

This is a post-review, non-copyedited version accepted on 2016-10-13 for publication in the *Journal of Pidgin and Creole Languages*. Please quote from the published version at [dx.doi.org/10.1075/jpcl.00002.deu](https://doi.org/10.1075/jpcl.00002.deu)

Singaporean internet chit chat compared to informal spoken language: linguistic variation and indexicality in a language contact situation

Dagmar Deuber, Jakob R. E. Leimgruber and Andrea Sand
University of Muenster, Germany / University of Freiburg, Germany /
University of Trier, Germany

This paper compares data from a Singaporean chit chat forum to informal spoken data. We first perform a qualitative analysis of text samples in a framework of indexicality. Then we present quantitative findings for two (sets of) features each of the contact variety Singlish (particles, the *kena*-passive) and spoken English in general (discourse markers, contractions). For the former some similarities are observed but we also find that the forum contributors tend to be creative and innovative in their choice of particles. In this connection we argue that they index specific subgroup identities and further point out that Singlish is a rather flexible set of resources. Our findings differ from those of previous research on Jamaican Creole as used in an internet forum, thus showing that the use of contact varieties in computer-mediated communication can take different forms. The general features of spoken English are used comparatively less in the forum data, indicating that using Singlish features in writing is not tantamount to writing down spoken language. Moreover, we draw attention to features of computer-mediated communication in the chit chat forum data. Finally we discuss implications in terms of the Dynamic Model of the evolution of Postcolonial Englishes (Schneider 2007).

Keywords: Singlish, computer-mediated communication, sociolinguistic approach

1. Introduction

Concomitant with the increasingly widespread use of new communication technologies, the past decade or so has seen a rising scholarly interest in the

use of creoles and other contact varieties in the context of computer-mediated communication (CMC). Lim and Ansaldo in their cutting-edge book-length treatment of *Languages in Contact* identify CMC as one of today's major topics that should be given more prominence in language contact research (2016:196-203). Based in particular on the example of Jamaican Creole, the contact variety whose use in CMC has so far been studied most widely, they observe that CMC is significant not just in broadening the scope of use of contact varieties but in providing a context where further evolution of varieties takes place (2016:200).

Singapore, the context in which the present investigation of CMC is situated, is a language contact environment par excellence. Intense contact between English and Asian languages in the city state has given rise to a characteristic vernacular form of English, Singlish, which is in fact so considerably restructured that it has been suggested to consider it as 'creolised' (Ansaldo 2004:143). It is true that Singlish is difficult to fit neatly into one of the traditional categories of contact languages (see e.g. Bao 2015:2 for a brief discussion). In any case, however, these were developed mainly on the basis of a limited range of varieties comprising in particular the Atlantic pidgins and creoles. For present purposes it is more productive to adopt an integrated approach as proposed by Lim and Ansaldo (2016:11), which deemphasizes traditional categories and, furthermore, brings language contact settings in Asia more into the picture.

The form and use of Singlish have continually been shaped by developments in Singapore's dynamic language situation. The informal speech of today's young generation is often characterized by a mix of Singlish and Standard English features, such that 'there is no easily identifiable "matrix language", or even a hint at which is the default code' (Leimgruber 2012:8). So pervasive is this type of mixed discourse that Wee (2011:78-79) argues that Singlish, because it is not normally used for entire exchanges, is not a 'fully extensive social language', although public discourse has treated it as such.

Our particular concern will be the nature of the written counterpart of the type of mixed speech that has been described e.g. by Leimgruber (2012). Therefore we have chosen data from a chat forum, where one is likely to find the most informal type of interaction available in a forum context (cf. Collot & Belmore 1996:18). We propose to combine aspects from different strands of CMC research, namely the one focusing on the consequences of using language in the new medium of CMC, a question which particularly preoccupied researchers at the first stages of work in the area, and a current more sociolinguistically oriented strand. In order to be able to situate language use in the forum in the context of the larger Singaporean speech community, we will, as a first step, compare it to informal spoken language. We will focus particularly on Singlish features, but, in view of the way they are intertwined with Standard English features in spoken discourse, we do not limit ourselves to these, but consider general features of spoken English as well. Besides these, our description of language use in the forum will take into account

CMC language features. The patterns of language use found will be interpreted in terms of indexicality on the levels of stances and identities. The indexical approach (Silverstein 2003; Eckert 2008) offers a comprehensive framework for analysing the sociolinguistic expression of interactional, interpersonal, attitudinal and cultural stances as well as socially and discursively constructed social identities. It is associated with the third wave of variation studies, according to Eckert (2012). As described by Eckert (2012), the earlier waves were both concerned with correlating linguistic variables and social categories. Speakers' identity was viewed just in terms of social categories, which linguistic variation was considered to reflect. The categories studied were broad demographic ones such as socioeconomic class and gender in the first wave, whereas second wave studies focused on more specific local categories. Third wave studies, in contrast, see social meaning as constructed through linguistic practice (Eckert 2008, 2012). In this perspective variables are considered to index specific stances and qualities which can be combined to constitute categories. An important notion in this context is that of the 'indexical field'; Eckert (2008:454) defines this as a 'constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable', further pointing out that such a field is fluid and open to change.

In applying a sociolinguistic approach grounded in indexicality we will be interested in the implications for the current and future development of English in Singapore. To explore these we will draw on Schneider's (2007) Dynamic Model of the evolution of Postcolonial Englishes, as this includes a major focus on identity constructions. Postcolonial Englishes, importantly for present purposes, are considered in Schneider's framework to comprise the full range of varieties in postcolonial contexts, from (emerging) standards to restructured varieties including creoles. So far, it has been received mainly as a model for the evolution of the former rather than the latter type of variety and it has thus become associated more with the field of New English studies than with creolistics. Such a distinction, however, is not pertinent for present purposes. We are, in any case, dealing with a situation of high heterogeneity, and from the point of view of indexicality, as Bao (2015:3) has rightly argued, what matters is not 'whether Singapore English is a pidgin, creole, or New English' but that it constitutes 'the linguistic tool with which the speakers navigate the social and cultural mosaic of Singapore'. Moreover, from a more theoretical point of view, not only is it productive to look across the boundaries of traditional categories of contact varieties, but across the disciplinary boundaries of creolistics and New Englishes studies as well (Bao 2015:2; Lim and Ansaldo 2016:13).

We proceed by first outlining the major relevant strands in CMC research, which provides the point of departure for our research questions (section 2). Then follow a survey of previous related research on Singaporean CMC (section 3) and sections on the sociolinguistic context and the data (4 and 5, respectively). Next we present a qualitative analysis of sample extracts

from the forum (section 6) and from there we move on to report and discuss the results of a comparative quantitative analysis of the forum data and the spoken data (section 7). The final section (8) is devoted to conclusions and implications.

2. Major strands in CMC research and present research questions

In view of the fact that many of these varieties are not standardized and were not widely used in writing prior to the advent of CMC, it is not surprising that questions of orthography and particularly informal orthographic standardization have been among the major concerns in creolistic approaches to CMC (see Hinrichs 2004, Hinrichs & White-Sustaíta 2011, and Moll 2012, 2015:chapter 4 on Jamaican Creole; Deuber & Hinrichs 2007 on Nigerian Pidgin and Jamaican Creole; Rajah-Carrim 2008 on Mauritian Creole; Oenbring 2013 on Bahamian Creole). In studies of CMC in standardized languages, in contrast, one major research focus has been the way CMC language differs from and challenges standard written language by means of nonstandard orthography and other special features. In fact, research into CMC-specific language use dominated what has been described as the ‘first wave’ of studies in the area, which goes back to the 1990s (Androutopoulos 2006:420). One question that received a large amount of attention in this first wave was the extent to which the language of CMC incorporated features of spoken language. At an early stage the new medium even made researchers think about whether the language of email, for instance, could be something like ‘written speech’ (Maynor 1994), i.e. similar to speech despite appearing in the written medium. Soon, however, it became increasingly clear that CMC defies such a label because it is hybrid and variable in nature, displaying features of speech and of writing to varying degrees in its different manifestations (see e.g. Collot & Belmore 1996). By now, it can be taken as established that the more synchronous and the more dialogic a type of CMC is, the more speech-like it is likely to be, although there is also a great deal of variation within CMC types (see e.g. Runkehl, Schlobinski & Siever 1998:116; Baron 2003; Dürscheid 2004). At the same time, it has been widely recognized that even in its more speech-like manifestations, CMC is not in fact ‘written speech’. Crystal writes in *Language and the Internet* that ‘Netspeak is better seen as written language which has been pulled some way in the direction of speech than as spoken language which has been written down’ (2006:51). This statement ties in with findings to the effect that features of spoken language are recurrent but not necessarily particularly frequent even in dialogic CMC (see e.g. Runkehl, Schlobinski & Siever 1998:60-63 on German language newsgroups; Lewin & Donner 2002 on English language bulletin boards). Crystal himself singles out response forms such as *mhm* and *yeah* as well as comment clauses like *you know* and *you see* as spoken language features that are notably lacking in dialogic CMC (2006:43).

Apart from nonstandard orthography and some use of spoken language features, Crystal describes several other features as characteristic of what he terms *Netspeak*, such as emoticons and abbreviations. Other researchers such as Dürscheid (2004) and Androutsopoulos (2006), while of course not denying that such features can be found in CMC, resist the homogenizing implications of a term like *Netspeak*, given the great variability of CMC. Squires (2010) argues that internet language has been enregistered as a variety through public discourse, illustrated in her article by the discourse of US print media, and that the perception of this variety is dominated by certain nonstandard features, namely acronyms, abbreviations, and respellings, although empirical studies have often shown these to be relatively rare. Furthermore, she finds that *Netspeak* has become linked to nonstandardness and youth in public perception.

As the limitations of the *Netspeak* approach to CMC language have been increasingly noted, research in the area has seen ‘a shift of focus from medium-related to user-related patterns of language use’, so that ‘[c]haracteristic features of “the language of CMC” are now understood as resources that particular (groups of) users might draw on in the construction of discourse styles in particular contexts’ (Androutsopoulos 2006:421). In a related vein, there have been more and more studies on CMC from a creolistic perspective that take a sociolinguistic approach. The pioneering work and a major focus in this strand of research have been on Jamaican Creole, as in the case of studies of orthography in CMC. There are monographs on English and Jamaican Creole in email communication (Hinrichs 2006) and on sociolinguistic styling in the Corpus of Cyber-Jamaican, compiled from the forum *jamaicans.com* (Moll 2015). Recently Nigerian Pidgin has also attracted some attention, with articles published on various aspects of sociolinguistic styling and metalinguistic discourse in a corpus composed of data from the *Nairaland* forum (Heyd 2013, 2015; Heyd & Mair 2014).

In a first study of the above-mentioned Corpus of Cyber-Jamaican, Mair (2011) took up a concern from the first wave of CMC studies, the relation to spoken language, and this turned out to be sociolinguistically illuminating as well. He observes a sometimes extreme overrepresentation of basilectal Creole features in comparison to informal spoken data from the Jamaican component of the International Corpus of English (ICE), whose contributors have a similar social profile to those in the Corpus of Cyber-Jamaican except for the fact that they are all Jamaicans resident in the country, whereas the forum has many contributors from the diaspora. Mair explains this as the result of ‘participants engag[ing] in anti-formal linguistic behaviour in a spirit of playfulness’ (2011:223); the term ‘anti-formal’, from Allsopp (1996:lvii), refers to a stylistic level in Caribbean English which is characterized by speakers or writers ‘[d]eliberately rejecting [f]ormalness’, using Creole forms to signal ‘an absence or a wilful closing of social distance’ (ibid.). Overall, the corpus contains, according to Mair (2011:226), the following three types of data:

- a. passages which read as if spontaneously produced spoken JC [Jamaican Creole] was transferred on the screen,
- b. passages ... in which a presumably unconscious element of stylization is evident, and
- c. passages which are consciously crafted with rhetorical skill.

Whether these findings are specific to Jamaican Creole on jamaicans.com or are representative of a more general tendency in the use of creoles and other language contact varieties on the internet is not clear at this point. The present study will contribute to further researching this issue. Specifically, against this background, we aim to address the following research questions:

1. To what degree are Singlish features present in the chit chat forum as compared to informal spoken language?
2. What evidence, if any, is there of playful or stylized use of Singlish?
3. Considered in both quantitative and qualitative terms, to what degree does the use of Singlish features in informal written language correspond to their use in informal spoken language?
4. To what extent are other spoken language features used in the chit chat forum as compared to informal spoken language?
5. What role do CMC features play in forum users' stylistic practices?
6. All the above aspects considered, how close to informal spoken language is the language of the chit chat forum?

Lastly, from our theoretical and methodological background as outlined in section 1, and the major research strands outlined here, the following guiding question emerges:

7. What indexical processes are served by Singlish features in the context of CMC and how do features of CMC language contribute to these processes?

