118.
- Tottie, Gunnel & Sebastian Hoffmann. 2006. Tag questions in British and American English. *Journal of English Linguistics* 34(4), 283–311.
- Trudgill, Peter & Jean Hannah. 1985. *International English: A guide to varieties of standard English*. 2<sup>nd</sup> edn. London: Edward Arnold.

(Received 11 February 2015)



# PUBLICATION

## III

### **Omission of direct objects in New Englishes**

Hanna Parviainen

Changing English: Global and Local Perspectives, 129–153  
<https://doi.org/10.1515/9783110429657-008>

**Publication reprinted with the permission of the copyright holders.**



Hanna Parviainen

# Omission of direct objects in New Englishes

**Abstract:** This chapter examines the omission of direct objects of transitive verbs as in A: *Do you know Malayalam?* B: *Oh yes I speak.* (ICE-IND) in contexts where the verb does not function intransitively. The English varieties examined for this study come from Fiji, Hong Kong, India, Jamaica, Kenya, the Philippines and Singapore and the superstrate varieties from Britain and America were also included in the data. The study analyses the frequency of the phenomenon in sentences where the verbs *bring*, *buy*, *enjoy*, *find*, *give*, *love*, *make*, *offer* and *show* are used transitively. The focus of this quantitative study is on spoken language while possible substrate influences on the varieties are also discussed. The data for Fiji, Hong Kong, Indian, Jamaican, Kenyan, Philippine, Singaporean and British English were obtained from the spoken sections of the *International Corpus of English* (ICE), whereas American English was studied by using the *Santa Barbara Corpus of Spoken American English* (SBCSAE). The results of the study indicate that the tendency to omit direct objects is strongest in IndE and SinE, while the feature was rarest in BrE, JaE and AmE.

**Keywords:** transitivity, structural vs semantic, transitive verb, monotransitive vs ditransitive, direct object, omission of direct object, substrate vs superstrate language, Founder Principle, indigenous language, agreement system, topic-comment, topic-prominence, universal, acquisitional, diffusion, areal convergence

## 1 Introduction

The tendency to omit objects is a common feature in many languages spoken around the world, but widespread use of the feature has not been attested in traditional inner circle varieties such as British (BrE) and American English (AmE). In contrast, Platt et al. (1984: 117) have noted that “there is a tendency [in New Englishes], particularly in colloquial speech, to imply the subject or the object pronoun of a sentence rather than state it explicitly.” More detailed descriptions of the spread of the feature include the English varieties of Hong Kong (HKE) (Platt 1982: 410), India (IndE) (Subbārão 2012: 28), Malaysia (Baskaran

---

Hanna Parviainen, University of Tampere

DOI 10.1515/9783110429657-008

2004: 1080), Sri Lanka (Platt et al. 1984: 117), Singapore (SinE) and the Philippines (PhiE) (Platt et al., cited in Bhatt 2004: 1026). In addition, Platt et al. (1984: 117) have stated that the feature is also used in “African Englishes”, though they do not elaborate on this issue further.

According to Kortmann and Lunkenheimer’s (2013, <http://www.ewave-atlas.org/parameters/42>) *Electronic World Atlas of Varieties of English*<sup>1</sup> (eWAVE), the omission of object pronouns is “pervasive or obligatory” in Maltese English, Kenyan English (KenE), IndE, HKE, “Colloquial Singapore English”, Malaysian English, Rope River Creole, “pure Fiji English”<sup>2</sup> and Tok Pisin. In addition, Englishes where the feature is argued to be “neither pervasive nor extremely rare” include such varieties as PhiE, Butler English and Sri Lankan English (ibid.). Interestingly, despite these estimations of the frequency of the feature, no quantitative studies based on comparative data have yet been conducted which would examine the exact spread and frequency of this feature in New Englishes.

This chapter presents a quantitative study of the frequency with which direct objects are omitted in Fijian English (FjE), HKE, IndE, Jamaican English (JaE), KenE, PhiE, and SinE. In addition, British (BrE) and American (AmE) Englishes have been included as the ‘source’ varieties of the other Englishes examined here. The main focus of the analysis is on the quantitative analysis of the data, although some qualitative aspects will also be addressed when considered relevant.

## 2 Transitivity in English

Kittilä (2010: 346) defines linguistic transitivity as “the formal and semantic features associated with [...] the linguistic coding of basic events [...] in which a volitionally acting, typically human agent targets its action at a thoroughly affected patient”. According to Kittilä (2002: 78–9), the marking of transitive clauses in English is primarily motivated by “the number of core participants”, which he defines as “participants that inhere in the semantics of events”. In intransitive clauses such as (1), there is only one core participant, the subject (S) (*he*), whereas in transitive clauses (examples [2] and [3]) the number of core

<sup>1</sup> eWAVE is an online database which presents estimates of the frequencies of different morpho-syntactic features in different varieties of English. The data comes from a questionnaire filled by 80 informants who are all experts in their field.

<sup>2</sup> This refers to the basilectal form of FjE (Kortmann and Lunkenheimer (<http://ewave-atlas.org/parameters/42#2/7.0/7.6>)).

participants is two, as both the S (*the dog, she*) and the direct object (Od) (*the ball, a book*) are present.

- (1) He read. (SV)
- (2) The dog ate *the ball*. (SVOd)
- (3) She was reading *a book* in the park. (SVOdA)

Although both verbs and clauses can be categorised as transitive, Huddleston and Pullum (2002: 216) note that it is ultimately the verb which determines the transitivity of both; for example, the verb *read* permits both the omission (1) and the presence (3) of the object and thus also determines the transitivity of the clause.

Huddleston and Pullum (2002: 218) divide the category of transitive verbs further to monotransitives, which can be either ordinary (2) or complex (examples [4] and [5]), and ditransitives, which can have only ordinary (6) structures:

- (4) She considered *him* quite handsome. (SVOdCo)
- (5) The neighbours kept *the dog* inside the house. (SVOdAo)
- (6) She sent him *a letter*. (SVOiOd)

As (2) shows, ordinary monotransitive verbs are followed by only one core argument, the Od, which according to Carter and McCarthy (2006: 498) is invariably a noun phrase or its equivalent such as a nominal clause. The second subclass of this category consists of complex transitive verbs, which, as Quirk et al. (1985: 55–6) note, are followed by an Od and an object complement (Co) as in (4), or an object-related adverbial (Ao) as in (5). Ditransitives, in contrast, have an additional core participant, the indirect object (Oi), which according to Carter and McCarthy (2006: 499) cannot occur without the Od. Kittilä (2006: 1) argues that “canonical three-participant events involve an animate agent, an inanimate theme (i.e. the thing transferred) and an animate recipient”, which in (6) are *she* (S), *a letter* (Od) and *him* (Oi). Furthermore, Dixon (2000: 41) suggests that in the English language, the syntactic functions of constituents are indicated by their order in the clause and therefore, the indirect object usually precedes the direct object, as Quirk et al. (1985: 54) note. In addition, indirect objects can frequently be placed after the Od, which turns them into prepositional objects (Op) (Quirk et al. 1985: 1208), as (7) shows.

- (7) She sent *a letter* to him. (SVOdOp)

Many researchers (see, for example, Kittilä 2010; Hopper and Thompson 1980) focusing on semantic transitivity have argued against the intransitive-transitive dichotomy of structural definitions presented above; in contrast, they prefer to view transitivity as a continuum, which takes into consideration the various nuances such as kinesis, agency, affectedness, volitionality, and the number of participants that contribute to the transitivity of a clause.<sup>3</sup> For example, although both (2) and (3) are transitive, the object (*the ball*) in (2) is more affected by the action than the object (*the book*) in (3) and therefore, the former is placed higher in the transitivity continuum than the latter. However, as Kittilä notes (2002: 79), transitivity in the English language is fundamentally determined by the number of core participants, which justifies the structural approach used here.

### 3 Data and Methods

#### 3.1 Data

With the exception of AmE, data for all varieties included in the study come from the *International Corpus of English* (ICE), a family of corpora where each corpus consists of 600,000 words of spoken and 400,000 words of written English. Since the focus of this study is on a feature that is most prominent in colloquial language, only the most informal category, ‘Private conversations’, will be examined here. ‘Private conversations’ consists of approximately 200,000 words in all Asian corpora, ICE-JA and ICE-GB, whereas only smaller amounts of data are available in ICE-EA (60,000) and ICE-FJ<sup>4</sup> (55,000). As the spoken component of ICE-US has not been published yet, the data representing spoken AmE come from the *Santa Barbara Corpus of Spoken American English* (SBCSAE). The use of the SBCSAE is a justified choice, since data from this corpus will also form a part of the unpublished ICE-US.<sup>5</sup>

#### 3.2 Methods

The nine verbs selected for the study are *bring*, *buy*, *enjoy*, *find*, *give*, *love*, *make*, *offer* and *show*<sup>6</sup>, which also represent different levels of transitivity: for example,

<sup>3</sup> For a detailed list of high and low transitivity features, see Hopper and Thompson (1980: 252).

<sup>4</sup> At the time of writing, ICE-FJ has not yet been published. I would like to thank Prof. Marianne Hundt for the opportunity to assist in the compilation of the corpora and granting me access to the unpublished data.

<sup>5</sup> For more information, see <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>.

<sup>6</sup> This selection of verbs that favour transitive uses is also supported by Mukherjee and Gries (2009: 40) to some extent.

the objects of verbs such as *make* and *give* tend to be more thoroughly affected by the action than the objects of *love* and *show*, and therefore, the former verbs are usually placed higher in the transitivity continuum than the latter. The searches were conducted using WordSmith Tools 6.0, which enables the search for not only individual words from the data, but also provides a link to the original file where the search hit is located.<sup>7</sup>

In the first phase of the analysis, all instances of verbs were collected from the data and analysed manually: all sentences where the verbs were used transitively but were not accompanied by an overt direct object were separated from the canonical transitive sentences, whereas instances of repetition, false starts and other unclear cases were discarded. The sentences where no Od was present were then subjected to a detailed analysis where several clues in the transcription were examined in order to ensure that the direct objects had indeed been omitted. For example, if the omitted pronoun was followed by a pause, as in (8), it was analysed as an example where the speaker did not intend to add an Od after the verb.

(8) A: Well do you *enjoy* <,>

B: Yeah <,> I like that work <w> I'll </w> *enjoy* more <,>

(ICE-IND:S1A-043#55-6)

Unfortunately, some files in the corpora did not contain any pause markings, as (9) shows; in such cases, that particular section of the dialogue was examined in greater detail.

(9) B: How 'bout the <.> ing </.> and the others <.> whe </.>  
when will we *buy*

B: When will we go out and *buy*

A: Friday

(ICE-PHI:S1A-014#36-8)

In (9), speaker B repeats his/her question, but both times does not add the Od after *buy*, and hence, the omission of the object in speaker B's repeated question

---

<sup>7</sup> The search parameters were set to exclude all instances of the corpora marked as 'extra-corpus material' (<X>\*</X>), 'editorial comments' (<&>\*</&>), 'untranscribed text' (<O>\*</O>) and transcriber's corrections (<+>\*</+>). The exclusion of the first three tags were based on the recommendations given in the ICE homepage (<http://ice-corpora.net/ice/importantnote.htm>). The decision to exclude also 'Transcriber's corrections' was based on the personal observation that many corpora (especially ICE-IND), contained many words and sentences which were corrected by the transcribers.

was considered to be deliberate. Also, no sections of speaker A and B's turns have been marked as overlapping and therefore, object omissions in sentences such as these were not considered to be caused by interruptions. In addition, instances where the speaker continued with their turn although the Od of the transitive verb had been left out were considered to be indications of deliberate object omission; an example of this is provided in (10).

- (10) A: I don't want a chocolate eclair but you cannot *give* to Ai Hui because  
 she <O> coughs </O> (ICE-SIN:S1A-021#100:1:)

After this detailed analysis, all instances of transitive verbs that were not followed by direct objects were included in the data.

In order to determine the frequency of the feature in New Englishes, the number of omitted objects was compared with the overall number of canonical transitive constructions in the data using the X-coefficient, which is an adaptation of the V-coefficient (Smitherberg 2005: 44).<sup>8</sup> The X-coefficient relates the number of transitive verbs with omitted objects to the total number of transitive verbs, as shown below

$$X = \text{NZERO Od} / \text{NTRANS} \times 10,000.$$

Since all nine verbs examined for this chapter can be used as intransitives and transitives (though they favour the latter), the use of the X-coefficient enables a variationist approach to the data, where the canonical transitive constructions and the transitive constructions with omitted direct objects form the two variants of the transitive clause.

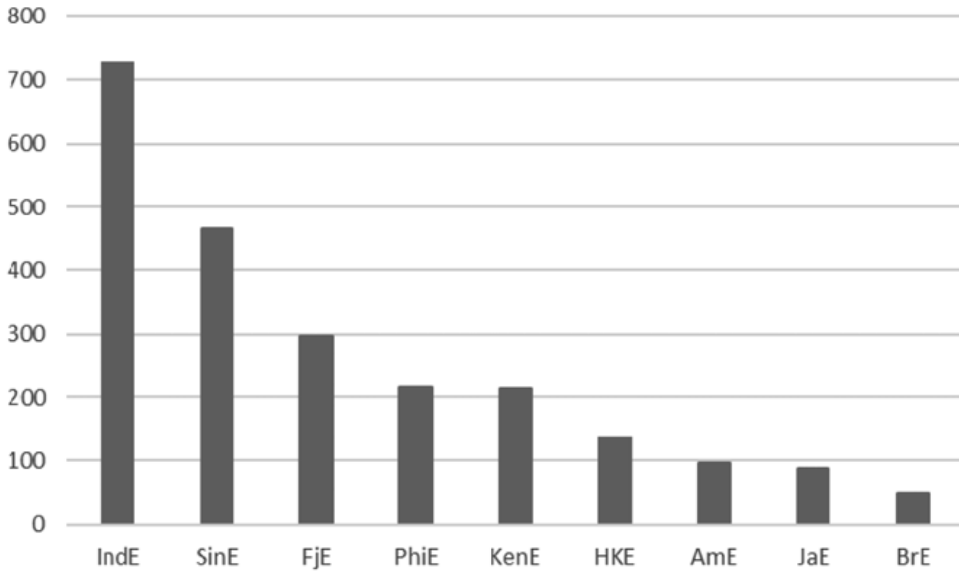
## 4 Results

As was mentioned in the introduction, the tendency to omit object pronouns has been argued to be a common feature in many New Englishes. Figure 1 and Table 1 below present the results on the nine varieties of English from the highest to the lowest.

---

<sup>8</sup> The V-coefficient (the *V* stands for 'verb') is used to relate the number of progressives to the number of verb phrases in a text (Smitherberg 2005: 44–5).





