SYNCHRONIC ANALYSIS OF THE PROGRESSIVE ASPECT IN THREE VARIETIES OF ASIAN ENGLISHES

SOLUNA SALLES BERNAL
Universidad de Málaga
solunasalles@alu.uma.es

1. Introduction

The diffusion of English worldwide, first as a consequence of British colonialism, and later as a result of the ongoing process of globalization, has greatly transformed the linguistic scenario of many countries. Thus, every continent has both *adopted* and *adapted* more than one recognized variety of English, either as a first or as a second language (foreign contexts excluded), each of them developing a particular "set" of distinctive linguistic features (Kortmann and Szmrecsanyi 2004: 1142). This distinctiveness has been a recurrent topic of research in the last few decades, with the publication of a number of insightful descriptions of divergent features in the different varieties (Platt, Weber and Ho 1984; Kachru 2005; Kachru et al. 2006; Mesthrie and Bhatt 2008). It is not always easy, however, to identify the intrinsic motivations behind such differences, multiple factors being thus postulated. One typical case in point is the progressive aspect and its particular use in non-standard varieties (Kortmann and Szmrecsanyi 2004: 1146).

On historical grounds, the progressive aspect is generally considered to stem from the Old English (OE) construction *be* + present participle as in *he wæs huntende*[1]. It is generally agreed that the progressive construction was an optional choice in OE and Middle English (ME), used stylistically rather than grammatically (Nuñez-Pertejo 2004: 20). Even though the frequency of the progressive remained low

until the beginning of the modern English period (EModE), it was more steadily used from the middle of the 16th century (Elsness 1994). Apart from being the main form for an ongoing action (Aitchison 1991: 109), the progressive later acquired a wide range of meanings, which henceforth contributed to increasing its frequency (Leech et al. 2009: 118; Aarts, Close and Wallis 2010). Thus, Elsness (1994) finds a rise of the progressive by a factor of more than 3 in Early Modern English (EModE), while Smitterberg (2005: 62) presents a growth rate of 71-81 % in the late modern period. A similar tendency is noted in present-day English by Smith (2002), Mair and Leech (2006) and Römer (2005). Leech et al.'s (2009) analysis of British English (BrE) and American English (AmE) covering the period 1960s-1990s associates the increase of the progressive to the higher frequency of certain forms (p. 124)[2]. The rise of the progressive is considered in the literature to be the consequence of several factors, such as *colloquialization* (i.e. linguistic features associated to the spoken language become common in written language), the development of new forms (modal and passive uses) and the occurrence of non-standard uses (i.e. stative verbs in the progressive form) (Collins 2008: 228). The variety of meanings is partly justified by the fact that the progressive is still evolving (Quirk *et al.* 1985: 202), and in this process new meanings are acquired as others decay.

New varieties of English seem to play an important role in the development of extended uses of the progressive form, which, in turn, is connected with the intrinsic evolution accomplished by each particular variety. According to Kachru's Concentric Circle model (1985: 11-36; 2005: 13-14) the South and Southeast Asian Englishes belong to the Outer Circle, where English functions as a second language (L2), developing its own rules for spoken language, but relying on the grammar of native varieties for written texts. If Schneider's Dynamic Model is considered instead, Indian and Hong Kong English are classified as belonging to "phase 3", though in an advanced state of *nativization* moving towards the next phase, while Singapore English is considered to be in "phase 4", already dealing with the process of *endonormative stabilization* (Schneider 2007: 160). Previous studies, such as Collins (2008), Sharma (2009), Van Rooy (2014) and Schilk and Hammel (2014), and more recently Rautionaho (2014), have done research on the question of the progressive aspect in regional varieties. Particularly interesting is the corpus-based investigation carried out by Collins (2008) on nine varieties of English, belonging to both the Inner and the Outer Circle. However, even though several 'variables'[3] are analysed, his results do not seem to be conclusive as to in which variety "the progressive [has] advanced the furthest" (Collins 2008: 246); and, in addition, he finds it difficult to explain the ordering within the Southeast Asian group. One possible explanation could be the restricted set of data used in his study, only 120,000 word samples from the *International Corpus of English*

88

(henceforth ICE), half from conversation and half from printed written registers. More recently, Rautionaho (2014) has studied the progressive in the ICE components of two native varieties (British and American English) and six non-native varieties (Irish, Jamaican, Indian, Philippine, Singaporean, and Hong Kong English) in detail. However, she only used samples from the spoken section of the corpora, from the *private conversation* section in particular. Therefore, research into the use of the progressive in non-native varieties still lacks a complete study that would encompass a larger set of data both from speech and from writing.