3. Previous research on Singlish and spoken language features in Singaporean CMC

There is some research on language use in CMC in Singapore but it has been fairly limited so far and reveals only a few aspects of the use of Singlish or general features of spoken language. Gupta (2006a) draws attention to 'Singlish on the web' in a study that uses a keyword sampling method in a broad approach comparing the online presence of Singlish and Geordie. Ooi et al. (2007) have presented a study on Singaporean weblogs based on two corpora of 100,000 words each, one of teenage blogs and one of undergraduate

blogs. While some findings on Singlish and other spoken language features can be gleaned from their comparative approach to these corpora, the study does not specifically focus on these. Another study by the same team (Tan et al. 2004) is actually devoted to ‘signalling spokenness’. This study, which considers Singapore and three other South East Asian countries (Malaysia, Brunei, and the Philippines), deals with personal advertisements on the web. Features investigated are augmenters – in which category the authors include the Singlish discourse particle *lah* and several general English items such as *really*, *very* – and mitigators (e.g. *just*, *a bit*). Overall the authors support Crystal’s dictum of Netspeak as ‘written language which has been pulled some way in the direction of speech’ (2006:63). This is because the frequency of the investigated features in the web data tends to be in between the frequency in printed and in spoken data from the ICE-Singapore corpus used for comparison. A notable exception, however, is the particle *lah*, which is practically absent from the web data. As an explanation the authors argue that ‘the notion of a “borderless” cyberspace might discourage advertisers from employing items like *lah* that point towards the local or suggest an insular or parochial outlook’ and speculate further that advertisers might want to be open to responses from nonlocals (2004:163). In a study of Internet Relay Chat (from Singapore only), in contrast, the same authors (2003) found *lah* to be rather common, and they conclude that ‘[t]he strong presence of *lah* in the sub-corpus [IRC] suggests that localisation is to be valued in this part of CMC’. We cannot be certain about the overall use of Singlish in these data, though, since *lah* is the only feature analysed in that study. With regard to the functions of *lah*, the authors note some not so typical uses, which, they suggest, are indicative of ‘chatters indulging in more self-conscious playfulness involving *lah*’. Particles and other Singlish features are also seen in Ooi and Tan’s (2014) more recent qualitative examination of Singaporean Facebook data; these, as the authors show, combine with Standard English and Netspeak features, as well as code-switches to non-English varieties.

Brown and Teo (2014) have compared the use of four discourse particles (*ah*, *lah*, *leh* and *lor*) in data from Singaporean online forums and spoken data, while their main focus is on the use of nonstandard spellings and certain specific lexical features for a particular type of identity stylization in the forums they analyzed. Leimgruber (2016) uses CMC data to investigate the previously undocumented discourse particle *bah*. His concern, however, is not CMC as such; he rather focuses on the description of *bah* and the evolving nature of the class of discourse particles in Singlish. Other recent studies whose concerns are closer to that of the present one are those by Sand (2013) on weblogs and by Deuber & Sand (2013) on the same corpus of weblogs as well as forum data. The former study shows that selected Singlish features as well as a global feature, discourse marker and quotative *like*, are present in weblogs, but at relatively low frequencies. The latter study finds the forum data to be closer to spoken language than the weblogs, but again the study considers only a limited number of features, which, moreover, display a rather

divergent behaviour, and the authors therefore highlight the need to study more features. The present study takes up this challenge. Only such a broader perspective will provide a sufficient basis to address the more general issues that we pursue.

4. The sociolinguistic situation in Singapore and its theoretical modelling

Present-day Singapore is a city-state characterized by a high level of ethnic and linguistic diversity. Like other world cities of its size, it boasts a population hailing from a large number of places around the globe, which has brought with it a large number of languages. What sets Singapore apart is that this diversity has been present on the island even since before British colonization: when the East India Company landed in Singapore in 1819, it found a population consisting of some 1,000 indigenous Malays and around 30 Chinese (Turnbull 1996:5). The landing vessel itself was manned by British officials as well as Indian soldiers and sailors (Sandhu 1993:774). At this point already, then, was the initial mix of the ‘traditional’ Singaporean ethnic groups brought together: Malays, Indians, and Chinese, with a few Europeans who, however, were in a position to impose their language – English – upon the ‘new’ settlement.

In the course of the following 100 years, immigration from around the Malay Archipelago, from the Indian Subcontinent, and from southern Chinese provinces resulted in the ethnic distribution that has roughly prevailed to this day: a majority of Chinese (74% in the 2010 census, Wong 2011), a strong Malay minority (13%), an Indian minority (9%), and a multitude of other ethnic groups (3%). These ethnic categories are the official ones used by the government for census and policy purposes. They gloss over more multifaceted realities within these categories (see e.g. Leimgruber 2013a; Lewis et al. 2013), with the same holding true for language: there are four official languages, English, Mandarin, Malay, and Tamil, but other languages spoken include several varieties of Chinese (Hokkien, Teochew, Cantonese, Hainanese, etc.), Dravidian languages (Malayalam, Telugu, Kannada), and Indo-Aryan languages (Punjabi, Bengali, Gujarati, Hindi, etc.), as well as Malayo-Polynesian languages (Javanese, Madurese, Tagalog, etc.).

The various educational and language policies implemented by the government since independence in 1965 (see e.g. Alsagoff 2012) have led to considerable language shift away from traditional ‘dialects’ to two of the official languages: Mandarin among the Chinese community (in large part due to the Speak Mandarin Campaign, launched in 1979, see e.g. Bokhorst-Heng 1999; Wee 2006) and English among all communities including the Chinese, with the trend being most pronounced among the Indian minority (see Bolton & Ng 2014; Tan 2014).

English in Singapore has been described as existing in two forms (see e.g. Gupta 1994): Standard Singapore English and Colloquial Singapore

English, commonly called Singlish, which is characterized by heavy lexical admixture and specific grammatical features. Lexical admixture into Singlish has come especially from the two main contact languages that have been spoken on the island, Hokkien (a Min Nan variety of Chinese) and Malay, with loanwords from other contact languages such as Cantonese also present. They come from many domains and include *kiasu* ‘afraid of losing out’ (from Hokkien), *paiseh* ‘embarrassed’ (from Hokkien), *kancheong* ‘nervous’ (from Cantonese), *ta pau* ‘take-away’ (from Cantonese), *makan* ‘food, to eat’ (from Malay), and even *kopitiam* ‘café’ (from Malay *kopi* ‘coffee’ and Hokkien *tiam* ‘shop’). An even more common class of items from the substrate languages comprises the discourse particles of Singlish. These are monosyllabic elements, usually in clause-final position, that can indicate a wide range of pragmatic aspects. Grammatically, Singlish is characterized by the variable absence of a range of elements, including verbal and nominal inflections, copular and auxiliary *be*, and subject and object noun phrases, as well as by several characteristic usages and constructions, e.g. invariant question tags, reduplication of various classes of words, and *got* in possessive, existential or perfective function (Alsagoff & Ho 1998; Fong 2004; Wee 2004b; Wee & Ansaldo 2004).

In the context of Singapore’s prescriptive language policies, Singlish became a target of the Speak Good English Movement, launched in 2000 to promote the use of the standard variety of Singapore English. However, the strong anti-Singlish stance of the early years of the movement (see Bokhorst-Heng 2000; Rubdy 2005) seems to have been gradually giving way to at least a grudging tolerance (see also Cavallaro et al. 2014:394). For example, when asked about the desirability of a ‘unique Singapore idiom’ in a 2005 interview, the then chairman of the movement, Professor Koh Tai Ann, replied that such an idiom, including some ‘Singlishisms’ – e.g. discourse particles – is in fact ‘in a sense ... desirable’ as ‘it is what gives a language its unique flavour and its users their identity, willy nilly’, though she also cautioned against it becoming ‘too distinctive and unique’, thus ‘depart[ing] too greatly from the standard form’ (Speak Good English Movement 2005). In 2010, the movement’s then and current chairman, Mr Goh Eck Kheng, even declared that ‘the government and the Speak Good English Movement is not out to eradicate Singlish’ and that ‘Singlish has become part of our national identity’, though he made it clear that ‘it’s not the role of the Speak Good English Movement to champion Singlish’ (Speak Good English Movement 2010). This is a rather remarkable recognition given that the government had for a long time not admitted any place for Singlish at all, instead assigning identity functions exclusively to Singapore’s Asian official languages while promoting English in its standard form as an ethnically neutral lingua franca and the language of international business (see e.g. Rubdy 2001:342; Wee & Bokhorst-Heng 2005:164-167; Alsagoff 2010:341-342). It supports Schneider’s (2007) description of Singapore as having reached phase 4 in the Dynamic Model of the evolution of Postcolonial Englishes. This is when, in

the post-independence period, former colonial subjects construct an identity as members of a new nation, and local varieties come to represent this identity. Thus, in Singapore, according to Schneider (2007:160),

The country's unique, territory-based, and multicultural identity construction has paved the way for a general acceptance of the local way of speaking English as a symbolic expression of the pride of Singaporeans in their nation. It encodes both sides of the national identity: its world language character expresses the country's global outreach and striving after economic prosperity, and its distinctively local shape on some levels ties up with the country's location and traditions.

The coexistence of Standard (Singapore) English and Singlish has been described and modelled in different ways. Gupta (1994) takes a Fergusonian diglossic view, attributing H ('high') status to Standard English and L ('low') status to Singlish. The two are, in this view, distinct, and have distinct purposes: H is taught in the classroom, and used in politics, in the media, and virtually all print, while L is acquired natively, used for informal everyday interaction, and generally unwritten. A departure from 'true' Fergusonian diglossia is observed in the concession that the Singaporean diglossia is 'leaky' (Gupta 2006:22), i.e. that features of L may appear in otherwise H discourse and vice-versa; this latter aspect is emphasized by Gupta in a later paper (2014), where she moves beyond diglossia entirely. Another approach to the variation between Singlish and Standard English is given by Platt (1975), who adapts the creole continuum model as developed by DeCamp (1971) for Jamaica to the Singaporean case. Here the idea is that there are not two varieties, but a multitude of subvarieties, ranging from the acrolect at the 'top' (equivalent to Standard English) through a number of mesolects all the way to the basilect Singlish. Speakers have at their disposal a range of the continuum that comprises all subvarieties between Singlish and that lect on the continuum which corresponds to their level of education: the lower one's level of education, the narrower the range of lects. Speakers then select a lect from this range based on the level of formality required. Other models have been proposed to explain the interaction between Singlish and Standard English, most taking educational attainment as one of the measures (Pakir 1991; Poedjoesodarmo 1995). The advantages and disadvantages of the different models have been analysed in some detail by Alsagoff (2007; 2010) and Leimgruber (2012, 2013b). To address the shortcomings of previous models, Alsagoff (2007, 2010) proposes a Cultural Orientation Model, which 'operates on the notion of language as a cultural resource, realized as a range of styles, where choice over culturally-loaded features may be used to indicate associative macrocultural orientations' (2010:344). In this framework features of Singlish (renamed 'Local Singapore English') are used for the discursive construction of a local orientation, whereas features of Standard English (or

‘International Singapore English’) are used for orientations that are more global. Alsagoff’s model is related to the framework of indexicality applied to Singapore by Leimgruber (2012, 2013b). Like the Cultural Orientation Model, indexicality rejects the diglossic premise that a segment of Singapore English can be readily identified as either Singlish or Standard on the basis of its constellation of linguistic features. It also considers linguistic features as resources in that their use is taken to index a particular social meaning. What indexicality offers beyond the Cultural Orientation Model is a more comprehensive theoretical framework, in that it recognizes indexical processes beyond cultural orientation alone. Furthermore, it can be extended to include instances of code-switching to non-English language material, with the switch indexing a particular social meaning (Leimgruber 2013b:59-60).

What distinguishes Alsagoff’s and Leimgruber’s approaches from a recent stylistic approach to Caribbean Creole continua (Deuber 2014), which works on the basis of a scale from Standard to Creole, is that features or resources are essentially seen as falling into two categories. This is not to say that they are all necessarily used in the same way. For example, Leimgruber (2013b:76) has shown that the Singlish variants of some variables are more favoured than those of others. Alsagoff also seems to suggest differences when she writes that speakers can vary not only the degree but also the type of Singlish and Standard English features (2010:345). However, a cline of features with different stylistic connotations that can be understood in terms of an adapted form of the continuum model, as proposed for Caribbean contexts by Deuber (2014), has not been described for Singapore.

5. Data

The internet chat data for the present paper was downloaded from the chat forum on sgforums.com in 2010 and amounts to about 100,000 words (this data set will subsequently be referred to as sgforums chat). So far as can be gleaned from the content of the forum posts, many (though not all) of the contributors are young people, including teenagers, and they are mostly resident in Singapore, though there are also a few who indicate that they live outside the country.