**Figure 1:** Frequency of direct object omission in nine varieties of English, X-coefficient

**Table 1:** Frequency of omitted direct objects, X-coefficient

	Canonical transitives	Transitives with zero Od	X-coefficient
IndE	981	77	728
SinE	1023	50	466
FjE	262	8	296
PhiE	1085	24	216
KenE	318	7	215
HKE	1011	14	137
AmE	818	8	97
JaE	1004	9	89
BrE	829	4	48

The results in Table 1 show that the tendency to omit direct objects is most frequent in IndE, followed by SinE, FjE, PhiE, KenE, HKE, AmE, JaE and BrE, in descending order. The overall distribution between the varieties is statistically significant ( $p < 0.001$ ), while the only statistically significant difference between individual varieties can be found between IndE and SinE ( $p < 0.05$ ) (for detailed figures, see Appendix 1). Examples from each variety are provided in (11) to (19):

- (11) A: Who doesn't loves <,> <indig> sarees </indig> <,> <{> <[>yeah <,>  
Kanchipuram <indig> sarees </indig> and <O> one word </O>  
B: <[> Yeah </[> </{>  
B: You don't *love*?  
C: Oh <w> I'm </w> not much interested in <,> <{> <[> dressing up </[>  
(ICE-IND:S1A-029#166-169)
- (12) A: Did your brother *make*  
A: I'm sure he did  
B: He no he didn't *make*  
B: He just bought it (ICE-SIN:S1A-066#93-6)
- (13) A: rum yeah rum I can drink  
B: hot stuff eh?  
A: yeah hot stuff it's good <,,>  
A: basically the imported ones <,,>  
B: what the imported ones?  
A: yeah imported <,> just like Red Label and all those <,> from abroad not  
the Fiji rum <,> <{><[>very strong</[> <,,> <O>laugh</O> I can *buy* <,,>  
(ICE-FJ-S1A-090)
- (14) B: ... I got so much interested that in her assignment in the <-\_>in the<-/>  
story  
I would miss doing anything else but not that not to get good marks but I  
*enjoyed* and I would read that book over and over again...  
(ICE-EA-convers2-K)
- (15) B: I am keeping my doors open but then I know I cannot *find*  
A: Uh uhm okay I understand I know (ICE-PHI:S1A-018#113-4)
- (16) B: So I uhm<,,> ask the parents to bring in the vouchers  
A: Yes  
B: Other parents who did not *bring* in uhm explained that <}> <->  
(ICE-JA:S1A-098#23-25)

- (17) A: Yeah they they had take some photographs uh and <,> they *showed* to me and I I think uh it's very beautiful yeah (ICE-HK:S1A-042#508:1:A)
- (18) B: Oh we were going through the uh <,> I was *showing* Mike and uh Dinah on Thursday  
 B: and we were going through the C V <,> (ICE-GB:S1A-016 #176-178)
- (19) NANCY: And so he needed change. Like two fives for a ten or something. So she *gave* him.  
 (H) And then, he kept it?  
 He didn't give her like the ten you know? (SBCSAE0050)

The structural and semantic transitivity of a clause can sometimes disagree and, as Kittilä (2002: 37) notes, the number of arguments in a clause in such cases is usually lower than the number of participants involved. This holds true also in (11) to (19), as they all lack a core argument which is implied in the surrounding context. For example, although the surface structure of *So she gave him* in (19) could be interpreted as an ordinary monotransitive, the surrounding context reveals that it is in fact a ditransitive where the direct object is omitted: the woman in question gave the man *change* (Od), although only *she* (S) and *him* (Oi) are mentioned in the clause.

Although direct object omission occurs in all the varieties, there are also great differences between them; the frequencies of BrE and AmE are among the lowest, which supports the argument that the feature was not transmitted to the other English varieties from the two superstrates. Thus, the origins of the tendency to omit direct objects must be located elsewhere. IndE has the highest frequencies of object omission and it is followed by SinE and FjE; interestingly, both Singapore and Fiji have ethnic Indian minorities which still speak Indian languages as their L1, and it is therefore possible that the heightened frequencies in SinE and FjE are caused by the presence of the ethnic Indian minorities in the data. However, considering that the ethnic Indians comprise only approximately nine per cent of the Singaporean population (Population trends 2013, Department of Statistics), whereas the frequency of direct object omission in the SinE data is over 60 per cent of the corresponding frequency of IndE, it is highly unlikely that the proportion of ethnic Indians in the SinE data is large enough to solely explain the high frequencies of SinE<sup>9</sup>.

---

<sup>9</sup> Unfortunately, the speaker information of ICE-SIN is not available and therefore, it is not possible to present estimations on the distribution or the strength of the feature among the different ethnic groups of Singapore.

In the case of FjE, the possibility of an ethnic Indian minority could indeed explain the results, as Indo-Fijians form almost 40 per cent of the Fijian population (2007 Population Census, Fiji Bureau of Statistics), and they form approximately 50 per cent of the informants in the ICE-FJ data used for this study. In fact, in six out of the eight instances where the direct object had been omitted, the speaker was of Indo-Fijian ethnicity. However, since the remaining two hits came from a Rotuman and a Fijian, it is clear that Indo-Fijians do not form the sole ethnic group using this feature.

Heightened frequencies of object omission can also be found in PhiE and KenE, where the tendency to use the feature is equally strong. Interestingly, HKE uses direct object omission the least when compared with the other Asian Englishes included in the study, and it is thus closest to its source variety, BrE. This supports the observations made by Mukherjee and Gries (2009), whose collostructional analysis on the patterns of transitive and intransitive verbs indicates that HKE aligns itself closer to BrE than the two other Asian varieties (SinE and IndE) included in their study. The closeness of HKE to BrE has also been observed by Rautioaho (2014). The results on JaE are also interesting, as the variety is aptly positioned between its past and present superstrates, BrE and AmE.

## 5 Omission of direct objects in substrate languages

As mentioned above, the frequency at which speakers of New Englishes omit direct objects could be linked to local language influence, and therefore, the existence of a similar linguistic feature in the substrates of these varieties must be examined in greater detail. Since the total number of languages spoken in India, Singapore, Fiji, the Philippines, Kenya, Hong Kong and Jamaica exceeds the scope of this study, only the major languages from each country were selected for closer examination. This restriction is a justified choice, as the languages with the largest numbers of speakers are also more prominently represented in the “feature pool” of each variety. The term is an integral part of Mufwene’s (2001: 3–6) “Founder principle”, which suggests that, in contact situations, the different languages (and varieties) spoken by the population form a pool from which different features emerge through competition and selection and become manifested in the new language (or variety). Thus, if the tendency to omit direct objects is common in the larger local languages, substrate influence could provide a plausible explanation for the increased use of the feature

in New Englishes.<sup>10</sup> However, before proceeding with the analysis, a more detailed account of the languages and the motivation for their inclusion should be presented.

Hindi is the only indigenous language of India with official status<sup>11</sup>, and since it is also the largest substrate language of IndE with 41 per cent of the population speaking it as their L1 (Graddol 2010: 51), the selection of Indian languages examined here is restricted to only Hindi. The official local languages of Singapore are Mandarin Chinese, Malay and Tamil, but only the first of these is of greater relevance to this study, since, as Lim (2007: 453, 456) argues, the forms of Malay and Tamil used by the Malay and Indian communities of Singapore are not significant in the contact dynamics of the area. In contrast, Bazaar Malay and Hokkien should be examined closer, as they both functioned as lingua francas in the Singaporean area (Gupta 1992: 327).<sup>12</sup> In addition, Cantonese and Teochew have been noted to be important substrates for SinE (Ansaldo 2004: 132), but only the former will be examined here, since Teochew differs from Hokkien only in its phonology.<sup>13</sup> Cantonese is also the only major substrate for HKE, as it is the L1 of approximately 90 per cent of the population (Thematic Household Survey Report No. 51, Census and Statistics Department 2013). The majority of Fijians speak either Fijian or Fiji Hindi as their L1 (2007 Population Census, Fiji Bureau of Statistics), and thus both languages are included in the analysis. The language situation in Kenya is more complex, as the official languages, English and Swahili, are mostly used as lingua francas in the country where multiple local languages are spoken as L1. Therefore, in addition to Swahili, two locally spoken languages, Kikuyu and Luo, are included here, as they are the biggest languages of the two major language branches (Bantu and Nilo-Saharan) spoken in the country (Population and Housing Census 2009, Kenya National Bureau of Statistics). The two official languages of the Philippines are English and the Tagalog-based Filipino, and the latter also has the largest number of speakers in the country, approximately 28 per cent of the population

---

**10** This approach is not new, as the founder principle has also been applied in studies on, for example, language evolution and ecology in Asia (Ansaldo 2009) and new dialect formation processes in London (Cheshire et al. 2011).

**11** There are also 22 other associate languages which are recognised in the Indian constitution, but none of the speaker populations of these languages represent even 10 per cent of India's total population (Graddol 2010: 51). Therefore, it can be argued that Hindi has the strongest representation in the feature pool of IndE.

**12** According to Lim (2007: 453–4), Bazaar Malay was used by the whole of the Singaporean population whereas Hokkien was used among the Chinese speaking population – both were replaced by English during the 1970s and 80s (Lim 2007: 453–4).

**13** Bao Zhiming, personal communication, 14.3.2013.

(2011 Philippines in Figures, National Statistics Office). Despite the relatively low proportion of native speakers of Tagalog in the Philippines, restricting the selection to this language is a justified choice since, as Schachter (1976: 493) notes, “the languages of the Philippines are sufficiently similar that examples from any one language can safely be taken as paradigmatic”. Jamaican Creole (JaC) is the most important substrate of JaE (Patrick 2004: 408), and thus the local creole should also be examined here.

## 5.1 Rich agreement

For some languages, the tendency to omit objects can be explained with the language’s rich agreement system, which makes stating the object explicitly redundant. This rationale can be applied to some of the languages listed above, as the morphologies of Hindi (20), Fijian (21), Swahili (22), Kikuyu (23) and Luo (24) are complex enough to explain to some degree the tendency to omit object pronouns in IndE, FjE and KenE.<sup>14</sup>

- (20) *mēne kahi.*  
 1SG-erg said-fs  
 ‘I said (it) to him/her.’ (Koul 2008: 214)

- (21) [*e ronqo-ta*] *tiko na marama.*  
 3SG hold-TR CNT DET woman.  
 ‘The woman is holding him.’ (Aranovich 2013: 467)

- (22) *ni-na-ku-ona*  
 1SG-PRES-2SG.OBJ-see  
 ‘I see you.’ (Myachina 1981: 64)

- (23) *nĩ-a-ra-mũ-rĩ-ĩr-a waru*  
 fp-3SG.SUB-PRES-3SG.OBJ-eat-A-fv 10potato  
 ‘She/he is eating food for her/him.’ (Mugane 1997: 159)

<sup>14</sup> The question of Fiji Hindi morphology remains open, as no examples were found where the direct object was omitted due to rich morphology. Siegel (1988: 121) has argued that Fiji Hindi is a morphologically simplified form of Hindi, which is “derived from several different Indian Hindi dialects”. Therefore, it is possible that the process of simplification has resulted in the weakening of the tendency in the language. Nevertheless, since Fiji Hindi has been mentioned to allow pro drop (see, for example, Gounder 2011), the scarceness of data on Fiji Hindi could also explain why no examples of this tendency were found for this paper.

- (24) *o-tedo-n(i)-a-gi*  
 3SG-cook-for/to-1SG.OBJ-them  
 ‘He cooks them for me’ (Stafford 1967: 17)

In (20), the object of the Hindi sentence can be omitted because “the verb [...] *kahna*: is inflected for an implied generic feminine object” (Koul 2008: 214). (21), in turn, is an example from Fijian, where the pronoun *him* is omitted; according to Kittilä (2002: 91), transitivity in (Boumaa) Fijian is signalled by verbal affixes, which is why, as Aranovich (2013: 467) points out, the omission of objects is allowed in some contexts. The omission of objects is also possible in Swahili, where “the independent arguments are very often eliminated and participants are referred to by agreement affixes only”, as Kittilä (2002: 98) notes. Interestingly, this tendency can be observed in both of the Bantu languages examined here, as the subject and object are signalled in the verbs of not only (22) from Swahili but also (23) from Kikuyu. The same can also be seen in (24) from Luo where, according to Stafford (1967: 17), object pronouns can be reduced to verbal affixes. As the examples above indicate, a rich morphology in the substrate languages can indeed explain why object omission is common in some New Englishes, although there are also some restrictions as to the applicability of this explanation in some languages.<sup>15</sup>

## 5.2 Topic-prominence

Although rich morphology in the substrates of IndE, FjE and KenE can indeed explain why objects are omitted in these three varieties, this cannot explain why the same feature is also found in other Asian Englishes such as SinE, where the substratum consists mostly of isolating languages with a poor morphology. As an answer to this question, Sato (2011: 362) suggests the following:

there is more than one grammatical source for the liberal omission of grammatical elements in the syntax. Classical pro-drop languages like Spanish [...] allow agreement-based drop because their inflectional morphology is rich enough to recover the missing element from agreement. Agreementless, topic-prominent languages allow [radical pro drop] because of the distinct topic structure underlain by topic prominence.

---

<sup>15</sup> Mugane (1997: 159) notes that in Kikuyu, the number of object prefixes attached to the verb is restricted to one and thus, the language allows the omission of only one object in a sentence. In addition, Aranovich (2013: 496) argues that objects belonging to the first of three categories in Fijian are “licensed by an *-a* suffix on the verb”, which he considers to be “an agreement marker [that identifies] an empty *pro* in the position of complement of V”. Therefore, object omission is possible also in Fijian, although the use of the feature is more constrained.

This view is supported by Li and Thompson (1976: 409), who note that “many structural phenomena of a language can be explained on the basis of whether the basic structure of its sentence is analysed as subject-predicate or topic-comment”; since many Asian languages are topic-prominent (see, for example, Sato 2011; Junghare 1988), this could explain some of the results presented in this paper.

According to Li and Thompson (1976: 484), the notion of topic is universal, but it is manifested differently in subject-prominent (Sp) and topic-prominent (Tp) languages; the differences between these two types are demonstrated in (25) and (26).

- (25) John hit Mary  
Subject Predicate (Li and Thompson 1976: 459)

- (26) *Nèi-xie shumu shù-shēn dà*  
 those trees tree-trunk big  
 Topic Comment  
 “Those trees, the trunks are big.”  
 (Adapted from Li and Thompson 1976: 462)

- |      |                   |                                       |
|------|-------------------|---------------------------------------|
| (27) | As for education, | John prefers Bertrand Russell's ideas |
|      | Topic             | Comment (Li and Thompson 1976: 459)   |

(25) shows the structure of a basic, unmarked sentence in English, which is an Sp language where the main components are the subject and the predicate. In contrast, (26) from Mandarin Chinese shows the unmarked structure of a sentence in a Tp language where ‘those trees’ functions as the topic and thus signals what the sentence is about, whereas the comment (‘the trunks are big’) provides further information about that topic. Li and Thompson (1976: 484) note that “subjects are essentially grammaticalized topics”, and therefore they are also assigned some of the properties of the topic. Here it should be noted that defining a language as Sp does not necessarily mean that it cannot have topic-comment structures or vice versa; as (27) shows, the topic-comment structure is also possible in English, where it is used as a special construction. Furthermore, the subject (*John*) and topic (*as for education*) are located in different arguments, which shows that the topic and subject do not always coincide although they are closely connected.