In the light of this, this paper attempts to extend the scope of previous studies by analysing the progressive in the complete ICE corpora of three Asian varieties of English, i.e. Hong Kong, India and Singapore, taking British English as a point of departure, with the following objectives: a) to find out if the frequency of the progressive is a distinctive feature among those varieties; b) to analyse its distribution according to tense, subject person and voice; and c) to evaluate the factors affecting the distribution of the progressive.

2. Methodology

The source of analysis comes from the ICE. These corpora fulfil the concept of *comparable corpora* required for a synchronic study, differing only in the territory where language examples were collected[4]. For this study the complete POS-tagged versions of the following ICE components were used: Hong Kong (HKE), India (IndE), Singapore (SingE) and British (BrE). Thus, each of the ICE corpora contains samples of approximately one million words, compiled since 1990 from native speakers aged 18 or above. Each corpus has 500 texts of approximately 2000 words each, both spoken and written (60% and 40%, respectively), with a slight emphasis on private conversations in the spoken mode. This analysis uses the complete ICE corpora, with the exception of extra-corpus material, which was removed from all the corpora[5]. Table 1 reproduces the word count of the ICE components.

	HKE	IndE	SingE	BrE
Spoken	735,082	693,463	625,112	643,015
Writren	496,473	387,713	402,710	428,826
Total	1,231,555	1,081,176	1,027,822	1,071,841

TABLE 1. Word-count for the ICE components analysed

The instances were automatically retrieved by means of AntConc 3.2.4, a freely-available software from the Antlab website (Anthony 2011). A wildcard combination was used to retrieve all the structures containing any form of the verb *to be* occurring with a present participle (*-ing* form). The queries allowed up to a maximum of four words as intervening material between the verb *to be* and the *–ing* form, giving room for different types of examples such as negatives and interrogative clauses, among others. Next, manual disambiguation was needed in order to ignore non-progressive forms *sensu stricto*, such as the catenative construction *to be going to*, and adjectives ending in *–ing* that appeared mistagged (e.g. interesting, boring). Examples of the progressive structures used in this study can be found in Table 2.

(Auxiliary) + *be* form + *-ing* main verb	Tense, mood, aspect, voice
am/'m, is/'s, are/'re + *-ing*	Present progressive active
was/were + *-ing*	Past progressive active
*can/could/may/might/must/shall/should/will/would + be + *-ing*	Modal progressive active
to be + *-ing*	to-infinitive progressive
has/have/had + been + + *-ing*	Present/past perfect progressive active
am/'m, is/'s, are/'re/was/were + being + past participle	Present/past progressive passive
can/could/may/might/must/shall/should/will/would + have + been + *-ing*	Modal perfect progressive

TABLE 2. Types of retrieval classified according to tense, mood, aspect and voice (following Leech et al. 2009)

In addition to these main forms, retrievals included up to four words between the verb *be* and the *–ing* participial, such as adverbs (including *not*) or noun phrases (as in questions). There were no retrievals for the perfect progressive construction in passive voice, e.g. *(s)he has been being taken*.