For comparison we use two sets of informal spoken data. The first is the ‘private dialogue’ section of the ICE-Singapore corpus (the files in question are labelled S1A-001 to S1A-100 and the data set will be referred to here as ICE-SIN S1A). With about 217,000 words it has a considerable size for a corpus of private conversations,¹ but for the purpose of comparison with

¹ The word count was retrieved by WordSmith WordList (‘tokens (running words) in text’). This word count may differ slightly from the word count provided, for instance, by Microsoft Word, but word counts for our other two data sets were retrieved in the same way, thus ensuring consistency. The exact word counts are: 216,913 for ICE-SIN S1A, 102,259 for

the sgforums chit chat data it nevertheless has two drawbacks: it dates from the 1990s, and the ICE corpora are generally restricted to speakers and writers aged 18 and over (Nelson 1996:28), so there is a time gap as well as a certain age gap. In contrast, our second data set, which comprises approximately 35,000 words, is of a later date (2006-7) and the speakers are younger: they were all at the time students at one of three postsecondary institutions where the data was collected (a junior college, a polytechnic, and a vocational training institute), and of an average age of 17.5 years. The data set comprises recordings of groups of informants carrying out a given task, namely to plan a holiday trip, as well as radio-microphone recordings of informal conversations without a set topic (our label for these recordings will be group/radio-mic).² Given the relative homogeneity of this data set in terms of the social profile of the speakers as well as its small size, we decided that it was ideal to be used as a complement to, but not in the stead of, ICE-SIN S1A.

In a few cases we furthermore draw on the Corpus of Global Web-Based English (GloWbE; <http://corpus.byu.edu/glowbe/>) to find out to what extent observed features are locally specific.

6. Analysis of text extracts

The two extracts below illustrate the range and diversity of language use on sgforums chit chat and will serve us to shed light on indexical processes in the context of the discourse environment of the forum.³

Extract 1

sgforums chit chat, and 34,675 for group/radio-mic. These word counts have been used as the basis for normalization to 100,000 words.

² More information about this data set, including a full explanation of the methodology used in the collection process and a breakdown of the informants' social and demographic background, can be found in Leimgruber (2009:36-47, 270-271).

³ For reasons of space the paragraphing of the original forum posts has not been preserved in the extracts printed here. Bold capital letters replace user names.

- A you are a pragmatist right? to you everything is replaceable, including one's roots and identity? to me, it isn't!!! i am already 45. i hope i won't live to see the day the chinese language goes into the museum. in fact i am quite sure i won't. there are millions and millions of chinese in china speaking dialects and mandarin. it would really take a farking big effort to make it go into obscurity. as for those mixed marriages you mentioned i don't have an answer and i don't care. only they know what race their children are. i am a chinese and i care only about my people.
- 5
- B** wah [A], relax la...dialects wun die off la....at least not in sg. You will be sure in China, it's very much alive. In Singapore, I give it another 100 years. Sometimes, it can't be helped. It's just that nowadays, a hokkien will marry a cantonese or whatever combinations. So lydat when the kids come out, they listen to the parents speak what? English or your mother's tongue (tongue in cheek!) of course.
- 10
- A** i am sure it's not difficult to learn both. i did!! my father was a hokkien and my mum a cantonese. i can speak both dialects well. i can speak english, chinese and japanese. i am sure the new generation is much smarter than people during my time. i have nothing against learning or speaking other languages. we must resist the tide of dropping dialects and mandarin because it make us who we are. [1 post by a further contributor omitted]
- 15
- B** well not all families share the same dynamics. why speak dialects when the dad and mum can speak english and chinese? You might resist the tide.....but I'm sure it will overwhelm you lol
- 20
- A** why do you keep telling me about foreigners? why do i care if foreigners (unless chinese) can speak chinese or not? every race should learn their own language. which part of this do you not understand? chinese learn mandarin. malay learn melayu. japanese learn nihongo. on top of that we learn english, the working language in singapore. it would be ideal if we can keep our dialects and mandarin as well. if it's too much drop the dialects but keep the language. english, for working reasons and chinese for your roots. is that not possible? i notice there are more and more people not even bothering to try. they just drop mandarin outright and focus on english. student nowadays can take so many subjects at one go but can't cope with 2 languages?
- 25
- 不是没能力 而是没那个心！！
- 30
- B** not true la....i took higher chinese back then but dropped it cos i couldn't cope with my other 8 subjects. Uncle at 45 years old, your education system is a bygone era. Even my time 15 years is now prehistoric. Children have a very very tough education system today.
- 35

The exchange in extract 1 is between two relatively older contributors to the chit chat forum. Contributor A is described as a 45-year-old (ll. 2, 36) and contributor B is probably somewhat above 30 (cf. l. 37). The discussion is on the topic of language issues in Singapore, about which A and B hold different opinions. Writer A defends a conservative view: in line with official policy, this writer's posts advocate a coexistence of English as a working language

and Chinese as a language of identity; furthermore A would ideally like to see the ‘dialects’, not encouraged by government policy, maintained in addition to Mandarin. The use of language is in line with the conservative views propounded in the posts. There are no abbreviations or respellings, though there is a conspicuous lack of capitals and apostrophes, features which, as Squires (2010:482) notes, are commonly used but relatively rarely discussed as part of internet language. Grammatically the English used is standard except for one uninflected verb form (l. 20) and an uninflected plural noun (l. 32). The switch to Mandarin in Chinese script in l. 34 (*bùshì méi nénglì ér shì méi nà ge xīn* ‘it’s not a lack of ability, it’s a lack of willingness’) underlines both the argument in favour of bilingualism and the writer’s own language competence.⁴

The language use of writer B, who is described by A as a ‘pragmatist’ (l. 1) and who is of the opinion that the ongoing language shift in Singapore is simply inevitable, differs in minor but important ways from that of writer A. While tending to use sentence-initial capitals and apostrophes most of the time, B employs an abbreviation typical of CMC, *lol* (l. 23) as well as two nonstandard spellings, *wun* ‘won’t’ (l. 9) and *lydat* ‘like that’ (l. 12). Unlike *lol*, the latter do not seem to be general features: in GloWbE most instances of *wun* ‘won’t’ are from Singapore, with a few from Malaysia as well, and hardly any from elsewhere, and there is only a single instance of *lydat*, which is precisely from sgforums.com. Also noteworthy is the use of the particle *la* (l. 9, 35). From the more macro point of view of the debate between A and B, these features can be read as being in line with a view of language in Singapore that is more open towards change and innovation than A’s conservative views. In the micro context of interactional dynamics, the cluster of local features at the beginning of B’s post in l. 9 (two instances of *la*, *wun*) appears to index a conciliatory stance, certainly as a reaction to A’s rather impassionate comment (as evidenced, for instance, in the liberal use of exclamation marks in l. 2 and the respelled swearword modifier in l. 5), and presumably in an attempt to facilitate the addressee’s reception of the writer’s own argument, itself presented for the most part in reasonably standard language (apart from the respelling of ‘like that’ as *lydat*), which underlines the importance of the point being made. The same can be seen in line 35, where B begins with the contradictory expression *not true* hedged with *la*: the particle fulfils here the role of solidarity marker, perhaps combined with its emphatic meaning, introducing an argument that is, again, presented without any use of particles. The absence of capitalisation and of the apostrophe in *couldn’t* in l. 35 may be seen as establishing a link to A’s writing style, easing the transition from the opening to the last two sentences, which are replete with standard characteristics, even to the extent of being non-locally specific

⁴ One may also note the writer’s display of linguistic erudition in l. 27, where the endonymous language names for Malay and Japanese (*melayu* and *nihongo*, respectively) – though not for Mandarin – are used.

(with the notable exception of *Uncle*, a term of respectful address to older males in several Asian Englishes).

Extract 2

- A** any of you poly students do MSN with your lecturer? At my poly, we have. A fren say got lecturer jio her tru MSN and they share sex jokes & chat. Nowadays she say she fall in love with this lecturer. Siao! Guy is married somemore. Sometimes I think teacher should not get too close with student.
- 5 **The teachers at my poly too frenly with student until I think student aso dun noe how to separate fren and teacher. U guys ever see this kind of thing ornot ?**
- B** what poly are euu in
- A** SP
- 10 **C** wa your female friend so easy to hook ar? intro leh hahahaha
- D** Only on TV.
- B** it's actually labeling. Most lecturers I have don't do that.
- E** this is wat we call 打成一片
- F** Unprofessional. [1 post by a further contributor omitted]
- 15 **G** Nothing to stop lecturers and students getting together on the MSN. but to share sex jokes etc makes the lecturer very unprofessional..... 😞 [3 posts by 2 further contributors omitted]
- H** Have. I am from rp and the lecturers give out their msn and yahoo chat id to students to receive Qns and guidance during non-class time. I hear ger classmates tease each other abt chat with lecturer many times. I think its commom at this age for gers to get emotionally swayed la. Get crushes over man teachers who look young or sporty.
- 20 **I** come let tok more, very soon it will appeared in lian he wan bao.⁵ Lol
- J** me from sp too. Its true some lecturers share msn with students. I got 2
- 25 lecturers & instructors who shared msn id with their class. But what they chat about only the 2 of them know la. Whether girl students fall in love with lecturers, all of us know it's possible also. Girls these days wear short short skirts, low low blouse, zao geng also maybe on purpose la. hahaha.. Lecturer or not, guy is guy ma. Ger want to show, guy sure see la. This is a free world.
- 30 Similarly, if ger want to make friends, go outside go hotel, as long as no one knows, no one can stop them. Last year, I ever see lecturer very friendly and act very close to ger student in school before. But that was quite late when have night class la.

⁵ *Lian he wan bao* is a Singaporean Chinese-language newspaper.

The forum contributor who initiates the discussion in extract 2 self-identifies as a student at a ‘poly’, i.e. polytechnic – one of the types of postsecondary institution that was also a site of fieldwork for our group/radio-mic data set – and explicitly addresses other polytechnic students. This discussion shows more use than extract 1 of nonstandard spellings. These include *fren* ‘friend’ (ll. 2, 6), *aso dun noe* ‘also don’t know’ (ll. 5-6), *eeu* ‘you’ (l. 8), *ger* ‘girl’ (ll. 19, 21, 29, 30). Some of these are apparently of restricted use. Ooi et al. (2007) found two of them, *noe* and *eeu*, to be characteristic specifically of their teenage blog corpus. For *noe* ‘know’,⁶ by far the largest number of examples in GloWbE is from Singapore (followed by Malaysia). The same applies to another of the nonstandard spellings seen here, *dun* ‘don’t’.⁷ *Eeu* ‘you’ is apparently even more restricted as there is no single example in GloWbE. Ooi et al. (2007) observe that ‘double letter additions’ as in *eeu* ‘are perceived as “cute” markers especially among teen female bloggers, who tend to signify their femininity by employing such visual cues’. Apart from these partly specific nonstandard spellings, one also finds an instance of *lol* (l. 23) and an emoticon (l. 16). Singlish features are far more numerous in this extract than in extract 1.⁸ They include:

- loanwords from Hokkien: *jio* ‘invite’ (l. 2), *siao* ‘crazy’ (l. 3), *zao geng* ‘expose’ (l. 28)
- discourse particles: *ar* (l. 10), *leh* (l. 10), *la* (ll. 21, 26, 28, 29, 33), *ma* (l. 29)
- zero copula, e.g. *teachers at my poly too frenly with student* (l. 5)
- uninflected verb forms: *she say she fall in love* (l. 3), *if ger want to make friends* (l. 30)
- *got* in possessive or existential function: *I got 2 lecturers* (ll. 24-25), *got lecturer* (l. 2)
- *ever* as an experiential aspect marker: *I ever see lecturer* (l. 31)
- uninflected plural nouns, e.g. *student* (l. 5)
- questions without *do*-support (ll. 1, 6)
- invariant question tag *or not* (l. 6)
- reduplication: *short short* (l. 27), *low low* (l. 28)
- zero subject, e.g. *have night class* (l. 33)
- absence of conditional conjunction: *Ger want to show* (l. 29)
- indefinite article omission, e.g. *lecturer* (l. 2)

⁶ The search string ‘noe.[v*]’ was used to identify relevant instances.

⁷ In this case the search string ‘dun [v*]’ was used.

⁸ As in extract 1, there is also one switch to Mandarin in Chinese script, in l. 13; it transliterates as *dǎchéngyīpiàn*, meaning ‘to become one’. This is a so-called *Chéngyǔ*, a four-character idiomatic expression usually translated as ‘proverb’ in English. Being semantically opaque, they have to be learnt by rote; their usage may potentially indicate a certain level of education. In this case, and given the topic at hand, this ‘proverb’ is clearly used with ironic intent.