In their analysis of aspects that distinguish Tp languages from Sp languages, Li and Thompson (1976: 466–71) provide a list of eight features that are summarised in Table 2.



**Table 2:** Characteristics of Tp languages (based on Li and Thompson 1976)

Feature	Description
a) Surface coding	The topic is coded (e.g. always in initial position) in Tp languages whereas coding of the subject might not be necessary
b) The passive construction	Passivization does not exist, it is rare, or it has special meaning
c) “Dummy” subjects	Not used
d) “Double subject”	Occur in all Tp languages
e) Controlling co-reference	The topic (and not the subject) controls the omission of the co-referent constituent in Tp languages
f) V-final languages	Tp languages are frequently verb-final
g) Constraints on topic constituent	Anything can function as the topic
h) Basicness of topic-comment sentences	Topic-comment sentence types are considered to be “basic”

Importantly, Li and Thompson (1976: 483) note that “[a]s with all typological distinctions, [...] it is clear that we are speaking of a continuum”. Therefore, the stronger these features are in a language, the more Tp it is. One specific feature, ‘Controlling co-reference’ is of special interest for the topic of the present paper, as it explains why certain languages allow the omission of objects despite their poor morphology. According to this principle, objects and other constituents can be freely omitted when their referent is the topic and thus recoverable from the context; (28) from Mandarin Chinese is an example of this.

- (28) *Nèike shù yèzi dà, suǒyǐ wǒ bu xǐhuan \_\_\_\_.*  
 that tree leaves big so 1SG not like  
 ‘That tree (topic), the leaves are big, so I don’t like (it)’

(Li and Thompson 1967: 469)

In (28), the object of dislike is *that tree*, which is also the topic and therefore, there is no need for the speaker to provide the coreferential object pronoun *it* in the comment.

Many Chinese dialects, together with Bazaar Malay, have been argued to be Tp (Li and Thompson 1976; Yip and Matthews 2007: 135; Khin Khin Aye 2005: 153). Examples of object omission controlled by the topic are provided in (29) to (30) from Cantonese and (31) from Bazaar Malay.

- (29) [TOPIC *ni1 gin6 saam1*]<sub>i</sub> *ngo5 hou2 zung1ji3 e<sub>i</sub>*.  
           this CL dress 1SG very like this  
           ‘This dress, I like [it] a lot.’ (Yip and Matthews 2007: 135)
- (30) [TOPIC  $\emptyset$ ]<sub>i</sub> *ngo5 hou2 zung1ji3 e<sub>i</sub>*.  
           1SG very like  
           ‘I like [it] a lot.’ (Yip and Matthews 2007: 135)
- (31) *Nanti saya kawan tahu*.  
       wait 1SG friend know  
       ‘Wait (for a while), my friends know (Malay, that is, how to speak Malay).’  
           (Khin Khin Aye 2005: 160)

Yip and Matthews (2007: 134–5) argue that “just as a missing object can refer back to an overt sentence topic” as in (29), “so it can refer to a topic which is implied but not stated” as in (30). As could be expected, Yip and Matthews (2007: 135) link this tendency to the “‘topic-prominent’ characteristics of Chinese as a whole”. The same tendency can be seen in (31) in Bazaar Malay where, as Khin Khin Aye (2005: 149–50) notes, “the object NP can [...] be omitted as long as it can be inferred either from the context or from knowledge shared between the interlocutors”.

The classification of Hindi is slightly more problematic, since some scholars, such as Kidwai (2004: 255), claim that Hindi is an Sp language<sup>16</sup>, whereas others (see, for example Junghare 1988: 316; Sato 2011: 362) argue that the language is in fact more Tp.<sup>17</sup> According to Junghare (1988: 322), “the zero pronouns occur more in conversational varieties of [Hindi]”; an example of object omission licensed by topic-prominence is provided in (32).

- (32) *jonko khat milā leki usne  $\emptyset$   $\emptyset$  dikhāyā nahī*  
       John letter got but 3SG.M showed not  
       ‘John received the letter but he didn’t show (it) (to me).’  
           (Junghare 1988: 325)

<sup>16</sup> Also Li and Thompson (1976) argue that Indo-European languages are Sp.

<sup>17</sup> Interestingly, Li and Thompson (1976) also distinguish a third category, languages that are both Tp and Sp, where “there are two equally important distinct sentence constructions, the subject-predicate construction and the topic-comment construction”. However, defining the exact location of Hindi on the Tp-Sp continuum according to the eight categories mentioned above would exceed the scope of this paper.

- (33) *ke surū kar-is rahā*  
 who start do-PF AUX  
 ‘Who started (it)?’ (Siegel 1988: 128)

The same phenomena can also be seen in (33) from Fiji Hindi, where the omission of the direct object is licensed by the anaphoric reference of the omitted pronoun to the topic, which is clear from the context.

Li and Thompson (1976: 459) suggest that there are some languages where “the subject and the topic have merged and are no longer distinguishable in all sentence types” and thus, they are neither Tp nor Sp. Interestingly, many Philippine languages, including Tagalog, belong to this category (see, for example, Li and Thompson 1976; Schachter 1976; Shibatani 1991). Despite this, Himmelmann (1999: 232) argues that “[z]ero-options exist for both actors and undergoers in all kinds of semantically transitive constructions in Tagalog” as (34) shows.

- (34) *Huhugasan ko and=mga=pinggan, at pupunasan mo*  
 FUT-wash-DV1. SG.GEN NOM=PL=dish and FUT-dry-DV 2.SG.GEN  
 ‘I will wash the dishes, and you will dry (them).’ (Kroeger 1993: 34)

According to Kroeger (1993: 33), “Tagalog allows [zero anaphora] to apply quite freely”, though it “requires that the antecedent actually precede the null pronoun, whether in the same sentence or in discourse context”, which is illustrated well in (34).

Interestingly, Li and Thompson (1976: 460) argue that languages of the Niger-Congo language family, which includes such Bantu languages as Swahili and Kikuyu, are Sp. Furthermore, the eight characteristics of Li and Thompson (1976) cannot be found in Swahili to the same extent as in prominently Tp languages<sup>18</sup>, and it is therefore unlikely that the Bantu substrates are as strongly Tp as, for example, Chinese dialects. JaC has also been noted to be Sp (Sato 2011: 362) which, according to Neeleman and Szendrői (2007: 690; also Kortmann and Lunkenheimer 2013), does not allow (radical) pro drop. However, Loftman Bailey (1966: 80) contradicts this view by stating that “if there is anything in the context which makes it clear what the direct object is, then that object is deletable” as in (35).

<sup>18</sup> See Augustin’s (2007) discussion on Swahili topic structures.

- (35) *huu sel yu di tebl*  
 ‘Who sold you the table?’

*Jak sel me*  
 ‘Jack sold (it to) me.’

(Loftman Bailey 1966: 80)

Again, the key issue seems to be the relatedness of the omitted object to the topic of the conversation. This issue will be discussed further below.

Kittilä (2002: 77) suggests that in a large number of languages “everything that can be inferred from the context is either omitted or is not marked explicitly”, and the discussion above supports this view. In fact, many of the sources cited above note that object omission occurs in the context of conversations, and since “the topic [...] is primarily a discourse notion”, as Junghare (1988: 322) points out, it is not surprising that examples of object omission licensed by the topic can be found in the spoken genres of virtually all substrates. Furthermore, since the data examined for this study represents informal spoken English, finding some instances of omitted objects in all colloquial varieties of English is something to be expected. What is important to note here is that, firstly, there are clear differences between the English varieties as to how frequently objects are omitted, and secondly, these differences seem to be linked to the topic prominence of the substrates. In Sp languages, the subject has many of the properties of the topic, and thus, sentences where the object can be omitted because it refers to the topic are rare. In contrast, there are fewer restrictions as to what can function as the topic in Tp languages and, as a consequence, omitted objects referring to the topic also occur more frequently.

Topic prominence does not, however, fully explain the results, since objects are omitted in IndE more frequently than in SinE, although Tp is more prominent in the substrates of the latter. In addition to Hindi, rich agreement and object omission are known to be common in many other South Asian languages (see, for example, Subbārão 2012: 28; Butt 2001: 2). Therefore, a possible explanation for the difference between the results on IndE and SinE is that the high frequency of object omission in IndE is caused by the interplay of Tp and rich morphology in the substrates. Another interesting observation is the difference between the frequencies of SinE and HKE, since Cantonese is a major substrate of both varieties. A possible explanation for this is that ICE-HK contains more acrolectal English when compared with the other ICE-corpora<sup>19</sup>, though it is difficult to estimate whether this detail alone is enough to explain the differences between the varieties. In contrast, the results on FjE can indeed be explained by substrate

<sup>19</sup> Lisa Lim, personal communication, 19.4.2013.

influence, as most of the ICE-FJ informants whose speech contained omitted objects were Indo-Fijians, and thus, their L1 is Fijian Hindi.<sup>20</sup> Interestingly, not all of the substrates of PhiE, KenE and JaE examined here were as prominently Tp as those of IndE, SinE, and FjE. Of these three varieties, KenE is the only one where the substrates have morphologies rich enough to explain some level of object omission in the local variety of English, and it is thus interesting that the tendency to omit objects in the variety is not higher than in PhiE.

## 6 Discussion

Whenever a feature is noted to be common in New Englishes, the explanations offered may vary from linguistic universals to substrate influence and second language acquisition, and often more than one of these alternatives can be regarded as a plausible explanation. Since object omission occurs more frequently in the New Englishes included in the present study than in BrE and AmE, the possibility of other explaining factors should also be addressed here. Sharma (2012) presents a useful theoretical framework, which provides the order in which these competing explanations should be considered. According to Sharma (2012), the explanations can be divided into four main types: ‘Properties shared with the superstrate’ (A), ‘Properties shared with the substrate’ (B), ‘Acquisitional universals’ (C) and ‘General universals’ (D). Furthermore, Sharma (2012: 221) argues that “[i]t is logical to *first* examine the languages in contact (Types A and B) for proximate causes of shared features before appealing to more general motivations [of Types C and D]”. This order will also be followed in this section. Importantly, Sharma (2012: 219) notes that these alternatives should be explored only after “genuine similarity [in the use of the feature] has been established” in the varieties. As was noted in the introduction of this chapter, broad remarks have been presented on the tendency in New Englishes to drop (object) pronouns. However, as the results of the present study indicate, a joint tendency in New Englishes can at least be observed in the case of direct object omission. Therefore, Sharma’s (2012: 219) requirement of genuine similarity is met.

According to Sharma (2012: 222), Type A explanations include ‘Founder effects’, ‘Exogenous prestige-driven mimicry’ and ‘Diffusion from one variety to another’. Since the tendency to omit object pronouns is weaker in the superstrate varieties, which are also the old (BrE) and new (AmE) prestige varieties,

---

<sup>20</sup> Despite the likelihood of this explanation, it should also be acknowledged that the results may be partially skewed by the smaller amount of data available on FjE.

‘Founder effects’ and ‘Exogenous prestige-driven mimicry’ do not provide plausible explanations for the prominent use of object omission in New Englishes. In contrast, the possibility of ‘Diffusion from one new variety to another’ should be explored further, as the frequencies of IndE, followed by SinE stand out from the results on the other New Englishes. Interestingly, some of the results of recent studies (Hundt et al. 2012: 163; Bernaisch and Gries 2015) suggest that IndE could indeed function as a linguistic epicentre in South Asia (see also Mair 2013: 263–4). Furthermore, although there is no diachronic data available on the majority of Asian Englishes, Parviainen and Fuchs (submitted) have used the apparent-time method to show that some innovative features are more established in IndE than in Southeast Asian Englishes, which supports the idea of IndE having some level of influence on Southeast Asian varieties. The differences between the results on SinE and HKE could further support the idea of diffusion from IndE to SinE: even if the low frequencies of HKE were caused by the differences in the ICE-SIN and ICE-HK data, it seems questionable whether this could solely explain the difference between SinE and HKE results. Therefore, the influence of IndE could have strengthened the tendency to omit object pronouns in SinE, although the possibility of independent development in the latter variety cannot be completely ruled out either.

Sharma (2012: 223) lists ‘Accidental resemblance’ and ‘Areal convergence’ under ‘Properties shared with the substrate’ (Type B) explanations. As the discussion in the previous section shows, substrate influence provides a plausible explanation for many of the results presented here. South and Southeast Asia have functioned as ‘melting pots’ of various cultures and languages for centuries, which has resulted in linguistic convergence of many languages spoken in the region.<sup>21</sup> Therefore, ‘Areal convergence’ should be considered the more likely explanation of the two. Furthermore, the levels of Tp in the substrates of the remaining Englishes fall somewhere between the strong Tp of SinE substrates and the strong Sp of English, and the frequencies of object omission in these varieties are also located somewhere between the two. Therefore, substrate influence is the most likely explanation for the differences in the tendency of omitting direct objects in the majority of New Englishes.

Gundel (noted in Junghare 1988: 322) suggests that it is ultimately the topic that controls the zero NP-anaphora in all languages, which would explain why

---

<sup>21</sup> Junghare (1988: 326) suggests that the reason why the Indo-Aryan language branch is more Tp than Sp, despite the fact that Sanskrit (together with other Indo-European languages) had become Sp already by the classical period, is because the other languages spoken in South Asia area are more Tp. Thus, their influence could have caused the Indo-Aryan languages such as Hindi to become more Tp in the course of time.

some examples of object omission were also found in spoken BrE and AmE. Despite this, generic explanations based on language universals should be considered with greater caution. Sharma (2012: 223) argues that “[t]rue emergence of universals would only be certain if the substrates do *not* have a particular feature, and yet the feature arises in offspring varieties of English”. Because the concept of topic is predominantly a discourse notion (see Junghare 1988: 322, above), the fact that the data examined comes from private conversations can explain why some instances of object omission are found in all varieties included in the present study. However, because there are also great differences between the English varieties, and because these differences can be explained with substrate influence, it is ultimately the substrate that explains the distribution of the feature among English varieties. Furthermore, as Sharma (2012: 224) notes, “[i]n the absence of either Type A or B explanations it is reasonable to *next* consider the possibility of SLA effects (Type C) before assuming the wider operation of genuine universals (Type D)”. Therefore, since Type B (and possibly Type A in the case of SinE) can explain the results presented here, there is no need to look further for explaining factors for object omission in New Englishes. As was mentioned in the introduction, according to Kortmann and Lunkenheimer’s (2013) eWAVE database, the omission of object pronouns is “pervasive or obligatory” in IndE, SinE<sup>22</sup>, FjE<sup>23</sup>, KenE and HKE, whereas “the feature is neither pervasive nor extremely rare” in PhiE. Furthermore, “attested absence of the feature” can be found in JaE (and JaC), AmE and BrE (Kortmann and Lunkenheimer 2013). The results of the present study support the majority of these arguments, although some conflicting observations can also be made. IndE, SinE, FjE and KenE indeed all favour the use of the feature more than the other varieties included in the study, thus supporting some of the arguments presented by Kortmann and Lunkenheimer (2013), but the differences between these four varieties are greater than one would expect based on the eWAVE classification.<sup>24</sup> The other interesting observation is that HKE, which according to Kortmann and Lunkenheimer (2013) favours object pronoun omission strongly, is actually located

---

<sup>22</sup> Kortmann and Lunkenheimer (2013) actually note that the feature is common in “Colloquial Singapore English (Singlish)”. Since the data used for this study comes from the category ‘Private conversations’ and represents informal language use in Singapore, it could be argued that the data examined for this paper is close to Singlish.