The non-parametric test Log-Likelihood index (G^2) was applied in order to determine whether the results obtained from each corpus differed significantly in each variety. The calculations were performed using an Excel spreadsheet designed by Paul Rayson and available for downloading at the UCREL website[6]. This index indicates to what degree two samples are different, the higher the log-likelihood, the more significant the difference between both frequencies (Rayson and Garside 2000). Frequencies were then normalized (*n.f.*) according to the total number of words of a given corpus or the number of words in a particular section, e.g. dialogue, as appropriate, according to the following equation:

n.f. = Number of retrievals * 10,000/ word count (whole corpus or corpus section)

3. Results

3.1. The progressive across the Asian varieties of English

The overall normalized frequency of the progressive in the four corpora is presented in Table 3, classified in terms of the variety and section (speech and writing). The data show that as overall BrE has the highest value, followed by IndE, SingE and HKE. The log-likelihood index (Table 4) confirms that the results obtained from HKE are significantly different from the others at a level of $p<0.001$, just as SingE is from BrE, whereas IndE is dissimilar from SingE and BrE at a level of $p<0.05$ and $p<0.01$, respectively. These varieties are observed to differ in their use of the progressive, insofar as HKE seems to be more distant from BrE, while IndE is the most similar, and the smallest difference is to be found between IndE and SingE.

	HKE		IndE		SingE		BrE	
	raw freq.	*n.f.*	raw freq.	*n.f.*	raw freq.	*n.f.*	raw freq.	*n.f.*
Spoken	3032	41,25	4454	64,23	3533	56,52	4165	64,77
Writren	1353	27,25	870	22,44	1308	32,48	1393	32,28
Total	4384	35,60	5324	49,24	4842	47,11	5558	51,85

Table 3. Distribution of the progressive in terms of register variation. *n.f.* stands for *normalized frequency*

	HKE	IndE	SingE	BrE
HKE		249.46***	176.53***	343.80***
IndE			4.98*	7.27**
SingE				23.88***
BrE				

TABLE 4. Log-likelihood index values for each corpus pair[7]

Although most previous studies analysing the progressive in World English varieties have related the occurrences of the progressive to the number of words in the corpus using the M-coefficient (Collins 2008; Sharma 2009; Schilk, 2014; van Rooy 2014), which is in fact the same as the *n.f.*, others prefer to consider the frequency of the progressive construction in relation with the number of verbal phrases (VPs) in the corpus (Smitterberg 2005; Aart, Close & Wallis 2010; Rautionaho 2014). In particular, Rautionaho (2014) has counted the number of VPs present in a fraction of 100,000 words obtained from the spoken section of the ICE corpora in native and non-native varieties. Considering that her selection is representative of the whole

91

corpus and that VPs are homogeneously distributed within each corpus, the V-coefficient was estimated for the data obtained in this study. Thus, Figure 1 shows the values for coefficients in both the Asian varieties and BrE.
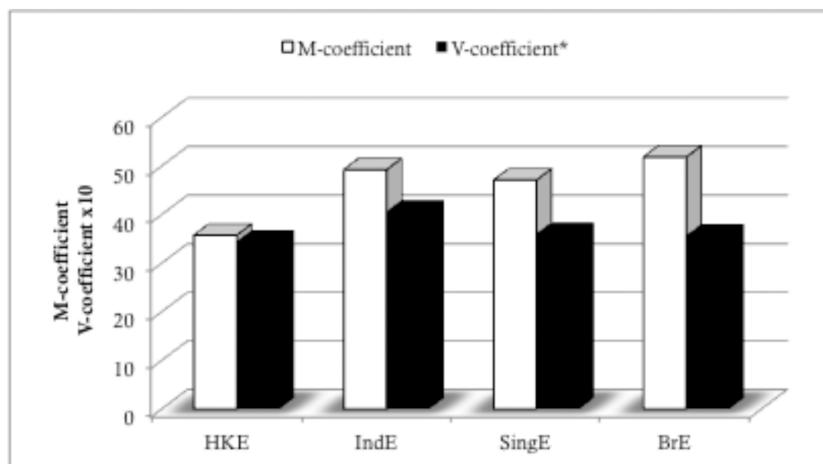
FIGURE 1. M-coefficient and V-coefficient for the progressive in the ICE corpora studied. Note that the V-coefficient must be multiplied by 10

According to our data the M-coefficient shows that the progressive is more frequent in BrE, followed by IndE, SingE and HKE. However, if the V-coefficient is considered, the order changes and IndE has the highest value, followed by SingE, BrE and HKE. The M-coefficient data of this study differs slightly from those presented by Rautionaho (2014), because in her study the order from the highest to the lowest value is IndE > BrE > SingE > HKE. This difference could be attributed to the fact that Rautionaho did not included the *to*-infinitive + progressive construction, which in fact is significantly higher (*p<0.001)* in BrE than in the other varieties considered. Thus, it is clear that the use of one index or the other (type of relative frequency) could slightly influence the results obtained.