The age of the contributors and the rather lighter topic compared to the one in extract 1 are factors likely to have influenced the language used, but beyond these it is important to pay attention to variation on the micro level. The features described above are not evenly distributed throughout the extract. The discussion starts off with an opening post that contains several features of Singlish, but this elicits quite different responses. For example, B's factual question in l. 8 contains a nonstandard spelling but no Singlish feature, whereas C's joking question in l. 10 has zero copula and two discourse particles. B (l. 12) and G (ll. 15-16) again use Standard English for rather serious comments. While G expresses a critical stance towards lecturers exchanging 'sex jokes etc.' with students, with the disapproval underlined by the emoticon, H, in a post that is overall in a similar vein (ll. 18-22), shifts the attention to girls attracted by lecturers; the lone Singlish feature in this post, the particle *la* in l. 21, could, in this context, index a certain feeling of understanding and solidarity. The last post, the longest in this extract (ll. 24-33), also expresses understanding and even tolerance and appeals to shared experience (*all of us know*, l. 27), hovering in tone between serious and playful. While being rich in Singlish features that support the stance of understanding and solidarity as well as the playful element, the post also contains many features that mark Standard English in contrast to Singlish such as use of copular *be*, the plural suffix *-s*, and verbs inflected for past tense and third person singular. By thus displaying quite a substantial knowledge of both Singlish and Standard English, J is able to create complementary stances of humour and maturity, insider status and general knowledgeability.

Leimgruber (2012, 2013b) has shown how, in the informal spoken language use of young Singaporeans, the juggling of various stances taken in the same conversation, the same turn, even the same utterance within such a turn, is achieved by a careful mix of language resources displayed in order to index those particular stances. Here it has become clear that the indexical approach to variation in Singapore is suitable to account for heterogeneity not only in the spoken but also in the written medium. The stances apparently indexed by the Singlish and Standard English features in the data fit in well with the indexical field of Singapore English presented by Leimgruber (2013b:106). This incorporates the aspects associated with Alsagoff's local versus global orientations (derived from Table 1 in Alsagoff 2007:39), which include closeness, camaraderie, informality versus authority, formality, distance, and adds further stances, e.g. relaxed versus serious. In keeping with the fluid nature of an indexical field other related stances and social identities could be added here, such as innovative, playful, young versus conservative and mature. Furthermore, as in speech, non-English language material can introduce further nuances to the discourse. Beyond this, we have seen that Netspeak features contribute in important ways to the construction of social meaning in the interactions presented, sometimes underlining and sometimes complementing the indexical work done by Singlish features. The indexical approach can easily accommodate this additional layer of variation. In fact, the

study of style and social meaning encourages us to look at a wider range of variables than has traditionally been the case in sociolinguistic studies (Eckert 2008:472).

7. Quantitative findings

7.1. Features associated with Singlish

Our analysis of features associated with Singlish will focus on the discourse particles (section 7.1.1) because these are highly distinctive, even stereotypical. This applies especially to *lah*, which has been described as ‘the most stereotypical particle’ (Leimgruber 2013b:86) and even ‘the hallmark of Singapore English’ (Besmeres & Wierzbicka 2003:3). The following example from an internet forum titled ‘Expatsingapore’ (www.expatsingapore.com/forum/) illustrates how a writer draws on this status of *lah*:

- (1) SG is NOT Beijing or Shanghai or Fujian or Canton, or UK or USA we're uniquely Singapore lah!!

In addition to the particles we have selected a grammatical construction that involves borrowing from a substrate language, namely the *kena*-passive, which employs *kena*, from Malay, as passive marker (section 7.1.2). ‘The Malay origin marks the *kena* passive as basilectal, and it is hardly used in formal writings in Singapore’, as Bao (2010:803) observes.

7.1.1. Discourse particles⁹

There is an extensive literature devoted to the Singlish discourse particles, in which the range of particles and their functions have been described in somewhat different ways. We begin with what may be considered a core set of particles, consisting of a total of nine particles mentioned in all or most studies on the subject since the late 1980s (see Leimgruber 2013b:92-94 for a detailed overview). As Lim (2007:465) reports, earlier studies mention only *lah*, *ah*, and *what* and it appears that only these three were already present in the 1970s. Lim (2007:465-466) identifies the source of *lah* and *ah* as Hokkien and/or Bazaar Malay (with reinforcement by Cantonese in the case of *ah*); Hokkien was the most widely spoken Chinese language in Singapore until the 1970s and the most important lingua franca in this period, followed by Bazaar Malay, a contact variety of Malay (Lim 2007:453-454). The origin of *what* is not entirely clear. Resembling Cantonese *wo* in form, it is functionally closer

⁹ There is some terminological variation in this area (see the discussion in Leimgruber 2013b:82-84) but we follow common practice in referring to the Singlish items discussed in the present section as discourse particles, and to the English items discussed in section 7.2.1 as discourse markers. Cf. also Wee (2003), who makes the same distinction.

to Hokkien *ma*, which, combined with its appearance in the same early period as *lah* and *ah*, leads Lim (2007:466) to favour explanations to the effect that it is a calque from Hokkien on English *what* (see Lim 2007:464). However, Cantonese influence on the form of the particle is considered possible by Gupta (1992b:42) and Leimgruber (2013b:91). The remaining particles, apparently introduced later than the other three, are, according to Lim (2007), all of Cantonese origin, with their appearance related to the increasing popularity of Cantonese in the period from the mid-1980s in connection with the rise of Cantonese pop culture.

Table 1 gives an overview of the core set of particles, based on Leimgruber (2013b:86-89), Lim (2007), Wee (2004a), and Gupta (1992b). Leimgruber (2013b:86-89) describes exactly this set of nine particles, while Wee's (2004a) and Lim's (2007) studies each cover eight of them, the former not including *ah* and the latter not including *hah*. Gupta (1992b) describes eleven particles, the nine listed in Table 1 and two more, namely *ge* and *na*. *Ge* does not occur in any of our three data sets and the status of *na* is questionable, as will be explained below, therefore these two are not included here.

There are no fixed spellings for the particles but in the recent linguistic literature one can observe a high degree of convergence. The spellings employed in Table 1 have been adopted from the most recent study (Leimgruber 2013b:86-89) but Lim (2007) differs only with respect to *mah* (<ma>), and Wee (2004a) with respect to *mah* (<ma>) and *what* (<wat>). Gupta used different spellings from those in Table 1 except for *what* (<ma, mei, lei, la, lo, ho, ha, a>) in her 1992b paper but has more recently adopted the same spellings as in Table 1, describing these as 'the conventional orthography as found in printed works which use Singlish' (2006b:250).

As regards the functions of the particles, the table incorporates, first, Gupta's division into three main groups:

The contradictory particles are used in utterances which explicitly contradict something a prior speaker has said. The assertive group are used to express speakers' positive commitment to what they are saying. The tentative group offer a less positive commitment. (Gupta 1992b:37)

For further characterization of the individual particles Wee's (2004a:125-126) succinct descriptions are quoted in the table. Lim (2007:260-261) provides an overview with mostly similar descriptions from which the one for *ah* is taken.

Particles	Functions	
	groups (Gupta 1992b)	individual particles (quoted from Wee 2004a, for <i>ah</i> from Lim 2007)
<i>mah</i>	contradictory	indicates information as obvious
<i>what</i>		indicates information as obvious and

		contradictory
<i>meh</i>	assertive	indicates scepticism
<i>leh</i>		marks a tentative suggestion or request
<i>lah</i>		indicates speaker's mood/attitude and appeals to addressee to accommodate the mood/attitude
<i>lor</i>		indicates obviousness or a sense of resignation
<i>hor</i>	tentative	asserts and elicits support for a proposition
<i>hah</i>		question marker
<i>ah</i>		<i>[ah24]</i> signals continuation (in narratives or explanations) and keeps interlocutors in contact; softens command; marks a question expecting agreement <i>[a21]</i> marks a question requiring a response

Table 1
Core set of Singlish discourse particles and their functions

It is important to note that although these descriptions helpfully summarize essential aspects of the functions of the particles, there are other studies that have offered different descriptions. This can be seen clearly in Leimgruber's (2013b:92-94) more comprehensive overview, which includes definitions of the particles by various authors.¹⁰ It is evident from this overview that an especially wide range of functions has been ascribed to *lah*, which ties in with Gupta's (1992b:42) observation that *lah* in fact 'covers the full range within the assertive continuum'.¹¹ Another assertive particle with a considerable range of possible functions, some of them overlapping with those of *lah*, is *leh*. In addition to tentativeness, its range of functions, according to the various studies cited by Leimgruber (2013b:93), include marking of emphasis and conveying information assumed to be new, both of which functions have been reported for *lah* as well. We will return to the functions of *leh* after presenting the quantitative findings.

The quantitative analysis obviously had to exclude other items homographous with the particles (e.g. Chinese expressions or proper names such as *ah beng* [see Wong 2014:116-121], *Nicholas Loh*; *what* as pronoun) but a number of other less straightforward cases had to be considered and taken a decision on as well. As already mentioned, the discourse particles

¹⁰ Note that some authors distinguish between different tonal patterns of the particles.

¹¹ For examples illustrating the different uses of *lah*, including several from ICE-SIN S1A, see Ler (2006:155-156).

typically occur at the end of a clause (or a nonclausal unit, e.g. *no lah*). However, Wee (2004a:124-125) mentions that *hor* and *hah* can also constitute a complete conversational turn, functioning as response elicitors, and our data shows that *ah* can occur on its own as well; in these cases it is used as a backchannelling response form, as in (2).

- (2) **B** <#> Because I I I think the problem also partly because of this computer drawing thing you know
A <#> **Ah**
B <#> Because we had we can't get a feel of what are the the drawings in the computer
 (ICE-SIN S1A-015)

Furthermore, *hah* and *ah* can occur at the beginning of a turn or clause:

- (3) **B** <#> But not long she told me before she has a screw inside her arm
A <#> **Ah** <#> You mean it hasn't been removed uh
B <#> No so when she go and take X-ray
A <#> Then see the screw
 (ICE-SIN S1A-077)
- (4) **Hah** not bad hor (ICE-SIN S1A-092)
- (5) **Ah** yes, CAS Paper Crane! (sgforums chit chat)

In turn/clause-initial position or when they stand on their own, *hah* and *ah* are not clearly distinguishable from similar items in general English (see e.g. Biber et al. 1999:1082-1092), and we have therefore excluded these uses,¹² combined forms such as *ah hah*, *uh ah* and *uh hah* have been excluded for the same reason. In contrast, *hor*, distinctive for Singlish, has been included wherever it occurs. The greatest problem in the analysis, however, was that the particle *ah* and the hesitator *uh* were apparently not consistently distinguished in the transcription of the ICE-SIN S1A data. While it does appear that <ah> often represents the particle or a response form or interjection as described above and <uh> the hesitator, as indicated in the ICE markup manual for spoken texts (Nelson 2002:6), there is evidently some overlap. Consider, for instance, *uh* at the end of A's first turn in example (3) above: given that it occurs at the end of a question to which the next speaker gives a response (and in the absence of other signs of hesitation such as syntactic incompleteness and repetition), this can reasonably be interpreted as the particle rather than the hesitator. Conversely, there are examples of <ah> where the obvious interpretation is that it represents the hesitator:

¹² The same might be said of *hah* in clause-final position since there is also a question tag *huh* especially in American English (Biber et al. 1999:1089) but we have nevertheless retained *hah* in what is its most prototypical position as a Singlish particle. At any rate, it is quantitatively of minor importance in our data.

- (6) I dread to see their **ah** their their their suffering (ICE-SIN S1A-028)
- (7) Wasn't like like **ah** like that (ICE-SIN S1A-031)
- (8) Then the sea part is quite **ah** the beach is quite broad also ah (ICE-SIN S1A-001)
- (9) Ya you return in two week's time and **ah** very troublesome lah to go there (ICE-SIN S1A-059)

The particle *ah* is not restricted to occurring in clause-final position. However, clause-medially it would be expected to occur after a complete clause constituent, typically a subject or an adverbial (see Gupta 1992b). In the examples above, in contrast, it occurs within a noun phrase (6), within a prepositional phrase (7), after an incomplete adjective phrase (8), and after the coordinator *and* (9), respectively. In examples (6) and (7) the impression of hesitation is reinforced by repetition. In the group/radio-mic data set, where a distinction between <ah> in the functions of particle, response form or interjection and <er> as hesitator was consistently made in the transcription, <ah> is never found in contexts like those above. In view of this, we have excluded occurrences of <ah> in atypical syntactic contexts of the type illustrated in (6) to (9) above from the count for the particle. As for <uh>, we searched for sequences of <uh> and the text unit marker <#> to find occurrences of <uh> in final position as in (3) and then manually identified all instances like those in (3) where it more plausibly represented the particle than the hesitator. Of course, it is likely that there are also instances of nonfinal <uh> where the speaker actually uttered a particle rather than a hesitator but given that <uh> does very often occur in contexts where it is clearly a hesitator and use as a hesitator is difficult to exclude clause-medially, <uh> in nonfinal position has not been considered in the analysis of particles. Finally, a further problem in the analysis of *ah* in ICE-SIN S1A involved the forms <nah> and <nuh>. As mentioned earlier, a separate particle *na* has been reported by Gupta (1992b). She classifies it among the assertive particles and notes that it is especially associated with rebukes and also found in directives (1992b:43). There are no attestations of such a particle in the sgforums chit chat data nor the group/radio-mic data set.¹³ What is striking in the ICE-SIN S1A data is that in the majority of instances <nah> and <nuh> occur after a word ending in <n> or a consonant cluster that would likely be reduced to /n/ in pronunciation. Most of these instances are in contexts where the particle *ah* could occur, as in (10), though specifically <nuh> also occurs in contexts where the hesitator *uh* could plausibly be found, as in (11).