<sup>23</sup> According to eWAVE, basilectal FjE does favour the omission of object pronouns, whereas acrolectal FjE does not. Although ICE-FJ contains acrolectal FjE, the data used for this study comes from the section ‘Private conversations’ and thus represents a more colloquial form of the local variety. This could explain the results obtained from ICE-FJ to some extent.

<sup>24</sup> ICE-IND has been argued to contain more basi/mesolectal English, but considering that only the most informal register of all the corpora were examined in this paper, it is unlikely that this detail would solely explain the high frequencies of IndE.

between KenE and AmE. Another observation worth noting is that the PhiE variety favours the feature as much as KenE, although Kortmann and Lunkenheimer (2013) suggest that the feature is more pervasive in the latter variety. In contrast, the low frequencies of JaE were consistent with the eWAVE data.

## 7 Conclusion

This paper has examined the frequency of direct object omission in nine varieties of English. The focus has been on spoken informal language and the results indicate that there are clear differences between even those varieties which the previous literature suggest to be convergent. The results support the conclusion that the omission of direct objects is motivated by substrate influence and, in the case of SinE, possible IndE influence. There are many features in New Englishes that descriptive studies have noted to be “common”, but as the results of the present study indicate, more comparative data-driven research is still needed, as they can provide interesting insights for the study of Englishes.

## References

- Ansaldo, Umberto. 2004. The evolution of Singapore English: Finding the matrix. In Lisa Lim (ed.), *Singapore English: A grammatical description*, 127–149. Amsterdam/Philadelphia: John Benjamins.
- Ansaldo, Umberto. 2009. *Contact Languages: ecology and evolution in Asia*. Cambridge: Cambridge University Press.
- Aranovich, Raúl. 2013. Transitivity and Polysynthesis in Fijian. *Language* 89. 465–500.
- Augustin, Maryanne. 2007. Topic and Focus in Swahili. *Graduate Institute of Applied Linguistics (GIAL) Electronic Notes Series* 1. 1–12.
- Baskaran, Loga. 2004. Malaysian English: morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A Handbook of Varieties of English, Vol. 2: Morphology and Syntax*, 1073–1085. Berlin: Mouton de Gruyter.
- Bhatt, Rakesh M. 2004. Indian English: syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A Handbook of Varieties of English, Vol 2: Morphology and Syntax*, 1017–1030. Berlin: Mouton de Gruyter.
- Bernaisch, Tobias and Stefan Th. Gries. 2015, May. Identifying linguistic epicentres empirically: The case of South Asian Englishes. Unpublished conference paper presented at ICAME36, Trier, Germany.
- Carter, Ronald and Michael McCarthy. 2006. *Cambridge Grammar of English*. Cambridge: Cambridge University Press.
- Cheshire, Jenny, Paul Kerswill, Sue Fox and Eivind Torgersen. 2011. Contact, the feature pool and the speech community: the emergence of Multicultural London English. *Journal of Sociolinguistics* 15. 151–196.



- Dixon, R. M.W. 2000. A typology of causatives. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Changing valency: case studies in transitivity*, 30–83. Cambridge: Cambridge University Press.
- Gounder, Farzana. 2011. *Indentured identities: resistance and accommodation in plantation era Fiji*. Philadelphia: John Benjamins Publishing Company.
- Graddol, David. 2010. *English Next India*. London: British Council.
- Gupta, Anthea Fraser. 1992. Contact features of Singapore Colloquial English. In Kingsley Bolton and Helen Kwok (eds.), *Sociolinguistics Today: international perspectives*, 323–345. London: Routledge.
- Himmelmann, Nicolaus P. 1999. The lack of zero anaphora and incipient personal marking in Tagalog. *Oceanic Linguistics* 38. 231–269.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hundt, Marianne, Hoffmann, Sebastian and Mukherjee, Joybrato. 2012. The Hypothetical Subjunctive in South Asian Englishes – Local Developments in the Use of a Global Construction. *English World-Wide* 33. 147–164.
- Hopper, P. J. and S.A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56. 251–99.
- Junghare, Indira Y. 1988. Topic-prominence and zero NP-anaphora in Marathi and Hindi. In Mohammad Ali Jazayery and Werner Winter (eds.), *Languages and Cultures*, 309–328. Berlin: Mouton de Gruyter.
- Khin Khin Aye. 2005. *Bazaar Malay: history, grammar and contact*. Unpublished PhD thesis, University of Singapore.
- Kidwai, Ayesha. 2004. The topic interpretation in universal grammar: evidence from Kashmiri (and German). In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, 253–289. Boston: Kluwer Academic Publishing.
- Kittilä, Seppo. 2002. *Transitivity: towards a comprehensive typology*. Turku: University of Turku.
- Kittilä, Seppo. 2006. Object-, animacy- and role-based strategies: a typology of object marking. *Studies in Language* 30. 1–32.
- Kittilä, Seppo. 2010. Transitivity typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 346–367. Oxford: Oxford University Press.
- Koul, Omkar N. 2008. *Modern Hindi Grammar*. Hyattsville: Dunwoody Press.
- Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI Publications.
- Li, Charles N. and Sandra A. Thompson. 1976. Subject and Topic: A New Typology of Language. In Charles N. Li (ed.), *Subject and Topic*, 458–489. New York: Academic Press.
- Lim, Lisa. 2007. Mergers and acquisitions: on the ages and origins of Singapore English particles. *World Englishes* 26. 446–473.
- Loftman Bailey, Beryl. 1966. *Jamaican Creole Syntax: A Transformational Approach*. New York: Cambridge University Press.
- Mair, Christian. 2013. The world system of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide* 34. 253–278.
- Mufwene, Salikoko S. 2001. *Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Mugane, John M. 1997. *A Paradigmatic Grammar of Gĩkũyũ*. Stanford: CSLI Publications.
- Mukherjee, Joybrato and Stefan Th. Gries. 2009. Collostructional nativisation in New Englishes. *English World-Wide* 30. 27–51.

- Myachina, E.N. 1981. *The Swahili language: a descriptive grammar*. Trans. G. L. Cambell. London: Routledge & Kegan Paul.
- Neeleman, A., Szendrői, K. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38. 671–714.
- Parviainen, Hanna and Robert Fuchs. (submitted). An apparent-time investigation into clause-final focus particles in Asian Englishes.
- Patrick, Peter L. 2004. Jamaican Creole: morphology and syntax. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider and Clive Upton (eds.), *A Handbook of Varieties of English. Volume 2. Morphology and Syntax*, 407–438. Berlin/New York: Mouton de Gruyter.
- Platt, John T. 1982. English in Singapore, Malaysia and Hong Kong. In Richard V. Bailey and Manfred Görlach (eds.), *English as a World Language*, 384–414. Cambridge: Cambridge University Press.
- Platt, John, Heidi Weber and Ho Mian Lian. 1984. *The New Englishes*. London: Routledge and Kegan Paul.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rautionaho, Paula. 2014. *Variation in the Progressive: a Corpus-based Study into World Englishes*. PhD dissertation, University of Tampere.
- Sato, Yosuke. 2011. Radical pro drop and fusional pronominal morphology in Colloquial Singapore English: Reply to Neeleman and Szendrői. *Linguistic Inquiry* 42. 356–365.
- Schachter, Paul. 1976. The subject in Philippine languages. In Charles N. Li (ed.), *Subject and Topic*, 491–518. New York: Academic Press.
- Sharma, Devyani. 2012. Shared features in New Englishes. In Raymond Hickey (ed.), *Areal Features of the Anglophone World*, 211–232. Berlin: De Gruyter Mouton.
- Shibatani, Masayoshi. 1991. Grammaticization of topic into subject. In Elizabeth Closs Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization, Vol 2: Focus on types of grammatical markers*, 93–133. Amsterdam: John Benjamins.
- Siegel, Jeff. 1988. The development of Fiji Hindustani. In Richard K. Barz and Jeff Siegel (eds.), *Language Transplanted: the development of overseas Hindi*, 121–149. Wiesbaden: Harrassowitz.
- Smitherberg, Erik. 2005. *The Progressive in 19th-century English: A Process of Integration*. Amsterdam: Rodopi.
- Stafford, Roy Lawrence. 1967. *An Elementary Luo Grammar: with vocabularies*. Nairobi: Oxford University Press.
- Subbārāo, Kārumūri V. 2012. *South Asian Languages: A Syntactic Typology*. Cambridge: Cambridge University Press.
- Yip, Virginia and Steven Matthews. 2007. *The Bilingual Child: early development and language contact*. Cambridge: Cambridge University Press.

## Electronic sources

- Butt, Miriam. 2001. *Case, Agreement, Pronoun Incorporation and Pro-Drop in South Asian Languages*. (Handout, talk held at the Workshop on The Role of Agreement in Argument Structure, Utrecht) <http://ling.uni-konstanz.de/pages/home/butt/> (Accessed 13.2.2014)
- Census and Statistics Department, Hong Kong Special Administrative Region. *Thematic Household Survey Report* No. 51. <http://www.statistics.gov.hk/pub/B11302512013XXXXB0100.pdf> (Accessed 9.10.2013)

- Department of Statistics, Singapore, *Population trends 2013* [http://www.singstat.gov.sg/publications/publications\\_and\\_papers/population\\_and\\_population\\_structure/population2013.pdf](http://www.singstat.gov.sg/publications/publications_and_papers/population_and_population_structure/population2013.pdf) (Accessed 7.11.2013)
- Fiji Bureau of Statistics, *2007 Population Census* [http://www.spc.int/prism/fjtest/Census2007/census07\\_index2.htm](http://www.spc.int/prism/fjtest/Census2007/census07_index2.htm) (Accessed 6.6.2014)
- International corpus of English –homepage <http://ice-corpora.net/ice/importantnote.htm> (Accessed 27.11.2011)
- National Statistics Office, Republic of the Philippines, *2011 Philippines in Figures*. <http://www.census.gov.ph/sites/default/files/2013%20PIF.pdf> (Accessed 16.2.2014)
- Santa Barbara Corpus of Spoken American English -homepage <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus> (Accessed 1.9.2014)
- Kenya National Bureau of Statistics. *Population and Housing Census 2009, Volume 2- Population and Household Distribution by Socio-Economic Characteristics*. [http://www.knbs.or.ke/index.php?option=com\\_phocadownload&view=category&id=109:population-and-housing-census-2009&Itemid=599](http://www.knbs.or.ke/index.php?option=com_phocadownload&view=category&id=109:population-and-housing-census-2009&Itemid=599) (Accessed 26.6.2014)
- Kortmann, Bernd and Kerstin Lunkenheimer, (eds.). 2013. *The Electronic World Atlas of Varieties of English*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://ewave-atlas.org> (Accessed 25.11. 2013)

## Appendix

### Appendix 1: Chi-square values of the figures presented in Table 1

Overall distribution	X-squared = 148.034, df = 8, p-value < 2.2e-16***
IndE-SinE	X-squared = 6.0564, df = 1, p-value = 0.01386*
SinE-FjE	X-squared = 1.1206, df = 1, p-value = 0.2898
FjE-PhiE	X-squared = 0.3097, df = 1, p-value = 0.5779
PhiE-KenE	X-squared = 0, df = 1, p-value = 1
KenE-HKE	X-squared = 0.5521, df = 1, p-value = 0.4574
HKE-AmE	X-squared = 0.3231, df = 1, p-value = 0.5698
AmE-JaE	X-squared = 0, df = 1, p-value = 1
JaE-BrE	X-squared = 0.5839, df = 1, p-value = 0.4448



# **PUBLICATION IV**

**'I don't get time only': An apparent-time investigation of clause-final focus particles in Asian Englishes**

Hanna Parviainen & Robert Fuchs

Asian Englishes 21(3), 285-304

<https://doi.org/10.1080/13488678.2018.1520594>

**Publication reprinted with the permission of the copyright holders.**





# 'I don't get time only': an apparent-time investigation of clause-final focus particles in Asian Englishes

Hanna Parviainen & Robert Fuchs

To cite this article: Hanna Parviainen & Robert Fuchs (2018): 'I don't get time only': an apparent-time investigation of clause-final focus particles in Asian Englishes, *Asian Englishes*, DOI: [10.1080/13488678.2018.1520594](https://doi.org/10.1080/13488678.2018.1520594)

To link to this article: <https://doi.org/10.1080/13488678.2018.1520594>



Published online: 08 Oct 2018.



Submit your article to this journal [↗](#)





View Crossmark data [↗](#)

ARTICLE



# 'I don't get time only': an apparent-time investigation of clause-final focus particles in Asian Englishes

Hanna Parviainen <sup>a</sup> and Robert Fuchs <sup>b</sup>

<sup>a</sup>Faculty of Communication Sciences, University of Tampere, Tampere, Finland; <sup>b</sup>Institute of English and American Studies, University of Hamburg, Hamburg, Germany

## ABSTRACT

Research on the historical development of varieties of English usually tracks influence from established, so-called Inner Circle, varieties of English (such as British English) to less established Outer Circle varieties (such as Indian English). Recently, evidence of convergent patterns of language use in a number of Outer Circle Asian varieties has emerged, which might have been caused by the influence of Indian English in the region, but the existing evidence is strictly synchronic. Using the apparent-time method, this article investigates the focus particles 'also' and 'only', which have developed new uses in Asian Englishes. We provide evidence of language change in progress in Indian English and Hong Kong English, suggesting that Indian English might have influenced Hong Kong English, while evidence of Indian English influencing Philippine English remains tentative. The results lend support to a description of Indian English as a 'super-central variety' that has some influence over other varieties, thus challenging the traditional account of the development of varieties of English.

## ARTICLE HISTORY

Received 11 August 2017  
Accepted 3 September 2018

## KEYWORDS

Language change in  
apparent time; focus  
particles; Indian English;  
Philippine English; Hong  
Kong English

## 1. Introduction

For years, research on the characteristics of New Englishes was mainly based on small-scale studies or descriptions of individual varieties (see Kachru, 1990; Labru, 1984; Nihalani, Tongue, & Hosali, 1979; Parasher, 1988), while studies extending through several varieties were often more qualitative in nature (see, e.g., Platt, Weber, & Ho, 1983). The situation has, however, changed dramatically over the past decades; with the availability of such corpora as the International Corpus of English (ICE; Greenbaum, 1996), which provides comparable data on many varieties of English, it is now possible to conduct detailed studies of not only the qualitative, but also the quantitative aspects of the linguistic innovations that have emerged in New Englishes spoken around the world (see, e.g., Collins, 2009; Fuchs, 2016a; Werner, 2014, 2016).

A good example of a linguistic innovation in New Englishes is the tendency to place the focus particles 'also' and 'only' after their referents: 'I do not have to work also' (ICE-HK:S1A-004#967) and 'I don't get time only' (ICE-IND:S1A-052#37). Although focus particles following their referents are also licensed by the grammars of 'standard'



varieties such as British English (BrE) and American English (AmE), previous studies (see, e.g., Lange, 2007; Fuchs, 2012; Hiramoto, 2015; Parviainen, 2012) show that speakers of these varieties prefer to place ‘also’ and ‘only’ before their referents. The innovative use of clause-final particles exists in a number of Asian Englishes such as Indian English (IndE) (Bhatt, 2000; Lange, 2007; Sedlatschek, 2009; Balasubramanian, 2009; Fuchs, 2012), Philippine English (PhiE) (Parviainen, 2012), Singapore English (SinE) and Hong Kong English (HKE) (Hiramoto, 2015b; Parviainen, 2012).<sup>1</sup> Furthermore, Parviainen (2012) compares the frequencies in which this syntactic feature is used in these Englishes and suggests that IndE – the variety where clause-final ‘also’ and ‘only’ are used most frequently – could have contributed to the emergence of the feature in SinE, PhiE and HKE. However, there is currently no further evidence to support the view that the innovative use of focus particles in IndE predates their use in Southeast Asian varieties. The present article aims to examine this question further by trying to find evidence of language change in progress in Southeast Asian varieties of English.

At present, there is no comparative diachronic data enabling a real-time analysis on the use of this feature in Asian Englishes.<sup>2</sup> Fortunately, this does not preclude the possibility of studying the phenomenon with the apparent-time method (Labov, 2001), as many of the ICE corpora provide background information on their participants. By examining the age and gender of those who use innovative ‘also’ and ‘only’ in ICE-IND, ICE-PHI and ICE-HK, it is possible to estimate how established the feature is in each variety and, consequently, provide further evidence on the potential influence of IndE in the Southeast Asian region. If clause-final ‘also’ and ‘only’ are used more frequently by younger speakers than by older speakers and/or more frequently by female speakers than by male speakers, we can infer that these features are becoming more frequent in these varieties. On a methodological level, we demonstrate how pre-existing, synchronic corpus data from several varieties of a language can be used to reconstruct patterns of diachronic influence through the use of the apparent-time method (for previous studies applying this method to World Englishes, see Fuchs & Gut, 2015; Hansen, 2017a, 2017b). In addition, we will discuss the challenges that the application of this method presents and suggest how these can be addressed.

## 2. Previous studies

Before proceeding to the details of this syntactic innovation in Asian Englishes, we will briefly discuss the canonical use of focus particles. The most common function of a focus particle is to draw attention to a specific part of the utterance, which is called the particle’s constituent or focus. König (1991, pp. 32–33) distinguishes additive (1) and restrictive (2) focus particles, which are represented here by ‘also’ and ‘only’ respectively:<sup>3</sup>

- (1) ‘**Creativity** is *also* important’. (ICE-IND:S1A-044#146:1:A; emphasis added)
- (2) ‘There was uh no body there was *only* **flowers**’. (ICE-IND:S1A-012#141:1:B; emphasis added)

In (1), the function of ‘also’ is not only to draw attention to the word ‘creativity’, but also to suggest that there are other factors in addition to ‘creativity’ that should be considered important. The restrictive focus particle ‘only’ in (2) draws attention to the constituent following it (flowers), but unlike ‘also’ in (1) it excludes other potential alternatives (a body) as its centre of focus (see also Nevalainen, 1991, p. 31). Biber, Johansson, Leech, Conrad, & Finegan (1999, p. 781) argue that the location of ‘also’ and ‘only’ signals which element in the clause is focused, and thus they ‘often cannot be moved without affecting their meaning in the clause’ as in (3):

(3)‘(Only) they (only) fed (only) the cats (only)’. (Nevalainen, 1991, p. 39)

Despite all these possible locations, Quirk, Greenbaum, Leech, & Svartvik (1985, p. 605; also Fuchs, 2012, p. 37) argue that the most common location for a focus particle is in a clause-medial position, between the operator and the main verb, and thus, if the focus is on the subject, the particle follows it (4); but if the focus is on the verb or the object, the focus particle precedes its constituent (5):

(4)‘A smaller magnet has a small magnetic field a huge magnet will have a bigger magnetic field. Similarly **earth** *also* has a magnetic field’. (ICE-IND:S1B-019#141f.:1:A, from Fuchs, 2012, p. 36; emphasis original)

(5)‘So here we can calculate ... all the sides or you can *also* **multiply the side of the square by** four’. (ICE-IND:S1B-013#298:1:A, from Fuchs, 2012, p. 36; emphasis original)

Because focus in English is often located on the last word of a clause (Carter & McCarthy, 2006, p. 778), it is not common for a focus particle to occupy this position (see also Nevalainen, 1991, p. 33; Fuchs, 2012, p. 35). Furthermore, clause-final ‘also’ and ‘only’ are used (in BrE and AmE) more frequently in written than in spoken language; and in spoken genres, more frequently in formal than in informal discourse (Parviainen, 2012; Fuchs, 2012; see also Huddleston & Pullum, 2002, p. 590). Speakers of many Asian varieties, in contrast, have reversed this pattern: according to Parviainen (2012), IndE has a strong tendency to use clause-final focus particles in informal speech, where, more generally, colloquial, innovative features are used more frequently. This pattern is, moreover, mirrored in SinE, PhiE and HKE with lower frequencies. Therefore, there is a clear difference between the four Asian Englishes and their superstrates (BrE and AmE) regarding the frequency and contexts where clause-final ‘also’ and ‘only’ are used, which warrants further investigation into the possible causes behind this innovative tendency in the Asian varieties.

## 2.1 Clause-final focus particles in the substrates of IndE

The different functions which ‘also’ and ‘only’ have acquired in IndE have been examined by a number of scholars, who trace these innovative uses back to Indian substrates (see, e.g., Balasubramanian, 2009; Fuchs, 2012; Lange, 2007). Hindi, for instance, has two focus particles, *bhii* and *hii*, which according to Sharma (2003, pp.60–62) often signal inclusive (*bhii*) or exclusive (*hii*) focus. Thus, their meanings

resemble the canonical uses of ‘also’ and ‘only’ discussed in the previous section. The Hindi particles do, however, differ from canonical ‘also’ and ‘only’ in that they always follow their constituents (Sharma, 2003, p. 64), and therefore it is likely that their location in relation to their referents has also given rise to the tendency to place ‘also’ and ‘only’ after their referents in IndE.

Another factor that has contributed to the increased use of clause-final focus particles in IndE involves the new semantic meanings assigned to ‘also’ and ‘only’. Lange (2007, p. 106; see also Bhatt, 2000), who examines the use of ‘only’ in IndE, refers to this as the ‘non-contrastive or presentational focus marking’ (for ‘also’, see Fuchs, 2012), which means that they merely signal the presence of ‘new, nonpresupposed information’, as noted by Kiss (1998, p. 246). Examples of this presentational use are provided in (6) and (7), and as they are not compatible with an additive/restrictive meaning, it is clear that they differ greatly from the canonical uses shown in (1) and (2):

(6) ‘That’s now you’re offering offering me [a party] without my guests *also*’.

(ICE-IND:S1A-003#27:1:B)

(7) ‘Means ma- calculation should be finished off you know first only’.

(ICE-IND:S1A-071#83:1:A, cited in Lange, 2007b, p. 109)

This new usage can again be traced back to the Hindi particles *hii* and *bhii*, which can also take on non-contrastive, presentational uses (Lange, 2007, pp. 112–113; Fuchs, 2012, pp. 47–48; Parviainen, 2012, pp. 229–230). The elevated use of clause-final ‘also’ and ‘only’ in IndE can therefore be explained by<sup>4</sup> the joint effect of two factors, both of which can be traced back to the substrate(s): the tendency to place focus particles after their referents and the new additional meanings they have acquired.

## 2.2 Clause-final focus particles in the substrates of HKE and PhiE

The question of substrate influence is more complex for the other two varieties included in this study. The dominant substrate of HKE is Hong Kong Cantonese, which according to Hiramoto<sup>5</sup> (2015, p. 644) has two particles that closely resemble clause-final ‘only’: the sentence-final particle *zaa* and the adverb *ji hai* (the latter occurs in a sentence-final position only in colloquial Cantonese). However, equivalent particles for ‘also’ are not found in Cantonese (Hiramoto, 2015, p. 644), and thus substrate influence alone cannot conclusively explain why both clause-final particles are used equally frequently in HKE, as noted by Parviainen (2012, p. 237).

Also Tagalog, the most widely spoken substrate of PhiE, has particles with meanings close to the canonical uses of ‘also’ (*naming, din*) and ‘only’ (*lang, lamang*) (Ramos, 1971, pp. 49–50; Ramos & Cena, 1990, p. 99). Ramos and Cena (1990, p. 98) argue that these particles ‘normally occur after the first word of the sentence’, which Schachter (1990, p. 211) notes is typically the predicate. Although this does not, of course, preclude clause-final use of these particles (see, e.g., Ramos, 1971, pp. 49–50), the fact that they do not commonly occupy this position lowers the likelihood of these particles being the sole explanatory factor behind the pronounced tendency of PhiE to use clause-final ‘also’ and ‘only’.

### **2.3 Previous studies of IndE as an emerging epicentre**

There is some evidence that the influence of IndE has already extended beyond the borders of India. Hoffmann, Hundt, and Mukherjee (2011) as well as Hundt, Hoffmann, and Mukherjee (2012) examine the influence of IndE on the use of light verb constructions and the hypothetical subjunctive in other South Asian Englishes. Although they attribute many of their findings to various other factors, Hundt et al. (2012, p. 163) discover one feature, the use of ‘on if’ as synonymous with ‘whether’, which, they note, shows the potential of being a genuine IndE lexico-grammatical innovation that has spread to the other English varieties spoken in South Asia. While Hundt et al. (2012, p. 163) qualify their results, arguing that they need to be verified with diachronic data, which is currently not available, Gries and Bernaisch (2016) argue more forcefully in favour of the international influence of IndE. Their analysis of the dative alteration in the English of South Asian newspapers leads them to argue that ‘[IndE] may be regarded as the linguistic epicentre of English for South Asia’ (Gries & Bernaisch, 2016, p. 1).

As yet, there are only a few studies which have extended this line of research into Southeast Asia. Hogue (2001) shows that words from the Anglo-Indian lexicon (spoken by both expatriate British and English-speaking Indians) spread to other emerging Southeast Asian Englishes ‘as a direct result of British colonialism – either through the British themselves or the English-speaking Indians they brought to other [British] colonies’ (p. 190) in the region. Furthermore, Parviainen’s (2012, 2016) studies on the use of clause-final ‘also’ and ‘only’ and the invariant tag ‘isn’t it’ in Asian Englishes indicate that there are other innovative features which are used most frequently in IndE, followed by SinE, PhiE and HKE, which, according to Parviainen (2012, 2016), supports the notion that IndE could have contributed to the emergence of these features in some Southeast Asian varieties.

In summary, there is some evidence of both past non-epicentric influence of IndE in Southeast Asia and of current epicentric influence in South Asia – and potentially in Southeast Asia. However, evidence for language change in progress is lacking, and the present study attempts to fill this gap.

## **3. Data and methods**

### **3.1 Data**

The data used in the study come from the International Corpus of English (ICE), a family of corpora that contains both spoken and written language from several English varieties spoken around the world. The corpora follow a common design and contain one million words each, split into approximately 600,000 words of spoken English and 400,000 words of written English. Since the innovative use of ‘also’ and ‘only’ occurs most frequently in more informal spoken genres (Fuchs, 2012, p. 45; Parviainen, 2012, pp. 235–236 and 238–239), the present article will focus on the most informal subsection of the corpora, called ‘private conversations’, which contain approximately 200,000 words. The speaker data provided in conjunction with the Indian, Philippine and Hong Kong subcorpora of ICE (ICE-IND, ICE-PHI and ICE-HK) contain a number of demographic variables such as the gender and the approximate age of the speakers,

while further information such as their first language (L1) and home region/town are also occasionally provided.<sup>6</sup> Although Parviainen (2012) argues that innovative ‘also’ and ‘only’ are also frequently used in SinE, this variety had to be excluded from the present study, as speaker data are not provided with the ICE-SIN corpus – we will therefore focus only on IndE, PhiE and HKE. Furthermore, it is unlikely that the innovation has entered the Asian varieties from AmE or BrE since the tendency to use clause-final ‘also’ and ‘only’ is significantly less pronounced in the two superstrates when compared with the four Asian varieties as indicated by Fuchs (2012) and Parviainen (2012). Therefore, the two superstrates have also been excluded from the present study.

### 3.2 Methods

The apparent-time method (first proposed by Labov, 1963) is based on the generally accepted assumption that ‘once the features of the sociolect are established in the speech of young adults, under normal circumstances those features remain relatively stable for the rest of their lives’ (Chambers, 2009, p. 197), usually withstanding (at least to a degree) even the emergence of new linguistic changes adopted by the following generations in the community. Therefore, speakers from different age groups can be argued to represent the type of language that was used in their adolescence; by examining the frequency with which a feature is used by speakers from different age groups, it is possible to find out when that feature was first taken up by the younger speakers of the population. While the apparent-time method has been subject to criticism, comparisons between studies investigating language change in real and apparent time in the same speaker populations (see, e.g., Bailey, Winkle, Tillery, & Sand, 1991) show that it is indeed a useful tool for mapping linguistic change (see also Chambers, 2009, p. 213). The apparent-time construct could be called into question if adults undergo language change during their lifetimes. However, if they do, then they follow the change in the speech community, so that the apparent-time method would in such cases tend to underestimate the extent of the change (Boberg, 2004, p. 258; Cukor-Avila & Bailey, 2013, p. 254; Hansen, 2017a, 2017b; Labov, 1994, p. 97; Sankoff & Blondeau, 2007, p. 582), making results more conservative rather than leading to spurious conclusions of ongoing change.