¹³ In the sgforums chit chat data there are 3 instances of <nah> but these have a different function. In two instances it is equivalent to *no* and in one case its meaning resembles 'here' (*Nah go read these by yourselves ...* [followed by two weblinks]).

(10) Oh you mean that one **nah** (ICE-SIN S1A-013)

(11) This is on **nuh** Thursday (ICE-SIN S1A-010)

Thus we seem to be dealing with transcription variants of the particle *ah* and all instances of <nah> and <nuh> in contexts like that in (10) have therefore been included in the count for that particle. Overall the exact number of instances of the particle *ah* in ICE-SIN S1A is impossible to determine in view of the inconsistencies in the transcription and without access to the sound files. We believe to have achieved the best possible approximation but nevertheless, the figure for *ah* in ICE-SIN S1A should be considered with a degree of caution.

Table 2 presents the frequency of each of the nine particles in Table 1 in each of the three data sets; the overall figures for each particle are also broken down according to orthographic variants.

	sgforums chit chat		ICE-SIN S1A		group/radio-mic	
	N	N/100,000 words	N	N/100,000 words	N ^a	N/100,000 words
mah	59	57.7	15***	6.9	7**	20.2
<mah>	27		14		7	
<ma>	32		1			
what	40	39.1	203***	93.6	32***	92.3
<what>	12		181		32	
<wah> ^b	1		22			
<wat>	6					
<wad>	3					
<wor>	17					
<worh>	1					
TOTAL CONTRA-DICTORY	99	96.8	218‡	100.5	39‡	112.5
meh	23	22.5	13***	6.0	3‡	8.7
<meh>	22		13		3	
<merh>	1					
leh	145	141.8	40***	18.4	8***	23.1
<leh>	96		38		8	
<le>	32		1			
<lei>	7		1			
<ley>	2					
<ler>	6					
<lerh>	1					
<reh>	1					
lah	219	213.2	1609***	741.8	164***	473.0
<lah>	78		1605		164	
<la>	118		2			
<lar>	21		2			
<rah>	1					
<luh>	1					
lor	131	128.1	144***	66.4	11***	31.7
<lor>	89		138		11	
<loh>	21		6			
<lo>	16					
<roh>	5					
TOTAL ASSERTIVE	518	505.6	1806***	833	186‡	536.4
hor	40	39.1	57‡	26.3	9‡	26.0
<hor>	40		57		9	

hah	13	12.7	47‡	21.7	2‡	5.8
<hah>			14		2	
<huh>	12		33			
<harh>	1					
ah	74	72.4	1372***	632.5	337***	971.9
<ah>	51		582		337	
<uh>			677			
<nah>			30			
<nuh>			83			
<ar>	14					
<arh>	9					
TOTAL TENTATIVE	127	124.2	1476***	680.5	346***	997.8
GRAND TOTAL	744	726.6	3500***	1613.6	625***	1802.5

Table 2

Frequency of discourse particles

^a These figures are from Table 5.8 in Leimgruber (2009:184), with subtraction of turn-initial instances in the case of *ah*.

^b <wah> occurs both as a spelling variant of *what* and as an interjection. In the latter case it occurs in turn/clause-initial position or on its own (see extract 2 in section 6 above, l. 8 for an example) and has not been included the count for the particle.

** The difference to the number in sgforums chit chat is statistically significant according to the chi-square test ($p < 0.01$).

*** The difference to the number in sgforums chit chat is statistically significant according to the chi-square test ($p < 0.001$).

‡ The difference to the number in sgforums chit chat is not statistically significant according to the chi-square test ($p > 0.05$).

The results for the two spoken data sets display a remarkable degree of similarity and resemble previous findings as well. Gupta (1992b:38) singles out *ah*, *lah*, and *what*, i.e. the three particles that seem to have appeared first (see the discussion above) as the most widespread; in her own data she finds *ah* and *lah* to be most common, with the former being more than three times as frequent than the latter (1992b:47). *Ah* and *lah* also stand out in terms of frequency in both of our spoken data sets, though there are rather notable differences in their individual frequencies. These are followed by *what*, which is of practically equal frequency in both data sets. The other particles have for the most part comparatively very low frequencies which are also fairly similar between the two data sets. The SG forums chit chat data contrasts with the spoken data in several respects while showing some parallels as well. The overall number of tokens of the nine discourse particles is far lower. The difference is due in considerable measure to one single particle, the tentative

particle *ah*, whose frequency is strikingly low in the written compared to the spoken data. The communicative situation probably plays an important role in this, as the function of *ah* that has been described as to signal continuation and keep interlocutors in contact (Lim 2007, see Table 1 above) is clearly more relevant to spoken interaction. No significant differences are seen for the other two tentative particles, *hor* and *hah*. With regard to the two contradictory particles, *mah* and *what*, there is no significant difference between either of the two spoken data sets and sgforums chit chat with regard to the total number for this group but there are significant differences for each of the two particles. Whereas *what* clearly dominates in the spoken data, the two show a more even distribution in the written data, with instances of *mah* actually outnumbering those of *what*. In the case of the assertive particles the total number in ICE-SIN S1A differs significantly from that in sgforums chit chat due to the extremely high number of instances of *lah*. There is no significant difference between group/radio-mic and sgforums chit chat in the total number of assertive particles. However, three of the four particles show a significant difference between both of the spoken data sets as compared to sgforums chit chat. *Lah* is less frequent in the written compared to the spoken data. *Lor*, in contrast, is more frequent in the written than in the spoken data. The same holds true for *leh* with the discrepancy being even greater. One may ask in view of this whether *leh* is used in a similar or different way in the written as compared to the spoken data. This is a difficult question as the exact function of a particle in a specific context is not always easy to pin down. However, it does appear that for the most part the instances of *leh* in the SG forum chit chat data are within the range of functions described in previous studies, and similar to those in the spoken data. Examples (12)-(19) are taken from the sgforums chit chat data as well as from ICE-SIN S1A and illustrate the use of *leh* in various functions. Tentativeness and emphasis as in examples (12)-(13) and (14)-(15), respectively, appear to be the most common functions. Use of the particle in contexts where new information is provided is shown in examples (16)-(17), while examples (18)-(19) illustrate another of the contexts for which the use of *leh* has been described previously, namely disagreement (Platt 1987:398; see also Leimgruber 2013b:93).

- (12) no debt for mi yet.... still thinking to buy house or not **leh**...
(sgforums chit chat)
- (13) Don't really intend to shop **leh** (ICE-SIN S1A-091)
- (14) Then you hang them on your wall for display.. UNIQUE
LEH!!!! (sgforums chit chat)
- (15) It's like oh another second half **leh** again so long (ICE-SIN S1A-
077)
- (16) err she's turning 15 tis yr liao **lei** (sgforums chit chat)
- (17) You know he was with the main branch **leh** for one year as a
cadet officer then subsequently he join Bedok branch as second
officer (ICE SIN S1A-064)

- (18) no **leh**.. some of those fling have was xiao mei mei too...
(sgforums chit chat)
- (19) Don't say always **leh** (ICE SIN S1A-041)

It appears thus that the expansion of the use of *leh* in the forum data is quantitative more than qualitative in nature. Why should the forum contributors show such a predilection for a particle that is relatively rare in speech? One relevant factor here seems to be linguistic creativity as a strategy of in-group identity marking. Consider in this connection the top 15 frequency ranked items in Ooi et al.'s (2007) corpus of teenage blogs as compared to the corpus of undergraduate blogs, according to Wmatrix: 9 are creatively respelled English words such as *euu* 'you' and *noe* 'know' (see Section 6 above) and 3 are common English words that are spelled regularly (*so*, *went*, *then*); there is also the abbreviation *LOL* and the remaining two are discourse particles, namely *leh* (spelled <le>) and *de* (on which see below). As teenagers are also among the forum contributors it is plausible that the informal written language use characteristic of this particular subgroup should have influenced the results. One reason why certain writers have apparently seized in particular on *leh* among the less common particles in speech may be that it is quite versatile in function, and can in some of its uses substitute for *lah*.

The ethnic composition of the samples might be surmised to be another possible reason why *leh* and several other particles are more common in the written than the spoken data, as there are some indications that particle use may differ to a certain extent by ethnic groups and the ethnic composition of these data sets could be different. We have no information on the ethnic background of the forum contributors, but in view of the fact that the Chinese ethnic group forms the majority of the population, writers of Chinese ethnicity are likely to be well represented (cf. also Deuber & Sand 2013:388 on the ethnic composition of the blog corpus they analysed). Thus it is not unlikely that more of the written data was produced by members of this ethnic group than is the case in the group/radio-mic recordings and at least part of ICE-SIN S1A, where speakers of Indian and Malay ethnicity are strongly represented. In the group/radio-mic recordings, the distribution of data by ethnic groups is as follows: Chinese 14,875 words, Indian 9,312 words, Malay 10,488 words. Social background information on ICE-SIN S1A is not available, but in the Grammar of Spoken Singapore English Corpus (GSSEC, see Lim & Foley 2004:10-14), a conversation corpus of over 60,000 words that has been integrated into ICE-SIN (Lim 2009), the majority of speakers are identified by ethnicity. In this corpus Smakman & Wagenaar (2013:313) counted 13,214 words by speakers of Chinese ethnicity, 34,874 words by speakers of Indian ethnicity, and 8,492 words by speakers of Malay ethnicity. As regards differences between the ethnic groups in particle use, Platt (1987:395) described *mah*, *hor* and *leh* as to some extent indicative of 'an ethnically Chinese basilectal subsection of the speech community'. Smakman & Wagenaar's (2013:317-318) results from the GSSEC seem to tie in with this

observation: they found that *hor* and *leh* were used exclusively by Chinese speakers (*mah* was not attested). They also found that *lor* was used by Chinese speakers more than by Malay and Indian speakers. The total numbers are very low, though: 9 and 6 instances respectively for *hor* and *leh*, and 32 for *lor*. The authors furthermore cite Leimgruber's (2009:185) findings to support their own. These may indeed be taken to point in the same direction but not only do the numbers appear too low to draw any definite conclusions, but the use of *hor* and *leh* is not even limited to speakers of Chinese ethnicity in this data set. Smakman & Wagenaar (2013:319) acknowledge this and conclude that 'these two particles may be moderate ethnic indicators rather than absolute ethnic indicators'. In conclusion they say that a 'relative detachment of particles from their roots' has taken place and suggest that '[h]eavy code-switching and everyday interethnic communication have probably added to speakers' linguistic versatility and a reduced ethnolinguistic awareness' (2013:320). Thus, even if the forum contributors were mostly of Chinese ethnicity, this would not necessarily have had a major effect on the use of particles. In fact, the differences between the spoken and written data that we have observed remain when only the Chinese speakers in Smakman & Wagenaar's analysis are compared to sgforums chit chat: as a percentage of the total of particles used, *ah* has a share of 46% in this group and *lah* of 38%, while the figures for *lor*, *hor*, and *leh* are 7%, 3% and 2%, respectively (Smakman & Wagenaar 2013:316). In sgforums chit chat, based on the results presented in Table 2, the percentages are as follows: *ah* 10%, *lah* 29%, *lor* 18%, *hor* 5%, and *leh* 20%. Therefore, it seems safe to assume that ethnicity is at most a minor contributory factor in our results for discourse particles.

As already indicated in the discussion so far, the Singlish particles are not limited to the core set that has been dealt with here so far. In fact, Wee (2010:46) argues that 'the category of particles in CSE [Colloquial Singapore English] is active and attracts "new recruits"'. Two such 'new recruits' described by Wee (2003, 2010) are *know* and *ya*, both from English. Another particle that is derived from English (in form) is *one*. This particle was in fact already mentioned by Gupta (1992a:336, 1992b:36) but only recently has it received detailed attention (Bao 2009; Wong 2014:180-191). Gupta (1992b:36) states that *one* is 'an assertive particle with a similar sense to *la*'. Bao (2009) and Wong (2014) both describe emphasis as the function of the particle *one*. There is agreement among the studies cited that this use of *one* is calqued on Chinese. Specific forms mentioned as the source are Mandarin *de* (Gupta 1992a:328; Bao 2009:341; Wong 2014:181) as well as Hokkien *e* or *ge* and Cantonese *ge* (Gupta 1992a:328). As Bao (2009) shows in detail, Singapore English *one* and *de* have parallel pronominal as well as emphatic functions. Wong (2014:181-190) has provided extensive examples of *one* as an emphatic particle. We limit ourselves to a few from both spoken and written data here:

(20) you see we take very long time **one** (group)

- (21) you guys are personally and manually going to build that that childcare yeah **one** (radio-mic)
- (22) she got alot of guy chasing her **one**... alot alot.. (sgforums chit chat)
- (23) if they ask ppl to go on a 1 way trip to stay on moon or mars i sure go **one**... (sgforums chit chat)

Bao (2009:351) reports 74 tokens of *one* used as an emphatic particle in ICE SIN-S1A, 34.1 per 100,000 words. In the group/radio-mic data we have found 11 instances, 31.7 per 100,000 words. Thus, as in the case of several other particles, the numbers for the two spoken data sets are very similar. Like most of the particles that are relatively uncommon in the spoken data, emphatic *one* is more frequently used in sgforums chit chat (67 occurrences, 65.5 per 100,000 words). The Chinese emphatic particle itself also occurs in this data set, in the Mandarin form *de*. To be sure, *de* (in various functions) is attested in the spoken data as well, but only in very small numbers and in contexts of wholesale code-switching, as in (24), or at least multilingual discourse where the preceding element is of non-English origin, as in (25) (*siao* is Hokkien for ‘crazy’).

- (24) zhen **de** zhen **de** <> in Mandarin) (ICE-SIN S1A-020)¹⁴
- (25) what he siao **de** go there (group)

In sgforums chit chat there is also a small number of such examples, but in this data set *de* also combines with English words, as in the examples below.¹⁵

- (26) Very special **de**!! Must have!!! (sgforums chit chat)
- (27) someone who got study wont get 39pts **de** la.. (sgforums chit chat)
- (28) my dai ti veri funny **de** (sgforums chit chat)
- (29) he mentioned he want ext hdd quite long ago, but i think give techie things to him weird weird **de**.....? (sgforums chit chat)
- (30) how come this ger keep having problems **de** (sgforums chit chat)

As Singapore is becoming an increasingly English-dominant society but with Mandarin as another language of major importance, it is plausible that a new particle from Mandarin should have entered Singlish (cf. also Leimgruber 2016 on *bah*),¹⁶ all the more so as Bao (2015:21) reports recent lexical

¹⁴ The editorial comment that the whole utterance is in Mandarin is present in the original corpus text; *zhen* is an adjective in Mandarin meaning ‘real’ and the whole utterance can be translated as ‘really, really’.

¹⁵ In a few instances it is not clear whether *de* is indeed the word from Mandarin in question but there are at least 30 examples of this type.

¹⁶ Examples like those in (26)-(30) have to the best of our knowledge not been presented so far. Platt (1987:395) lists *de* among other particles but without any explanations or examples.

borrowing from Mandarin. According to our findings it may be a CMC-specific usage but in view of the small size of our more recent spoken corpus we cannot of course exclude the possibility that it may occur in speech as well.

6.1.2. *The kena-passive*

The *kena*-passive has been described in detail by Bao & Wee (1999). They summarize the syntax and semantics of the construction as follows:

- a. The lexical verb can be in the bare form or in the past participial form
- b. The agentive *by*-phrase is optional
- c. The subject must be adversely affected
- d. Stative verbs cannot be passivized with *kena*
(Bao & Wee 1999:3)

Bao (2010:802) states that according to his personal experience, ‘the *kena* passive is common in spontaneous speech’. As he also states (*ibid.*), the ICE-SIN S1A files contain 6 instances (2.8 per 100,000 words). Two of these, with the spelling <kena>, are given by Bao (*ibid.*) as examples. In the other four instances the variant <kana> is used; three of them occur in the same text and with the same verb, *sexual harassed* [sic], which in one instance is ellipitd:

- (31) C <#> I **kana** sexual harassed again you know
[...]
A <#> The other way round uh <#> You harass the men
B <#> She she just said the **kana** right
A <#> Ya
B <#> So she must be object of the harassment (ICE-SIN S1A-031)

An example from another spoken corpus cited by Bao (2010:802n10), *kena electric shock* ‘suffer electric shock’, shows that *kena* can also combine with an object noun phrase where the meaning of the construction corresponds to that of the *kena*-passive.

Our group/radio mic data contains 2 instances of the *kena*-passive (5.8 per 100,000 words), both with the verb *scold*. The sgforums chat data has by far the highest number of instances of *kena* of our data sets, namely 29 (28.4 per 100,000 words).¹⁷ Of these, 3 seem semantically or structurally exceptional. In (32) and another instance with the same object noun the adversative aspect is absent.¹⁸ In (33), where *kena* is combined with Hokkien *buay song* ‘not satisfied’, the subject (*ppl*) follows *kena*.

¹⁷ Spelling variants are <kena>, <kenna>, <kana> and <kanna>.

¹⁸ According to Bao and Wee (1999: 3), the *kena*-passive can be grammatical with positively connoted words when an adverse interpretation is possible in a specific context, but it is not evident that this should be the case here.

- (32) A arena more fun, can get some special awards
 B i try the groword one i **kena** awards. (sgforums chit chat)
- (33) also **kena** ppl buay song (sgforums chit chat)

The remaining 26 examples are in line with what has been found in spoken data so far. They are adversative semantically and *kena* is used either in the verbal construction (with the verb in the participle form as in (34), in the base form as in (35), or ellipted as in (36); 21 instances) or with a following object noun phrase (see e.g. (37); 5 instances).

- (34) last time in the news got this boy **kena** stabbed in real in LAN cafe (sgforums chit chat)
- (35) heard hubby said recently got a ger **kena** stab in tampines by a robber (sgforums chit chat)
- (36) I know it feels shitty when you fall prey. i oso **kena** b4. (sgforums chit chat)
- (37) dont hide [the cane] if not **kena** hanger then worse le (sgforums chit chat)

With 26 examples of the type also found in spoken data (25.4 per 100,000 words) the *kena*-passive is significantly more frequent in sgforums chit chat ($p < 0.001$ as compared to ICE-SIN S1A and $p < 0.05$ as compared to group/radio-mic, according to chi square tests). However, the topics probably play an important role in this. There are three threads on topics which lend themselves particularly to the use of this construction, namely ‘Parents allowed to cane children?’, ‘Should females that commit serious crimes be caned as well?’, and ‘Crimewatch’, and half of the 26 examples occur in these three threads. The *get*-passive, which typically has an adversative meaning as well, is also more frequent in sgforums chit chat than in our spoken data. There are 60 instances (58.7 per 100,000 words), and of these, 21 are in the same three threads mentioned. In the group/radio-mic data, there are only 7 instances of the *get*-passive (20.2 per 100,000 words) and for ICE-SIN S1A Bao (2010:802) reports 56 (25.1 per 100,000 words).

In view of the specific semantics of the *kena*-passive, being that it is especially relevant to certain topics, we cannot necessarily conclude from the numbers that it is favoured more in the type of CMC we have analysed than in speech, but what can be said is that given appropriate contexts it appears to be used no less in informal writing than in speech.

7.2. General features of spoken English

This part of the analysis will look at discourse markers (section 7.2.1) and contractions (section 7.2.2). These were chosen because they are highly characteristic of spoken English and specifically informal conversation. While

this has so far been demonstrated mainly for traditional native varieties of English, our spoken data – compared to data from other ICE components, so far as available – will show to what extent these aspects are characteristic of informal spoken English in Singapore as well. The category of discourse markers of course overlaps with that of comment clauses, which Crystal (2006:43; see section 2 above) has described as lacking in CMC.

7.2.1. Discourse markers

The present analysis covers four discourse markers that are highly typical of spoken and particularly informal, conversational English in traditional native varieties, namely *well*, *you know*, *so*, and *like* (see e.g. Biber et al. 1999:887, 1096; Müller 2005:7; Carter & McCarthy 2006:2008; Miller 2009:321-323; Aijmer 2013:26-27), the latter being especially common in teenage speech (see e.g. Andersen 2001:224-226; Tagliamonte 2005:1900-1904). The same four discourse markers have been dealt with in detail by Müller (2005) and we have followed her detailed descriptions of the discourse marker functions of the four items in question in identifying the relevant instances in the data, thus see Müller (2005) for the types of uses included and examples.

Comparing several New Englishes, including Singapore English, as represented in ICE (direct conversations) with British English, Gilquin (2015:111-112) found a set of five two-word discourse markers to be used significantly less in the New Englishes, but with major divergences between individual items, *I mean* and *sort of* being underused and *you know* overused. To be able to better assess the status of the discourse markers as characteristic features of informal spoken English in Singapore we provide figures for ICE-GB S1A as well; these were retrieved automatically, which is possible in ICE-GB as the only ICE corpus so far that is parsed.¹⁹ According to the results (Table 3), *well* is remarkably rare in informal spoken English in Singapore. This is especially the case in the group/radio-mic data set. It is interesting to note in this connection that Aijmer (2013:25), comparing ICE-GB to the earlier London-Lund Corpus and also considering the Bergen Corpus of London Teenage Language, suggests for British English that *well* may sound ‘middle-age and old-fashioned’. As also found by Gilquin (2015:114), *you know* is extremely frequent in ICE-SIN S1A, its frequency exceeding that of the most common Singlish particle, *lah*, whereas it is not that frequent in the group/radio-mic data. *So* is very frequent in both Singaporean data sets. In connection with the low frequency of *like* in the private conversations in ICE-GB S1A it should be noted that these show the lowest frequency of this discourse marker among private conversations from eight ICE corpora analysed by Schweinberger (2014) and that the number in ICE-SIN S1A compares well with the number in the private conversations in several other

¹⁹ The figures were obtained from ICE-GB Release 2 by entering the items in question in combination with the discourse marker node (<dismk>) in the retrieval software ICECUP 3.1 that comes with the corpus. Normalization and statistical significance testing are based on the word count of 205,627 given in ICECUP 3.1 for the private dialogue component.

ICE corpora (cf. ICE-JAM 253.1, ICE-NZ 230.8, ICE-PHI 234.1; frequencies per 100,000 words based on Schweinberger 2014:152).²⁰ That it is twice as frequent in the group/radio-mic data set than in ICE-SIN S1A may be due to both the age of the speakers and the later date of the latter corpus (cf. Schweinberger 2014:199, 202). In sum, the spoken data indicate specific preferences and dispreferences for discourse markers in Singaporean English and suggest sociolinguistic variation and potentially diachronic change. Clearly, however, at least *you know*, *so* and *like* are common in spoken English in Singapore, including or even especially in the language of teenagers. As regards sgforums chit chat as compared to the spoken Singaporean data, the result is clear: all of these three discourse markers are far less frequent.

	Singapore				GB			
	sgforums chit chat		ICE-SIN S1A		group/radio-mic		ICE-GB S1A	
	N	N/100,000 words	N	N/100,000 words	N	N/100,000 words	N	N/100,000 words
<i>well</i>	37	36.2	265***	122.2	11‡	31.7	1495***	727
<i>you know</i>	14	13.7	1898***	875	103***	297	799***	388.6
<i>so</i>	252	246.4	1912***	881.5	351***	1012.3	1057***	514
<i>like</i>	39	38.1	518***	238.8	166***	478.7	140***	68.1
TOTAL	342	334.4	4593***	2118.8	631***	1819.8	3491***	1697.7

Table 3
Frequency of discourse markers

*** ICE-SIN S1A and group/radio-mic: the difference to the number in sgforums chit chat is statistically significant according to the chi-square test ($p < 0.001$); ICE-GB S1A: the difference to the number in ICE-SIN S1A is statistically significant according to the chi-square test ($p < 0.001$).

‡ The difference to the number in sgforums chit chat is not statistically significant according to the chi-square test ($p > 0.05$).