In addition to age, gender has also been shown to be an important indicator of the changes that are taking place in the language used by a community. As Labov (2001) notes, ‘[w]omen have been found to be in advance of men in most of the linguistic changes in progress studied by quantitative means in the past several decades’ (p. 280). Although women frequently function as the instigators of linguistic change, it should be noted that this often applies to features that have overt prestige, whereas men are more likely to favour features with covert prestige (Chambers, 2009, pp. 235–238). At present, there are no attitudinal studies that have examined whether covert or overt prestige is assigned to the use of clause-final ‘also’ and ‘only’ by speakers of IndE, but some inferences can be drawn from the results of previous studies. Both Lange (2007, p. 115) and Parviainen (2012, p. 238) note that innovative ‘only’ is predominantly used in the local vernacular, and hence its use in more formal and written contexts is avoided due to a certain level of stigmatization of this feature in the variety. The case of innovative ‘also’, however, is more complex; Fuchs (2012, pp. 48–49) and Parviainen

(2012, p. 244) show that ‘also’, although used most frequently in informal spoken contexts, is also used frequently in the category of student essays, which could indicate that the feature is becoming increasingly accepted by younger speakers of IndE. This would also support the following argument by Balasubramanian (2009):

it is possible that younger users of Indian English, those who didn’t necessarily grow up with the same external norm (British English) that older speakers did, are responsible for the spread of Indianisms ... to the more formal registers. (p. 225)

Unfortunately, there are no studies addressing the possible stigmatization (or lack thereof) that might be associated with the use of ‘also’ and ‘only’ in HKE or PhiE. We will return to the question of prestige of the feature in different varieties in Section 4.3.

According to Labov (1994, pp. 79–82), linguistic change can be divided into five stages: completed, nearing completion, mid-range, new and vigorous, and incipient. This development is often illustrated with an S-curve where, at the beginning, the curve is low as the new form is used only by a small number of speakers. This is followed by a rapid rise in the curve as the feature is adopted by a growing proportion of the population, after which, in the final stage, the curve again stabilizes as the feature reaches saturation point. Therefore, if the relative proportion of the use of innovative ‘also’ and ‘only’ by young (and) female speakers is higher than that of older (and) male speakers in both PhiE and HKE (when compared with IndE), this would indicate that the use of the feature is a more recent innovation in the former two varieties. This, in turn, would lend further support to the argument that IndE is an emerging epicentre in South(east) Asia.

Previous studies show that for some Asian varieties such as IndE, the innovative use of ‘also’ and ‘only’ is not restricted to clause-final instances, but that the two particles can also follow their referents while occupying sentence-medial positions. Despite this, there are two reasons why the focus of the present paper is restricted to clause-final cases. Firstly, as noted earlier, speakers of BrE often place ‘also’ after its referent when the focused constituent is the subject (Fuchs, 2012, p. 36). Therefore, the innovative tendency of placing focus particles after their referents only applies to focused elements in the predicate. Secondly, since determining the referent of a focus particle in the clause-medial position is often difficult, our focus was restricted to cases where ‘also’ and ‘only’ unquestionably follow their referents; that is, when they are in the clause-final position (see also Parviainen, 2012, p. 232; Nayak, Singh, & Caldwell-Harris, 2016, p. 275).

The first step of the analysis consisted of searching for all instances of ‘also’ and ‘only’ in the data using WordSmith Tools 6.0 (Scott, 2012) (extra-corpus material was excluded). After careful consideration of each hit and its context, only clear instances of clause-final ‘also’ and ‘only’ were retained. A clause-final position was assumed in cases where the particle was followed by a pause or a coordinating conjunction, or where the speaker clearly ended their turn without interruption by another speaker. Furthermore, only tokens uttered by speakers for whom age and gender is known were used in the analysis. While the documentation of gender information in the corpora is straightforward (either male, female or no information), the categorization of age differs. All corpora use age ranges for some speakers, while for other speakers sometimes exact and at other times approximate ages are given. The latter cases were assigned to the respective age ranges for our analysis.

Furthermore, the age ranges differ between the corpora, and had to be subsumed to a common scale (below 26 years, 26–35 years, 36–50 years, above 50 years).

The final step in the analysis consisted of counting the number of tokens of clause-final ‘also’ and ‘only’ uttered by speakers of the different age and gender groups in each corpus, and counting the number of words contributed to the corpus by each of these groups. This allowed us to calculate the relative frequency of the two features per million words (pmw), and to apply linear regression models in order to determine: whether younger and/or female speakers are more likely to use the feature in question; and whether the three varieties differ in this regard as well as in the overall frequencies of clause-final ‘only’ and ‘also’. Two regression models were computed in R with relative frequency of clause-final ‘only’ and ‘also’, respectively, as dependent variables, and age, gender and variety as independent variables. Model selection was conducted using the step function, with *F*-tests as the selection criterion, and allowed for interactions of up to three variables. After model selection, post-hoc Tukey tests (corrected for multiple comparisons) were conducted with the lsmeans function from the eponymous R package (Lenth & Hervé, 2015).

## 4. Results

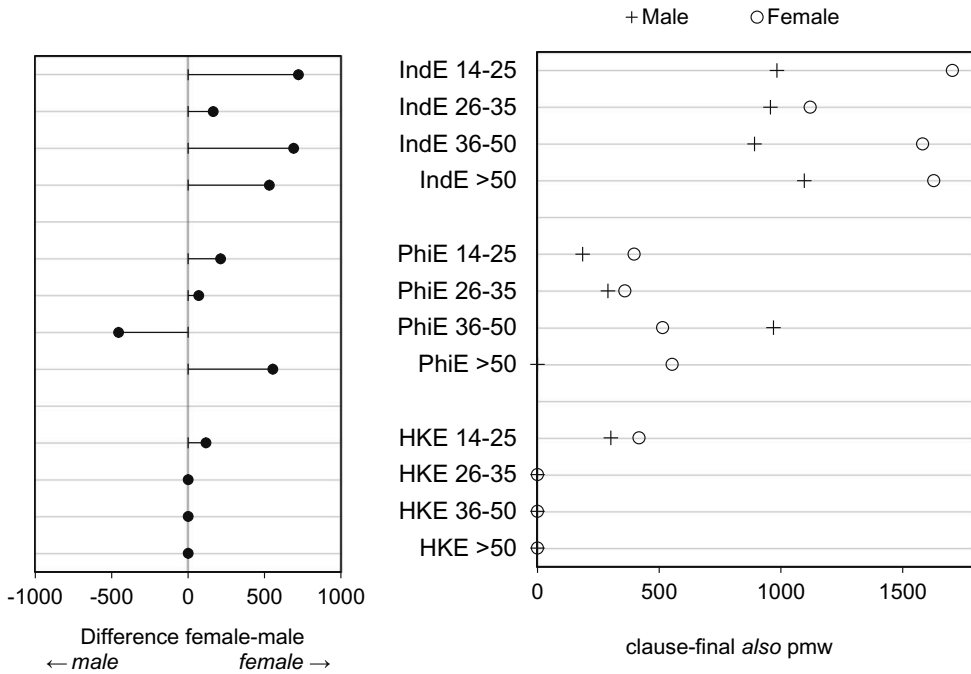
### 4.1 ‘Also’

For clause-final ‘also’, the regression model involved a main effect for independent variables variety and gender, as well as an interaction between these factors, but no main effect for age. As Figure 1 illustrates (see Appendix 1 for raw and relative frequencies), this focus particle is much more frequent in IndE than in PhiE (female  $t = 6.3$ ,  $p < 0.0001$ ; male  $t = 3.7$ ,  $p < 0.01$ ;  $df = 18$  for these and all following tests) and HKE (female  $t = 8.5$ ,  $p < 0.0001$ ; male  $t = 5.5$ ,  $p = 0.0001$ ). It is also more frequent in PhiE than in HKE, where it is in fact used exclusively by speakers below the age of 26 years, but the difference fails to reach significance (female  $t = 2.1$ ,  $p = 0.11$ ; male  $t = 1.7$ ,  $p = 0.22$ ). Clause-final ‘also’ is more frequent in female speech across all age groups in IndE, in all but one group in PhiE and in one group in HKE (as the focus particle was not used by any speaker above the age of 25 years). Nevertheless, the difference is only significant in IndE ( $t = 3.2$ ,  $p < 0.01$ ), but not in PhiE ( $t = 0.6$ ,  $p = 0.57$ ) or HKE ( $t = 0.2$ ,  $p = 0.86$ ).

While age was not included in the regression model as an influential factor, a glance at the data reveals that, in HKE, clause-final ‘also’ is used exclusively by speakers from the youngest group. In IndE and PhiE, by contrast, it is equally frequent in all age groups.

### 4.2 ‘Only’

For clause-final ‘only’, the regression model involved a main effect for independent variables variety and age, as well as an interaction between these factors, but no main effect for gender. This focus particle is much more frequent in IndE than in both PhiE and HKE in all but the oldest group of speakers (see Figure 2). This difference is significant in the youngest group of speakers, aged 14–25 years (IndE vs. PhiE  $t = 6.7$ ,  $p < 0.0001$ ; IndE vs. HKE  $t = 6.0$ ,  $p < 0.001$ ;  $df = 12$  for these and all following tests), the group aged 26–35 years (IndE vs. PhiE  $t = 3.8$ ,  $p < 0.01$ ; IndE vs. HKE  $t = 2.8$ ,  $p < 0.05$ )



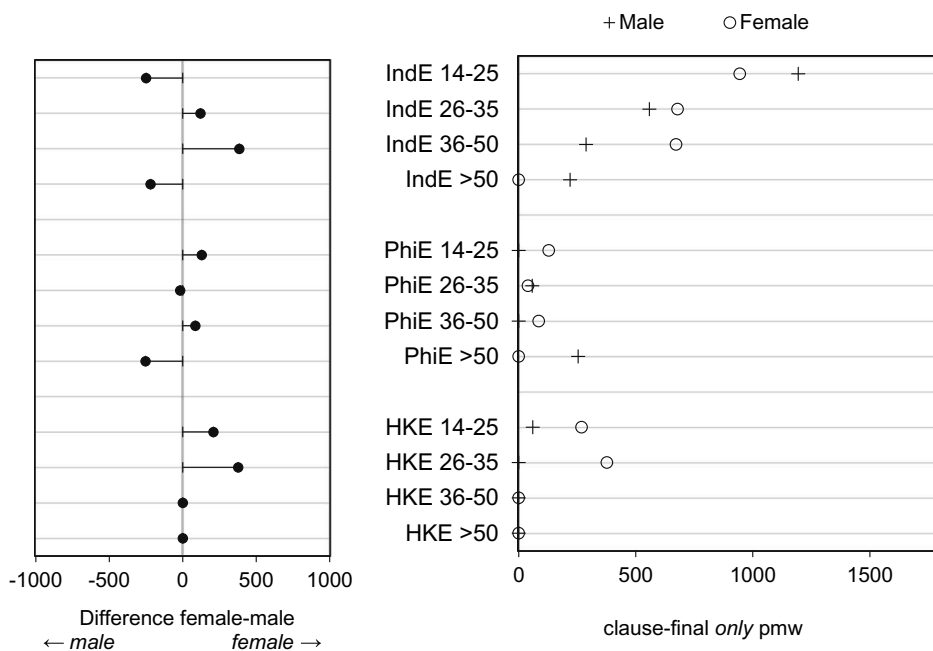
**Figure 1.** Relative frequency (pmw) of clause-final 'also' in IndE, HKE and PhiE by speaker age and gender (left-hand panel), and difference in relative frequency between male and female speakers across all varieties and age groups (right-hand panel). Figure created using a template provided by Sønning (2016).

and the group aged 36–50 years (IndE vs. PhiE  $t = 2.9$ ,  $p < 0.05$ ; IndE vs. HKE  $t = 3.2$ ,  $p < 0.05$ ). Among speakers above 50 years of age, the focus particle is more frequent in IndE than in HKE, but not significantly so ( $t = 0.7$ ,  $p = 0.8$ ), and is just as frequent in IndE as in PhiE ( $t = 0.1$ ,  $p = 0.99$ ).

As for speaker age, we observe a clear age trend in IndE, with clause-final 'only' decreasing almost steadily in frequency from the youngest to the oldest group of speakers. It is significantly more frequent in the youngest group of speakers compared to those aged 26–35 years ( $t = 3.0$ ,  $p < 0.05$ ), 36–50 years ( $t = 3.9$ ,  $p < 0.01$ ) and older than 50 years ( $t = 6.4$ ,  $p < 0.001$ ). The difference between those aged 26–35 and 36–50 years is not significant ( $t = 0.9$ ,  $p = 0.80$ ), but that between the 26–35 and 50+ age groups is ( $t = 3.4$ ,  $p < 0.05$ ). Finally, the difference between the 36–50 and 50+ age groups fails to reach significance ( $t = 2.5$ ,  $p = 0.12$ ). In HKE, we also observe an age trend, as the focus particle was found only in the speech of the two youngest groups, and was not used by older speakers. However, the differences between the groups do not reach significance. Finally, in PhiE there is no age trend, as clause-final 'only' is used at a low frequency across all age groups.

While gender was not included in the regression model as an influential factor, a glance at the data reveals that, in HKE, clause-final 'also' is more frequent in the language of female compared to male speakers. In IndE and PhiE, by contrast, the analysis of speaker gender reveals no clear trend. The focus particle is more frequent in





**Figure 2.** Relative frequency (pmw) of clause-final 'only' in IndE, HKE and PhiE by speaker age and gender (left-hand panel), and difference in relative frequency between male and female speakers across all varieties and age groups (right-hand panel). Figure created using a template provided by Sönning (2016).

the language of female speakers in some age groups, and more frequent in the language of male speakers in others.

### 4.3 Summary of results

Focusing on the extralinguistic variables of age and gender, and using the apparent-time method, this study set out to determine whether there is ongoing change in the usage of the clause-final focus particles 'only' and 'also' in HKE and PhiE (and, for comparison, IndE), and whether this is a feature that might have spread from IndE to the two Southeast Asian varieties. Specifically, we expected to find that younger speakers use these features more often than older speakers, and that female speakers use them more often than male speakers. This is based on the assumption that younger and/or female speakers are in most cases more advanced in ongoing language change than older and/or male speakers (Chambers, 2009; Labov, 1963, 2001). Furthermore, we expected that these differences would be more pronounced in HKE than in PhiE, as contact between IndE and HKE has historically been closer than with PhiE (Hogue, 2001).

For clause-final 'also', there was evidence of ongoing language change in IndE, HKE and, to a lesser extent, PhiE (see Table 1 for a summary of the results). In IndE, female speakers use this innovative focus particle significantly more frequently than male speakers, which also supports the notion that clause-final 'also' is not as stigmatized

**Table 1.** Overview of results on age and gender differences in the use of clause-final ‘also’ and ‘only’.

Clause-final		IndE	HKE	PhiE
‘also’	Age: younger > older speakers	–	(n.s.)	–
	Gender: female > male speakers	( $p < 0.01$ )	(n.s.)	(n.s.)
‘only’	Age: younger > older speakers	(up to $p < 0.001$ )	(n.s.)	–
	Gender: female > male speakers	–	(n.s.)	–

in IndE as clause-final ‘only’ has been argued to be (see Section 3.2). In HKE, younger speakers use clause-final ‘also’ more frequently than older speakers, and female speakers more frequently than male speakers. Although the differences are not statistically significant, they do give reason to suspect that clause-final ‘also’ is not considered to be a highly stigmatized feature in HKE either. In PhiE and IndE, younger speakers do not use this feature more often than older speakers. Female speakers use it somewhat more frequently than male speakers, which would suggest that the feature is becoming more frequent in these varieties, but for PhiE the differences were not statistically significant. In the case of HKE, we found tokens of clause-final ‘also’ exclusively in the youngest group of speakers. The focus particle was also used more frequently by female than by male speakers, and the gender difference was greater than for the other two varieties, although still not significant. Finally, in PhiE, female speakers use it more frequently than male speakers in all but one age group, although this difference is, again, not significant.

For clause-final ‘only’, there is evidence of its continuing diffusion in IndE, where it is used more frequently by younger compared to older speakers, but there is no gender difference. The fact that women did not favour ‘only’ to the same extent as ‘also’ further supports the notion that clause-final ‘also’ and ‘only’ are not equally acceptable among speakers of IndE. In HKE, clause-final ‘only’ was found in the speech of those 35 years old or younger, and not at all in the speech of those above the age of 35 years. As this feature is also used more frequently by female than by male speakers, we conclude that the evidence suggests that clause-final ‘only’ is becoming more frequent in HKE where it might not be considered to be as stigmatized a feature as in IndE, although this would need to be confirmed with attitudinal studies. Finally, in PhiE, no clear age or gender effects were found, suggesting that the frequency of clause-final ‘only’ is stable in this variety.

Overall, the evidence suggests that clause-final ‘also’ is a relatively established feature of IndE and PhiE, and may be increasing in frequency in the former. For HKE, the evidence is more clearly in favour of clause-final ‘also’ being a new and expanding feature. Furthermore, the lower frequency of this feature in HKE and PhiE, compared to IndE, is commensurate with the hypothesis of spread from IndE to HKE and possibly PhiE, although available evidence is significantly weaker for the latter variety. In summary, the results for clause-final ‘also’ are suggestive, and should be tested more explicitly once historical data on these varieties become available.

The evidence on clause-final ‘only’ is more conclusive. There are indicators that it is becoming more established in IndE, as there are clear age differences. For HKE, we found that this feature is used predominantly by younger and female speakers, while for PhiE no age or gender trends were revealed by the analysis. Consequently, the evidence suggests that the diffusion of clause-final ‘only’ is more advanced in IndE than in either

HKE or PhiE, and hence may have spread from IndE to HKE, while the results for PhiE remained inconclusive.

## 5. Discussion

Suggestive as the results are, they will still have to be tested using historical data once they become available for the three varieties. One potential drawback of our approach, which looks for gender and age differences in existing corpus data, is that the corpora are not balanced for these variables. For example, the private conversation sections of the corpora used in this study tend to provide more data for younger speakers than for older speakers. While this is not ideal, it is only a problem in so far as it decreases statistical power, and thus our chances of uncovering an existing difference between the two groups. Therefore, the data are not more likely to produce spurious results. Furthermore, even in cases where the available data are limited by corpus linguistic standards, these still compare favourably with non-corpus-based variationist research (see, e.g., Tagliamonte & D'Arcy, 2009). For example, for HKE, speakers above 35 years of age are represented by only about 10,000 words in ICE-HK, but this still represents 11 speakers for this group alone. Finally, except in one case (clause-final 'also' in IndE), our results do not reveal the U-shaped pattern characteristic of age-grading (Downes, 2005, p. 224), wherein a (usually informal) variant is more frequent in the speech of younger and older speakers than among middle-aged speakers, who avoid it due to its informality. In summary, our approach of using existing corpus data to track language change in progress is not more likely to lead to fictitious results than more traditional approaches that rely on balanced data sources. A larger and more balanced amount of data would permit the use of more sophisticated statistical techniques such as mixed effects regression modelling to account for variation between speakers and thus lead to more reliable results. However, in the absence of such a source of data, the ICE corpora and their rich metadata provide a ready source of evidence on ongoing language change in varieties of English (Hansen, 2017a). More generally, the apparent-time method is a well-established tool in research on language change. Even though adults may vary in their speech over their lifetime, if they change, they follow the change in the community (Boberg, 2004, p. 258; Cukor-Avila & Bailey, 2013, p. 254; Hansen, 2017a, 2017b; Labov, 1994, p. 97; Sankoff & Blondeau, 2007, p. 582).

### 5.1 *IndE as an epicentre*

Hundt (2013, p. 185) argues that for a variety to qualify as a potential epicentre, it must meet two criteria: firstly, it must be endonormatively stabilized (stage 4 in Schneider's, 2007, model); and, secondly, it must be able to function as a model for other neighbouring varieties. IndE can indeed be argued to meet the first of these requirements; although Schneider (2007, p. 165) suggests that IndE has not progressed further than the structural nativization stage (i.e. stage 3). Mukherjee (2007, pp. 168–169) argues that IndE had already progressed to the exonormative phase by the 1960s (if not earlier), and the latter view is now supported by a growing body of evidence (Hoffmann et al., 2011; Lambert, 2014). Further evidence comes from sociolinguistic interviews, which reveal an endo-normative orientation among young well-educated speakers, who are as accepting of IndE and its sub-varieties as they are critical of Indians (who have grown up in India) who

try to emulate a British or American accent (Fuchs, 2015, 2016b; Sridhar, 1989, p. 48). The results of the present article support a description of IndE as an endonormative variety, because the clause-final use of focus particles appears to be increasing among the younger generations (see also Balasubramanian, 2009, p. 225).

The second criterion mentioned by Hundt (2013, p. 185), whether IndE is able to function as a model for other varieties in the region, is also supported by some scholars; Mair (2013) categorizes IndE as a '(standard) super-central variety' in his World System of Standard and Non-Standard Englishes, and thus places the Indian variety on the same level as British, Australian, Nigerian and South African English. When discussing the role of IndE as a super-central variety, Mair (2013) notes the following:

There is (as yet) anecdotal evidence that hundreds of thousands of expatriate Indians working abroad in business and information technology are beginning to leave their mark on British and American English. It would be interesting to note whether migration, modern communication and media technology, combined with a craze for Bollywood-style entertainment, will be sufficient to establish Indian English norms as one relevant factor in the future development of the varieties spoken in the Indian diaspora communities in Asia, the Pacific and the Caribbean ... (p. 263)

The influence of these expatriate Indians has also been addressed by Kachru (1985, p. 28), who notes that the vast number of English-speaking South Asians who work in various expert positions (e.g. engineers and English teachers) in Southeast Asia have passed on some of their independently developed norms to other local Asian varieties of English, which might also explain how some Indian features could have spread to some Southeast Asian varieties of English.

Of the conditions that must hold for an alleged linguistic epicentre (such as IndE) to influence a potential recipient variety (such as HKE), so far we have considered conditions that must be met by the potential epicentre; that is, endonormativity and the potential to function as a model for other varieties. In addition, some properties of the potential recipient variety might also make influence from the epicentre more or less likely. A variety that has progressed further in its development (as defined by developmental stage in the Dynamic Model) is more likely to rely on its own norms and in turn less likely to be influenced by another variety, all other factors being the same. PhiE has remained in the structural nativization phase since the 1940s, while some restricted signs of endonormativity have also been detected in the local variety of English (Schneider, 2003, pp. 263, 2007, pp. 140–143), and hence the variety has advanced further in the model than HKE, which reached structural nativization in the 1960s but still exhibits some traces of exonormativity (Schneider, 2003, p. 258, 2007, p. 135). As a consequence, speakers of HKE are likely to be more receptive to influence from other major English varieties (such as IndE) spoken in the region. This would not be unprecedented, since Hogue (2001) has shown that some words from Anglo-Indian lexis spread to the English spoken in Hong Kong during colonial times, whereas fewer Anglo-Indian lexical items were borrowed into PhiE. In summary, we conclude that the preconditions for IndE to function as a linguistic epicentre influencing HKE and PhiE are met.

## 5.2 *IndE influence on HKE and PhiE*

While epicentral influence from IndE might have occurred, as argued in the previous section, the question remains whether there is enough linguistic evidence to support this scenario. Our results indicate that, for PhiE, only one criterion, the gender difference for clause-final ‘also’, supports our hypothesis of IndE influence in the variety. Furthermore, since Tagalog has focus particles which can sometimes be placed in clause-final position (see Section 2.2), substrate influence might be sufficient to explain our results on the use of clause-final ‘also’ and ‘only’ in PhiE. The possibility of independent development cannot be fully excluded in the case of HKE either. However, since Cantonese, the most significant substrate for HKE, lacks a particle that could be linked with the increased tendency to use clause-final ‘also’, substrate influence is unlikely to be the sole factor explaining our results for HKE. Thus, the remainder of this section will discuss the putative influence of IndE on HKE.

Although there is some proof of a historical connection between IndE and HKE, the results of the present study indicate that the increased tendency to use clause-final focus particles in HKE is a more recent development. The majority of the data for the three corpora was collected during the 1990s (ICE-IND, ICE-HK and ICE-PHI), while some data samples for ICE-HK and ICE-PHI were also collected during the first years of the twenty-first century. Because the speech of an individual reflects the type of language used in their adolescence, the language used by the informants of the younger groups (35 years or under) in all corpora represents trends in the varieties from the last decades of the twentieth century; this indicates that the use of the feature has become more common in IndE and HKE during this time period, although the feature is most established in the former, as noted earlier.

Interestingly, this time frame coincides with the rise of modern communication technology and a new wave of highly educated expatriate Indians moving to major metropolitan centres in Southeast Asia (and around the world), a development that has continued to this day.<sup>7</sup> This also means that during the past decades, younger generations in Southeast Asia have been increasingly exposed to IndE. Even though the overall number of Indians in Hong Kong is still small, in absolute terms it has doubled in recent decades (Singhvi, 2001, p. 277). Because of the positive image that many Southeast Asian countries have of the skilled Indian work force, it is likely that the variety of English that they use will also be viewed more positively – a further requirement for any variety functioning as a linguistic epicentre, as Hundt (2013, p. 198) notes.

In addition to the direct connections between IndE and HKE discussed earlier, other varieties spoken in the Southeast Asian region might also have contributed to the elevated influence of IndE on HKE; Hundt (2013) suggests that ‘[i]nfluence of an epicentre on a spatially non-adjacent variety of English is ... more like the spread of change from one urban centre to another’ (p. 189). Although SinE could not be included in the present study, the variety has the second-highest frequencies of clause-final focus particle use after IndE (Parviainen, 2012). In addition, both varieties are more advanced in Schneider’s (2007) Dynamic Model when compared with the other Southeast Asian varieties included in the present paper; IndE had already reached endonormativity by the 1960s (latest) (Mukherjee, 2007, p. 169) and SinE followed suit a decade later, in the 1970s (Schneider, 2007, p. 160).

Indians form a prominent minority in Singapore, some nine per cent of the population (Department of Statistics, Singapore, 2016), and they have always been well represented in trade and several branches of the government, thus contributing to their positive image in the Singaporean community (Rai, 2008, p. 36). Throughout its history, Singapore has received a steady flow of new immigrants from India (Rai, 2006, 2008), and hence SinE has also been constantly exposed to influence from IndE. For example, Platt et al. (1983) argue that due to the presence of Indian teachers in the Straights Settlements, which included today's Singapore, '[t]here are similarities in lexical and syntactic usage' (p. 8) between IndE and SinE. Furthermore, Parviainen's (2016; see also Mesthrie & Bhatt, 2008, p. 86) study on the use of the invariant tag 'isn't it' in Asian Englishes supports the notion that IndE could have contributed to the emergence of this feature in SinE.

Rai (2006, p. 188) notes that India and Singapore's close cultural, educational and economic connections have only increased during the past decades, which would also match with the time frame of our results. Shantakumar and Mukhopadhyaya (2008, p. 573) argue that the 1990s marked an unprecedented rise in Indian migration flows to Singapore, remaining at an elevated level ever since, partially due to the active support provided by the Singaporean government (Rai, 2006, p. 185). One of the consequences of this growing interaction between speakers of IndE and SinE could be the increased use of clause-final focus particles by speakers of SinE, which in turn could have promoted the use of clause-final 'also' and 'only' in Southeast Asia even further. Singapore is a major English-speaking metropolitan city-state in the region, and therefore it has the potential to function as one of the channels through which influence from IndE could spread to the region's other major urban centres such as Hong Kong. Indeed, this would mirror the patterns in which Indian influences have spread in the region in the past (see, e.g., SarDesai, 1997, p. 77; Hogue, 2001; Sridharan & Sevea, 2006, p. 198). However, to gain any insights into this matter, speaker information for the ICE-SIN informants would be required, which is not at present available.

## 6. Conclusions

The results of the present study indicate that the use of clause-final 'only' and, to some extent, 'also' are relatively established in IndE, whereas in HKE, and possibly also PhiE, their use seems to be of more recent origin. Although the results do lend support to the notion that influence from IndE on Southeast Asian varieties is not restricted to colonial times but has continued in the postcolonial era, the issue should be further examined by using real-time diachronic data. Furthermore, the potential role of IndE as an emerging epicentre should be investigated further based on attitudinal and sociolinguistic data (as argued by Hoffmann et al., 2011, p. 277), which would add to the existing corpus linguistic evidence, and provide insights into the role and status that users of English in Southeast Asia ascribe to IndE features. Our results, in turn, demonstrate how the apparent-time method can be usefully applied to existing, synchronic corpus data in order to contrast diachronic developments in distinct, but related, varieties. Where real-time data are available, it is arguably to be preferred over apparent-time analyses. Nevertheless, real-time data are often unavailable or costly to compile, making apparent-time analyses a viable source of evidence for research on varieties of English and beyond.