7.2.2. Contractions

Biber et al.'s figures for British and American English (1999:1129-1131) show that contraction rates for both verb forms and negatives are very high in conversation, with pronominal subjects particularly favouring contracted forms. Mair (2009) has analysed contractions of the verb *be* with pronominal

²⁰ Schweinberger's figures are not exactly comparable to the present ones but certainly reasonably so (cf. his figure of 63.1 per 100,000 words for ICE-GB S1A; 2014: 152)). Schweinberger removed speakers who uttered less than 100 words from the data and excluded outliers, i.e. speakers who produced extremely high numbers of the discourse marker *like*, from the analysis (see pp. 137, 150). Furthermore, he did not count *like* preceding numeric expressions (see p. 142), which, following Müller (2005: 210-11), we included.

subjects – both subject-verb contractions and contractions of *be* and *not* – in the private dialogues of four ICE corpora, namely those for Great Britain, New Zealand, Jamaica and India; the respective contraction rates he found are 94.8%, 97.7%, 84.7%, and 40.9% (2009:53). For the sake of comparability we have replicated his analysis, using the search strings listed by Mair (2009:52) and limiting the analysis to those uncontracted forms that could in theory have been contracted (cf. Mair 2009:53n19).²¹ As can be seen in Table 4, we found fairly high contraction rates – most similar to those in Jamaican English among the varieties analysed by Mair – in the spoken corpora. In sgforums chit chat, in contrast, contractions are used significantly less (chi square tests for the uncontracted and contracted variants in sgforums chit chat versus ICE-SIN S1A as well as sgforums chit chat versus group/radio mic are significant at $p < 0.001$).

Table 4

	uncontracted	contracted	total	contraction rate in %
sgforums chit chat	679	640	1319	48.5
ICE-SIN S1A	933	3492	4425	78.9
group/radio-mic	106	510	616	82.8

Frequency of *be*-contractions

8. Conclusion

The present study has taken as a point of departure Mair's (2011) observation on the overrepresentation of basilectal Creole features on a Jamaican forum. As discussed in section 4, the category of basilect is not unproblematic in the sociolinguistic context of present-day Singapore. A continuum analogous to the one in Jamaica was postulated at an earlier stage, but it has not been established that current stylistic uses of Singlish can be described in such terms. Nevertheless, the features associated with Singlish that we have chosen have been described as 'stereotypical' or 'basilectal' and therefore seem good candidates for features that might be overrepresented if a similar tendency as in the Jamaican forum held on sgforums chit chat. We have not found evidence of such an overrepresentation, however. In fact, the results for features associated with Singlish suggest a fair degree of correspondence between informal written and informal spoken usage, except where a feature is strongly constrained by the communicative situation, as in the case of the tentative particle *ah*. Apart from this case, the quantitative analysis of particle use revealed some notable parallels between spoken and written language use

²¹ For sgforums chit chat non-standard spelling of course had to be taken into consideration. In addition to the standard forms results were also found, and included in the figures, for forms without an apostrophe, with or without a space instead (e.g. <Im>, <I m> 'I'm', <its> 'it's', <isnt> 'isn't') and for the non-standard spellings <u> 'you' and <r> 'are'.

especially on the level of groups of particles with similar functions. On the level of individual particles, however, the forum writers tend to eschew the common ones, in particular the stereotypical *lah*, in favour of alternatives not as common in speech, with *leh* especially favoured. We have even found an innovative use of *de* as a particle in English-language contexts not attested in the spoken data. The *kena*-passive did appear to be overrepresented at first sight but closer inspection revealed the topics to be an important factor in this. On a qualitative level we have been able to show that the use of linguistic resources traditionally associated with a variety ‘Singlish’ can be explained, in the case of written CMC, in a framework of indexicality with a range of meanings in interaction in the same way as in speech. In sum, although there are differences, the use of Singlish features in sgforums chit chat is closer to informal spoken language than the use of basilectal Jamaican features on jamaicans.com is. The reason for this surely is that both the contributors and their discussions tend to be for the most part firmly rooted in the local community. Thus the identificatory function that Singlish has for Singaporeans may often be relatively taken for granted in the present context; there is not necessarily a need to emphasize a Singaporean identity in the way diasporic Jamaicans may emphasize a Jamaican one on Jamaicans.com, and to engage in performative, stylized vernacular language use for this purpose. On sgforums chit chat participants’ identity may rather be contested on other grounds:

(38) You don't seem like a 14/15 year old to me. In fact you sound like a bloody pervert [...].

While the high degree of stylization found by Mair (2011) on jamaicans.com is not generally characteristic of the data we have analysed, there is an element of linguistic playfulness in the way discourse particles are used. That the contributors creatively modify the patterns of use and the repertoire of individual particles can be explained by their aiming to index not so much their identity as Singaporeans as their membership in a subgroup such as young social media users. However, the fact that subgroup identities are foregrounded is probably not the only reason why the linguistic creativity of the Singaporean forum writers takes the form it does. Rather, the fact that what is called Singlish is not a focused variety but a fairly flexible set of resources would seem to provide fertile ground for such creativity. In Jamaica, although there is a high degree of variation in speech, there is at the same time a strong consensus on the linguistic norms of the Creole basilect (Patrick 1999:271).

Although there are some parallels between the use of Singlish features in internet chit chat and informal spoken language we cannot say that the language of sgforums chit chat is ‘written speech’, or spoken Singlish transferred to the screen. Not only is it problematic to speak of Singlish as a variety in the first place, but in speech features of Singlish combine with general features of spoken English such as discourse markers and contractions,

which are present only to a much lesser degree in the written medium. In the forum data, in contrast, Singlish features work together with Netspeak features. These include respellings of which some appear to be local or at least regional as well, and some even group-specific. The two aspects converge in the spelling variation that we have documented for the discourse particles.

When compared to previous research on Jamaican Creole in CMC, our findings indicate that the use of contact varieties in informal contexts on the internet can take different forms depending not only on the social groups of users and their degree of mobility versus rootedness in a national community, but also on the nature of the spoken language transferred to the written medium itself, a relevant factor in our case being the degree of diffusion as opposed to focusing (cf. Le Page & Tabouret-Keller 1985).

Research so far has provided much evidence that Singapore English in its standard and colloquial manifestations has reached phase 4 in Schneider's (2007) Dynamic Model of the evolution of Postcolonial Englishes, i.e. the stage where local forms of English become accepted as means of expressing a newly evolved national identity. Looking at the language use of young Singaporeans in the context of an internet forum through the lens of indexicality theory has afforded us a glimpse of developments beyond that stage, namely of phase 5: the forum we have looked at is a context where national identity and the linguistic resources that can serve to express it are largely taken for granted, allowing room for users to creatively adapt these resources and combine them with other resources for their communicative and social purposes in this specific context. In doing so, they display the 'internal differentiation' central to phase 5 (Schneider 2007:53). At this stage citizens of a postcolonial nation increasingly define their identity in terms of smaller subgroups and these develop distinct language varieties. Schneider's general discussion of developments in phase 5 (Schneider 2007:52-55) seems to be grounded predominantly in a rather traditional sociolinguistic approach centred on broad social categories and varieties or variables correlated to these. We would like to suggest that the study of (potential) phase 5 developments can be enriched by a more fine-grained, indexicality-based perspective of the type we have adopted here. We also argue that in addition to spoken face-to-face interaction as the traditional domain of sociolinguistic study, social media needs to be given more attention in this connection as an increasingly important site of social life where linguistic variation and indexicality can profitably be studied.

References

- Aijmer, Karin. 2013. *Understanding pragmatic markers: A variational pragmatic approach*. Edinburgh: Edinburgh University Press.
- Alsagoff, Lubna. 2010. English in Singapore: Culture, capital and identity in linguistic variation. *World Englishes* 29. 336-348.

- Alsagoff, Lubna. 2012. The development of English in Singapore: Language policy and planning in nation building. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language use*, 137-154. Amsterdam: Benjamins.
- Alsagoff, Lubna & Ho Chee Lick. 1998. The grammar of Singapore English. In Joseph A. Foley et al. (eds.), *English in new cultural contexts: Reflections from Singapore*, 127-151. Singapore: Oxford University.
- Allsopp, Richard. 1996. *Dictionary of Caribbean English usage*. Oxford: Oxford University Press.
- Andersen, Gisle. 2001. *Pragmatic markers and sociolinguistic variation*. Amsterdam: Benjamins.
- Androutsopoulos, Jannis. 2006. Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics* 10. 419-438.
- Ansaldo, Umberto. 2004. The evolution of Singapore English: Finding the matrix. In Lim (ed.), 127-149.
- Bao, Zhiming. 2009. *One* in Singapore English. *Studies in Language* 33. 338-365.
- Bao, Zhiming. 2010. A usage-based approach to substratum transfer: The case of four unproductive features in Singapore English. *Language* 86. 792-820.
- Bao, Zhiming. 2015. *The making of vernacular Singapore English: System, transfer, and filter*. Cambridge: Cambridge University Press.
- Bao, Zhiming & Lionel Wee. 1999. The passive in Singapore English. *World Englishes* 18. 1-11.
- Baron, Naomi S. 2003. Language of the internet. In Ali Farghali (ed.), *The Stanford handbook for language engineers*, 59-127. Stanford: CSLI Publications.
- Besemeres, Mary & Anna Wierzbicka. 2003. Pragmatics and cognition: The meaning of the particle *lah* in Singapore English. *Pragmatics & Cognition* 11. 3-38.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.
- Bokhorst-Heng, Wendy. 1999. Singapore's Speak Mandarin Campaign: Language ideological debates in the imagining of the nation. In Jan Blommaert (ed.), *Language ideological debates*, 235-265. Berlin: Mouton de Gruyter.
- Bokhorst-Heng, Wendy. 2005. Debating Singlish. *Multilingua* 24. 185-209.
- Bolton, Kingsley & Bee Chin Ng. 2014. The dynamics of multilingualism in contemporary Singapore. *World Englishes* 33. 307-318.
- Brown, David West & Teo Shi Jie. 2014. Singapore English and styling the *Ah Beng*. *World Englishes* 33. 60-84.
- Carter, Ronald & Michael McCarthy. 2006. *Cambridge grammar of English: A comprehensive guide*. Cambridge: Cambridge University Press.

- Cavallaro, Francesco, Bee Chin Ng & Mark Fifer Seilhamer. 2014. Singapore Colloquial English: Issues of prestige and identity. *World Englishes* 33. 378-397.
- Collot, Milena & Nancy Belmore. 1996. Electronic language: a new variety of English. In Susan Herring (ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*, 147-170. Amsterdam: Benjamins.
- Crystal, David. 2006. *Language and the internet*. 2nd edn. Cambridge: Cambridge University Press.
- DeCamp, David. 1971. Towards a generative analysis of a post-creole speech continuum. In Dell Hymes (ed.), *Pidginisation and creolisation of languages*, 349-370. Cambridge: Cambridge University Press.
- Deuber, Dagmar. 2014. *English in the Caribbean: Variation, style and standards in Jamaica and Trinidad*. Cambridge: Cambridge University Press.
- Deuber, Dagmar & Lars Hinrichs. 2007. Dynamics of orthographic standardization in Jamaican Creole and Nigerian Pidgin. *World Englishes* 26. 22-47.
- Deuber, Dagmar & Andrea Sand. 2013. Computer-mediated communication in Singapore: Spoken language features in weblogs and a discussion forum. In Röder & Wischer (eds.), 385-396.
- Dürscheid, Christa. 2004. Netzsprache – ein neuer Mythos. *Osnabrücker Beiträge zur Sprachtheorie* 68. 141-157.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12. 453-476.
- Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87-100.
- Fischer, Kerstin (ed.). 2006. *Approaches to discourse particles*. Amsterdam: Elsevier.
- Fong, Vivien. 2004. The verbal cluster. In Lim (ed.), 75-104.
- Gilquin, Gaëtanelle. 2015. At the interface of contact linguistics and second language acquisition research: New Englishes and Learner Englishes compared. *English World-Wide* 36. 90-123.
- Gupta, Anthea Fraser. 1992a. Contact features of Singapore Colloquial English. In Kingsley Bolton & Helen Kwok (eds.), *Sociolinguistics today: International perspectives*, 323-345. London: Routledge.
- Gupta, Anthea Fraser. 1992b. The pragmatic particles of Singapore Colloquial English. *Journal of Pragmatics* 18. 31-57.
- Gupta, Anthea Fraser. 1994. *The Step-Tongue: Children's English in Singapore*. Clevedon: Multilingual Matters.
- Gupta, Anthea Fraser. 2006a. Singlish on the web. In Azirah Hashim & Norizah Hassan (eds.), *Varieties of English in South East Asia and beyond*, 19-37. Kuala Lumpur: University of Malaya Press.

- Gupta, Anthea Fraser. 2006b. Epistemic modalities and the discourse particles of Singapore. In Fischer (ed.), 243-263.
- Gupta, Anthea Fraser. 2014. One world, one English. In Katrin Schöpe, Renate Belentschikow, Angelika Bergien, Armin Burkhardy & Karen Ebeling (eds.), *Pragmantax II: Zum Aktuellen Stand der Linguistik und ihrer Teildisziplinen. Akten des 43. Linguistischen Kolloquiums in Magdeburg, 2008*, , 613-634. Frankfurt on Main: Lang.
- Heyd, Theresa. 2014. Doing race and ethnicity in a digital community: Lexical labels and narratives of belonging in a Nigerian web forum. *Discourse, Context and Media* 4-5. 38-47.
- Heyd, Theresa. 2015. The metacommunicative lexicon of Nigerian Pidgin. *World Englishes* 34. 669-687.
- Heyd, Theresa & Christian Mair. 2014. From vernacular to digital ethnolinguistic repertoire: The case of Nigerian Pidgin. In Véronique Lacoste, Jakob Leimgruber & Thiemo Breyer (eds.), *Indexing authenticity: Sociolinguistic perspectives*, 242-266. Berlin: Mouton de Gruyter.
- Hinrichs, Lars. 2006. *Codeswitching on the web: English and Jamaican Creole in e-mail communication*. Amsterdam: Benjamins.
- Hinrichs, Lars & Jessica White-Sustaíta. 2011. Global Englishes and the sociolinguistics of spelling: A study of Jamaican blog and email writing. *English World-Wide* 32. 46-73.
- Leimgruber, Jakob R.E. 2009. *Modelling variation in Singapore English*. Oxford: University of Oxford Ph.D. dissertation. <http://jakobleimgruber.ch/papers/dissertation.pdf> (12 December, 2014.)
- Leimgruber, Jakob R.E. 2012. Singapore English: An indexical approach. *World Englishes* 31. 1-14.
- Leimgruber, Jakob R.E. 2013a. The management of multilingualism in a city-state: Language policy in Singapore. In Peter Siemund, Ingrid Gogolin, Monika Edith Schulz & Julia Davydova (eds.), *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education*, 229-258. Amsterdam: Benjamins.
- Leimgruber, Jakob R.E. 2013b. *Singapore English: Structure, variation, and usage*. Cambridge: Cambridge University Press.
- Leimgruber, Jakob R.E. 2016. *Bah* in Singapore English. *World Englishes* 35: 78-97.
- Le Page, Robert B. & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Ler, Soon Lay Vivien. 2006. A relevance-theoretic approach to discourse particles in Singapore English. In Fischer (ed.), 149-166.
- Lewin, Beverly A. & Yonatan Donner. 2002. Communication in internet message boards: A quantitative analysis of usage in computer-mediated conversation. *English Today* 71. 29-37.

- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the world*. 17th edn. Dallas: SIL International. <http://www.ethnologue.com/> (12 December, 2014.)
- Lim, Lisa (ed.). 2004. *Singapore English: A grammatical description*. Amsterdam: Benjamins.
- Lim, Lisa. 2007. Mergers and acquisitions: On the ages and origins of Singapore English particles. *World Englishes* 26. 446-473.
- Lim, Lisa. 2009. The Grammar of Spoken Singapore English Corpus (GSSEC): Ground rules and conventions. <http://www.english.hku.hk/staff/lisalim.htm> (9 February, 2015.)
- Lim, Lisa & Umberto Ansaldo. 2016. *Language in Contact*. Cambridge: Cambridge University Press.
- Lim, Lisa & Joseph A. Foley. 2004. English in Singapore and Singapore English. In Lim (ed.), 1-18.
- Mair, Christian. 2009. Corpus linguistics meets sociolinguistics: Studying educated spoken usage in Jamaica on the basis of the International Corpus of English. In Lucia Siebers & Thomas Hoffmann (eds.), *World Englishes: Problems, properties and prospects. Selected papers from the 13th LAWE Conference*, 39-60. Amsterdam: Benjamins.
- Mair, Christian. 2011. Corpora and the New Englishes: Using the 'Corpus of Cyber-Jamaican' to explore research perspectives for the future. In Fanny Meunier, Sylvie De Cock, Gaëtanelle Gilquin & Magali Paquot (eds.), *A taste for corpora: In honour of Sylviane Granger*, 209-236. Amsterdam: Benjamins.
- Mair, Christian. 2013a. Corpus approaches to the vernacular web: Post-colonial diasporic forums in West Africa and the Caribbean. In Röder & Wischer (eds.), 397-406.
- Mair, Christian. 2013b. The world system of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide* 34. 253-278.
- Maynor, Natalie. 1994. The language of electronic mail: Written speech? In Greta D. Little & Michael Montgomery (eds.), *Centennial usage studies*, 48-54. Tuscaloosa: University of Alabama Press.
- Miller, Jim. 2009. *Like* and other discourse markers. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative studies in Australian and New Zealand English*, 317-337. Amsterdam: Benjamins.
- Moll, Andrea. 2012. Orthographic practices in diasporic Jamaican online communities: Between idiosyncratic usage and 'grassroots' conventionalization. In Anne Schröder, Ulrich Busse & Ralf Schneider (eds.), *Codification, canons, and curricula: Description and prescription in language and literature*, 283-302. Bielefeld: Aisthesis.
- Moll, Andrea. 2015. *Jamaican Creole goes web: Sociolinguistic styling and authenticity in a digital 'Yaad'*. Amsterdam: Benjamins.
- Müller, Simone. 2005. *Discourse markers in native and non-native English discourse*. Amsterdam: Benjamins.

- Nelson, Gerald. 1996. The design of the corpus. In Sidney Greenbaum (ed.), *Comparing English worldwide: The International Corpus of English*, 27-53. Oxford: Clarendon.
- Nelson, Gerald. 2002. Markup manual for spoken texts. <http://ice-corpora.net/ice/manuals.htm> (12 December, 2014.)
- Oenbring, Raymond. 2013. *Bey* or *bouy*: Orthographic patterns in Bahamian Creole English on the web. *English World-Wide* 34. 341-364.
- Ooi, Vincent B.Y. & Peter K.W. Tan. 2014. Facebook, linguistic identity and hybridity in Singapore. In Rani Rubdy & Lubna Alsagoff (eds.), *The global-local interface and hybridity*, 225-244. Bristol: Multilingual Matters.
- Ooi, Vincent B.Y., Peter K.W. Tan & Andy K.L. Chiang. 2007. Analyzing personal weblogs in Singapore English: The Wmatrix approach. In Päivi Pahta, Irma Taavitsainen, Terttu Nevalainen & Jukka Tyrkkö (eds.), *Towards multimedia in Corpus Studies*. Helsinki: Research Unit for Variation, Contacts and Change in English. http://www.helsinki.fi/varieng/series/volumes/02/ooi_et_al/ (12 December, 2014.)
- Pakir, Anne. 1991. The range and depth of English-knowing bilinguals in Singapore. *World Englishes* 10. 167-179.
- Patrick, Peter L. 1999. *Urban Jamaican Creole: Variation in the mesolect*. Amsterdam: Benjamins.
- Platt, John T. 1975. The Singapore English speech continuum and its basilect 'Singlish' as a 'creoloid'. *Anthropological Linguistics* 17. 363-374.
- Platt, John. 1987. Communicative functions of particles in Singapore English. In Ross Steele & Terry Gold (eds.), *Language topics: Essays in honour of Michael Halliday*, vol. 2, 391-401. Amsterdam: Benjamins.
- Poedjosoedarmo, Gloria R. 1995. Lectal variation in the media and the classroom: A preliminary analysis of attitudes. In Teng Su Ching & Ho Mian-Lian (eds.), *The English language in Singapore: Implications for teaching*, 53-67. Singapore: Singapore Association for Applied Linguistics.
- Rajah-Carrim, Aaliya. 2009. Use and standardisation of Mauritian Creole in electronically mediated communication. *Journal of Computer-Mediated Communication* 14. 484-508.
- Röder, Katrin & Ilse Wischer (eds.). 2013. *Anglistentag 2012 Potsdam: Proceedings*. Trier: Wissenschaftlicher Verlag.
- Rubdy, Rani. 2001. Creative destruction: Singapore's Speak Good English Movement. *World Englishes* 20. 341-355.
- Runkehl, Jens, Peter Schlobinski & Torsten Siever. 1998. *Sprache und Kommunikation im Internet: Überblick und Analysen*. Opladen: Westdeutscher Verlag.
- Sand, Andrea. 2013. Singapore weblogs: Between speech and writing. In Magnus Huber and Joybrato Mukherjee (eds.), *Corpus linguistics and variation in English: Focus on non-native Englishes*. Helsinki:

Research Unit for Variation, Contacts and Change in English.
<http://www.helsinki.fi/varieng/journal/volumes/13/sand/> (9 February, 2015.)

- Sandhu, Kernial Singh. 1993. Indian immigration and settlement in Singapore. In Kernial Singh Sandhu & A. Mani (eds.), *Indian Communities in Southeast Asia*, 775-788. Singapore: Times Academic Press.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Squires, Lauren. 2010. Enregistering internet language. *Language in Society* 39. 457-492.
- Schweinberger, Martin. 2014. The discourse marker LIKE: a corpus-based analysis of selected varieties of English. Hamburg: University of Hamburg Ph.D. dissertation. <http://ediss.sub.uni-hamburg.de/volltexte/2014/6969/pdf/Dissertation.pdf> (12 December, 2014.)
- Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23. 193-229.
- Smakman, Dick & Stephanie Wagenaar. 2013. Discourse particles in Colloquial Singapore English. *World Englishes* 32. 308-324.
- Speak Good English Movement. 2005. Professor Koh Tai Ann's interview with Straits Times Deputy Editor, Felix Soh. <http://www.goodenglish.org.sg/category/movement/over-the-years/2005/media-coverage-2005> (14 August, 2013.)
- Speak Good English Movement. 2010. Interview with SGEM chairman, Goh Eck Kheng. <http://www.goodenglish.org.sg/about/read-blog/news-updates/stand-on-singlish/> (27 June, 2011.)
- Tagliamonte, Sali. 2005. *So who? Like how? Just what?* Discourse markers in the conversations of young Canadians. *Journal of Pragmatics* 37. 1896-1915.
- Tan, Peter K. W., Vincent B.Y. Ooi & Andy K.L. Chiang. 2003. Pragmatic particles in computer-mediated communication: The case of internet relay chat in Singapore. Paper presented at the Second International Conference on Speech, Writing and Context, Osaka.
- Tan, Peter K.W., Vincent B.Y. Ooi & Andy K.L. Chiang. 2004. Signalling spokenness in personal advertisements on the Web: The case of ESL countries in South East Asia. In Karin Aijmer & Bengt Altenberg (eds.), *Advances in Corpus Linguistics: Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23)*, 151-167. Amsterdam: Rodopi.
- Tan, Ying Ying. 2014. English as a 'mother tongue' in Singapore. *World Englishes* 33. 319-339.
- Turnbull, Constance Mary. 1996. *A history of Singapore: 1819-1988*. 3rd edn. Singapore: Oxford University Press.

- Wee, Lionel. 2003. The birth of a particle: *Know* in Colloquial Singapore English. *World Englishes* 22. 5-13.
- Wee, Lionel. 2004a. Reduplication and discourse particles. In Lim (ed.), 105-126.
- Wee, Lionel. 2004b. Singapore English: Morphology and syntax. In Bernd Kortmann & Edgar W. Schneider (eds.), *A handbook of varieties of English*, vol. 2: *Morphology and syntax*, 1058-1072. Berlin: Mouton de Gruyter.
- Wee, Lionel. 2006. The semiotics of language ideologies in Singapore. *Journal of Sociolinguistics* 10. 344-361.
- Wee, Lionel. 2010. The particle *ya* in Colloquial Singapore English. *World Englishes* 29. 45-58.
- Wee, Lionel. 2011. Metadiscursive convergence in the Singlish debate. *Language & Communication* 31. 75-85.
- Wee, Lionel & Umberto Ansaldo. 2004. Nouns and noun phrases. In Lim (ed.), 57-74.
- Wee, Lionel & Wendy D. Bokhorst-Heng. 2005. Language policy and nationalist ideology: Statal narratives in Singapore. *Multilingua* 24. 159-183.
- Wong, Jock O. 2014. *The culture of Singapore English*. Cambridge: Cambridge University Press.
- Wong, Wee Kim (ed.). 2011. *Census of Population 2010 Statistical Release 1: Demographic characteristics, education, language and religion*. Singapore: Department of Statistics. http://www.singstat.gov.sg/publications/publications_and_papers/cop2010/census_2010_release1/cop2010sr1.pdf (12 December, 2014.)

Authors' addresses

Dagmar Deuber
University of Muenster
English Department
Johannisstraße 12 – 20
48143 Münster, Germany

Jakob R. E. Leimgruber
University of Freiburg
English Department
Rempartstraße 15
79085 Freiburg im Breisgau, Germany

deuber@uni-muenster.de

jakob.leimgruber@anglistik.uni-freiburg.de

Andrea Sand
University of Trier
FBII / Anglistik - English Department
Universitätsring 15
54286 Trier, Germany

sand@uni-trier.de