## Notes

1. Innovative uses of focus particles are not unique to IndE or South(east) Asian Englishes. In a comparable case, Nigerian English has developed innovative uses of the focus particles 'even' and 'still' (Fuchs, Gut, & Soneye, 2013).
2. The only available diachronic corpora on Asian Englishes are the Phil-Brown corpus, which provides data on PhiE from the late 1950s to the early 1960s (Collins, Borlongan, & Yao, 2014), and the Kolhapur Corpus, which contains data on IndE from the year 1978. However, since neither corpus contains spoken data, they could not be used for this study.
3. Fuchs (2012, pp. 30–31) mentions two additional uses for 'also', as a conjunctive adverbial and as a correlative of coordinating conjunctions. However, the additive use is the most frequent of the three.
4. Similar emphatic particles following their referent can also be found in, for example, Tamil, Marathi and Malayalam (see Asher, 1985, p. 94; Asher, Annamalai, & Kumari, 1997, p. 381; Pandharipande, 1997, pp. 239–241).
5. Hiramoto (2015) studies the use of sentence-final adverbs in SinE and HKE and their substrates. However, based on an example from ICE-HK (Hiramoto, 2015, p. 640), data tagged as extra corpus text (<X> ... </X>) were not excluded from the searches. Therefore, the numerical results presented for ICE-SIN and ICE-HK should be interpreted with caution.
6. We would like to thank Beke Hansen for not only making us aware of the mistakes in the original ICE-IND metadata, but also for her help in allocating these data to the correct speakers.
7. Although the growing number of Indian white-collar workers has been described as one of the driving forces behind the rise of IndE (see, e.g., Mair, 2013), unfortunately no statistics are available on their exact numbers or their professional fields.

## Acknowledgements

During work on this manuscript, the first author was supported by the Academy of Finland.


## Disclosure statement


No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the Academy of Finland [Grant 269114].

## ORCID

Hanna Parviainen  <http://orcid.org/0000-0002-0723-5692>

Robert Fuchs  <http://orcid.org/0000-0001-7694-062X>

## References

- Asher, R. E. (1985). *Tamil*. Amsterdam: Routledge.
- Asher, R. E., Annamalai, E., & Kumari, T. C. (1997). *Malayalam*. London, New York: Routledge.
- Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change*, 3, 26–264.
- Balasubramanian, C. (2009). *Register variation in Indian English*. Amsterdam: John Benjamins.

- Bhatt, R. (2000). Optimal expressions in Indian English. *English Language and Linguistics*, 4, 69–95.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson ESL.
- Boberg, C. (2004). Real and apparent time in language change: Late adoption of changes in Montreal English. *American Speech*, 79(3), 250–269.
- Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide: Spoken and written English grammar and usage*. Cambridge: Cambridge University press.
- Chambers, J. K. (2009). *Sociolinguistic theory: Linguistic variation and its social significance* (Revised ed.). Malden: Blackwell.
- Collins, P. (2009). *Modals and quasi-modals in English*. Amsterdam: Rodopi.
- Collins, P., Borlongan, A. M., & Yao, X. (2014). Modality in Philippine English: A diachronic study. *Journal of English Linguistics*, 42(1), 68–88.
- Cukor-Avila, P., & Bailey, G. (2013). Real time and apparent time. In J. K. Chambers & N. Schilling (Eds.), *The handbook of language variation and change* (2<sup>nd</sup> ed., pp. 239–262). Malden, Mass: Wiley-Blackwell.
- Department of Statistics, Singapore, *Population trends 2016* Retrieved from [http://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications\\_and\\_papers/population\\_and\\_population\\_structure/population2016.pdf](http://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/population_and_population_structure/population2016.pdf) (Accessed 18.9.2018)
- Downes, W. (2005). *Language and society* (2nd ed.). Cambridge: Cambridge University Press.
- Fuchs, R. (2012). Focus marking and semantic transfer in Indian English. *English World-Wide*, 33(1), 27–52.
- Fuchs, R. (2015). You're not from around here, are you? – A dialect discrimination experiment with speakers of British and Indian English. In E. Delais-Roussarie, A. Mathieu, & S. Herment (Eds.), *Prosody and language in contact: L2 acquisition, attrition and languages in multilingual situations* (pp. pp. 121–146). Berlin: Springer.
- Fuchs, R. (2016a). The frequency of the present perfect in varieties of English around the world. In V. Werner, C. Suárez-Gómez, & E. Seoane (Eds.), *Re- assessing the present perfect in English: Corpus studies and beyond* (pp. pp.223–258). Berlin: de Gruyter.
- Fuchs, R. (2016b). *Speech rhythm in varieties of English: Evidence from educated Indian English and British English*. Berlin: Springer.
- Fuchs, R., & Gut, U. (2015). An apparent time study of the progressive in Nigerian English. In P. Collins (Ed.), *Grammatical change in English world-wide* (pp. pp. 373–387). Amsterdam: Benjamins.
- Fuchs, R., Gut, U., & Soneye, T. (2013). “We just don’t even know”: The usage of the pragmatic focus particles even and still in Nigerian English. *English World-Wide*, 34(2), 123–145.
- Greenbaum, S. (Ed.). (1996). *Comparing English worldwide: The international corpus of English*. Oxford: Clarendon Press.
- Gries, S. T., & Bernaisch, T. (2016). Exploring epicentres empirically: Focus on South Asian English. *English World-Wide*, 37(1), 1–25.
- Hansen, B. (2017a). The ICE metadata: A window to the past? An exploratory study of Hong Kong English. To appear. *World Englishes*, 36(3), 4714–4786.
- Hansen, B. (2017b). Modal verbs of obligation and necessity in World Englishes. Unpublished PhD dissertation, University of Kiel.
- Hiramoto, M. (2015). Sentence-final adverbs in Singapore English and Hong Kong English. *World Englishes*, 34(4), 636–653.
- Hoffmann, S., Hundt, M., & Mukherjee, J. (2011). Indian English – An emerging epicentre? A pilot study on light-verbs in web-derived corpora of South Asian Englishes. *Anglia*, 12(3–4), 258–280.
- Hogue, C. (2001). The spread of Anglo-Indian words into South-East Asia. In B. Moore (Ed.), *Who’s centric now? The present state of post-colonial Englishes* (pp. pp. 165–197). Melbourne: Oxford University Press.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.



- Hundt, M. (2013). The diversification of English: Old, new and emerging epicentres. In D. Schreier & M. Hundt (Eds.), *English as a contact language* (pp. pp.182–203). Cambridge: Cambridge University Press.
- Hundt, M., Hoffmann, S., & Mukherjee, J. (2012). The hypothetical subjunctive in South Asian Englishes – Local developments in the use of a global construction. *English World-Wide*, 33(2), 147–164.
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. pp. 11–30). Cambridge: Cambridge University Press.
- Kachru, B. B. (1990). A nativized variety: The South Asian case. In B. B. Kachru (Ed.), *The alchemy of English* (pp. pp. 33–56). Urbana/Chicago: University of Illinois Press.
- Kiss, K. É. (1998). Identificational focus versus information focus. *Language*, 74, 245–274.
- König, E. (1991). *The meaning of focus particles. A comparative perspective*. London: Routledge.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19, 273–309.
- Labov, W. (1994). *Principles of linguistic change. Vol 1. Internal Factors*. Oxford: Blackwell.
- Labov, W. (2001). *Principles of linguistic change. Vol 2. Social factors*. Oxford: Blackwell.
- Labru, G. L. (1984). *Indian newspaper English*. Delhi: B. R. Publishing.
- Lambert, J. (2014). Diachronic stability in Indian English lexis. *World Englishes*, 33(1), 112–127.
- Lange, C. (2007). Focus marking in Indian English. *English World-Wide*, 28, 89–118.
- Lenth, R., & Hervé, M. (2015). lsmeans: Least-Squares Means. R package version 2.17. Retrieved from <http://CRAN.R-project.org/package=lsmeans>
- Mair, C. (2013). The world system of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide*, 34(3), 253–278.
- Mesthrie, R., & Bhatt, R. (2008). *World Englishes*. Cambridge: Cambridge University Press.
- Mukherjee, J. (2007). Steady states in the evolution of new Englishes: Present-day Indian English as an equilibrium. *Journal of English Linguistics*, 35, 157–187.
- Nayak, S., Singh, I., & Caldwell-Harris, C. (2016). Familiarity, comprehension and use of Indian English only. *English World-Wide*, 37, 267–292.
- Nevalainen, T. (1991). *But, only, just: Focusing adverbial change in modern English 1500–1900*. Helsinki: Société Néophilologique.
- Nihalani, P., Tongue, R. K., & Hosali, P. (1979). *Indian and British English: A handbook of usage and pronunciation*. Delhi: Oxford University Press.
- Pandharipande, R. V. (1997). *Marathi*. London, New York: Routledge.
- Parasher, S. V. (1988). Certain stylistic features in Indian English and their relationship to degrees of formality. *Indian Journal of Applied Linguistics*, 14(2), 95–122.
- Parviainen, H. (2012). Focus particles in Indian English and other varieties. *World Englishes*, 31(2), 226–247.
- Parviainen, H. (2016). Use of the invariant tag isn't it in Asian Englishes. *World Englishes*, 35(1), 98–117.
- Platt, J., Weber, H., & Ho, M. L. (1983). *Singapore and Malaysia*. Amsterdam: John.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.
- Rai, R. (2006). Singapore. In B. V. Lal, P. Reeves, & R. Rai (Eds.), *The Encyclopedia of the Indian Diaspora* (pp. pp. 176–188). Honolulu: University of Hawai'i Press.
- Rai, R. (2008). 'Positioning' the Indian diaspora: The South-east Asian experience. In P. Raghuram, A. K. Sahoo, B. Maharaj, D. Sangha (Eds.), *Tracing an Indian diaspora: Contexts, memories, representations* (pp. 29–51). New Delhi: Sage.
- Ramos, T. V. (1971). *Tagalog Structures*. Honolulu: University of Hawaii Press.
- Ramos, T. V., & Cena, R. M. (1990). *Modern Tagalog: Grammatical explanations and exercises for non-native speakers*. Honolulu: University of Hawaii Press.
- Sankoff, G., & Blondeau, H. (2007). Language change across the lifespan:/R/in Montreal French. *Language*, 83(3), 560–588.
- SarDesai, D. R. (1997). *Southeast Asia: Past and present* (4<sup>th</sup> edition). Boulder: Westview Press.

- Schachter, P. (1990). Tagalog. In B. Comrie (Ed.), *The major languages of East and South-East Asia* (pp. pp. 208–230). London: Routledge.
- Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. *Language*, 79, 233–281.
- Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Scott, M. (2012). *Wordsmith Tools version 6*. Liverpool: Lexical analysis software.
- Sedlatschek, A. (2009). *Contemporary Indian English: Variation and change*. Amsterdam, Philadelphia: John Benjamins.
- Sharma, D. (2003). Discourse clitics and constructive morphology in Hindi. In M. Butt & T. Holloway King (Eds.), *Nominals: Inside and out* (pp. pp. 59–84). Stanford, CA: CSLI.
- Singhvi, L. M. (2001). Report of the high level committee on the Indian diaspora. *Government of India, Ministry of External Affairs, Non-Resident Indian and Persons of Indian Origin Division*.
- Sönning, L. (2016). The dot plot: A graphical tool for data analysis and presentation. In H. Christ, D. Klenovsak, L. Sönning, & V. Werner (Eds.), *A Blend of MaLT. Select Contributions from the Methods and Linguistic Theories Symposium*. Bamberg: University of Bamberg Press. (Vol. 2015, pp. 101–129).
- Sridhar, K. K. (1989). *English in Indian Bilingualism*. New Delhi: Manohar.
- Sridharan, K., & Sevea, R. S. (2006). The Philippines. In B. V. Lal, P. Reeves, & R. Rai (Eds.), *The encyclopedia of the Indian diaspora* (pp. pp. 198–199). Honolulu: University of Hawai'i Press.
- Tagliamonte, S. A., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation and language change. *Language*, 85(1), 58–108.
- Werner, V. (2014). *The present perfect in world Englishes: Charting unity and diversity*. Bamberg: University of Bamberg Press.
- Werner, V. (2016). Rise of the undead? BE-prerfects in world Englishes. In V. Werner, C. Suárez-Gómez, & E. Seoane (Eds.), *Re-assessing the present perfect in English: Corpus studies and beyond* (pp. 259–294). Berlin: de Gruyter.

## Appendix 1. Frequency of clause-final focus particles ‘also’ and ‘only’ in ICE-IND, ICE-HK and ICE-PHI (section S1A, private conversations) by speaker age and gender

Focus particle	Variety	Gender	Age (years)	Number of speakers	Number of tokens	Number of words	Relative frequency (pmw)
‘also’	HKE	Female	14–25	84	45	107,796	417
	HKE	Female	26–35	3	0	5304	0
	HKE	Female	36–50	6	0	5888	0
	HKE	Female	> 50	0	0	0	0
	HKE	Male	14–25	15	5	16,599	301
	HKE	Male	26–35	2	0	2892	0
	HKE	Male	36–50	4	0	3204	0
	HKE	Male	> 50	1	0	902	0
	IndE	Female	14–25	52	74	43,410	1705
	IndE	Female	26–35	33	33	29,456	1120
	IndE	Female	36–50	27	40	25,282	1582
	IndE	Female	> 50	9	11	6757	1628
	IndE	Male	14–25	18	14	14,236	983
	IndE	Male	26–35	26	24	25,082	957
	IndE	Male	36–50	37	31	34,771	892
	IndE	Male	> 50	37	35	31,927	1096
	PhiE	Female	14–25	170	37	93,197	397
	PhiE	Female	26–35	78	9	25,083	359
	PhiE	Female	36–50	54	6	11,674	514
	PhiE	Female	> 50	0	2	3613	554
	PhiE	Male	14–25	69	6	32,386	185
	PhiE	Male	26–35	75	5	17,283	289
	PhiE	Male	36–50	106	6	6190	969
	PhiE	Male	> 50	1	0	3937	0
‘only’	HKE	Female	14–25	84	29	107,796	269
	HKE	Female	26–35	3	2	5304	377
	HKE	Female	36–50	6	0	5888	0
	HKE	Female	> 50	0	0	0	0
	HKE	Male	14–25	15	1	16,599	60
	HKE	Male	26–35	2	0	2892	0
	HKE	Male	36–50	4	0	3204	0
	HKE	Male	> 50	1	0	902	0
	IndE	Female	14–25	52	41	43,410	944
	IndE	Female	26–35	33	20	29,456	679
	IndE	Female	36–50	27	17	25,282	672
	IndE	Female	> 50	9	0	6757	0
	IndE	Male	14–25	18	17	14,236	1194
	IndE	Male	26–35	26	14	25,082	558
	IndE	Male	36–50	37	10	34,771	288
	IndE	Male	> 50	37	7	31,927	219
	PhiE	Female	14–25	170	12	93,197	129
	PhiE	Female	26–35	78	1	25,083	40
	PhiE	Female	36–50	54	1	11,674	86
	PhiE	Female	> 50	0	0	3613	0
	PhiE	Male	14–25	69	0	32,386	0
	PhiE	Male	26–35	78	1	17,283	58
	PhiE	Male	36–50	106	0	6190	0
	PhiE	Male	> 50	1	1	3937	254