3.2. Distribution of the progressive in the spoken and written registers

The frequency of the progressive is higher in the spoken than in the written sections, in all the corpora studied with ratios (*n.f.* spoken/ *n.f.* written) varying from 1.52 in HKE to 2.86 in IndE (Fig. 2). Once again, IndE and BrE present similar frequencies of the progressive, notably in the spoken language, with *n.f.* values of 65.23 and 64.77, respectively.
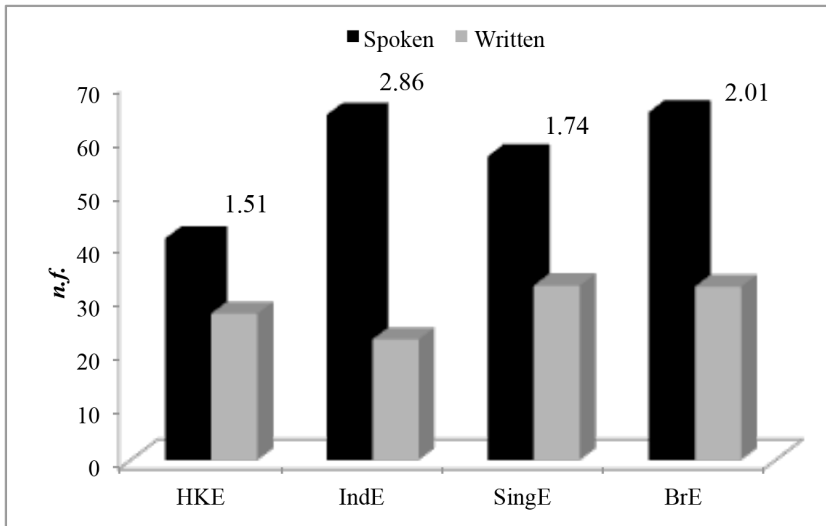
FIGURE 2. Distribution of the progressive in the spoken and written components. (The numbers above the bars indicate the spoken/written ratio)

Contrariwise, IndE has the lowest frequencies of the progressive in the written form (22.44), followed by HKE (27.25), while SingE and BrE present the highest values, 32.48 and 32.28, respectively. While IndE is considered to be syntactically close to the native varieties (particularly BrE), it also exhibits a higher degree of formality, "with a preference for certain syntactic forms" (Sailaja 2009: 39). This could explain the low proportion of the progressive in the writing samples in comparison with the other varieties, such as in *social letters*. Thus, while SingE frequency of the progressive is 112.44 in *social letters* (347 positive retrievals), IndE only shows a value of 52.00 (170 cases).

The spoken register of ICE also distinguishes between dialogue and monologue, the former taken from private and public conversations, such as phone-calls, classroom lessons and parliamentary debates (the latter divided into scripted and unscripted situations, for instance, broadcast talks and commentaries). The data show that the progressive is differently distributed in the spoken samples in the Asian varieties (Table 5). Thus, while in IndE and SingE the progressive is more frequent in dialogue than in monologue (with a ratio of 1.73 and 1.47, respectively), in HKE the use of the progressive seems to be slightly favoured in monologues (ratio 0.88). The phenomenon is found to present a higher frequency in public

than in private conversations in HKE, with a *n.f.* of 41.74 and 37.63 in each case, while the opposite occurs in IndE and SingE (Fig. 3). Within the monologues, however, the unscripted samples show a higher proportion of the progressive than the scripted ones in all corpora.

	HKE		IndE		SingE	
	raw freq.	*n.f.*	raw freq.	*n.f.*	raw freq.	*n.f.*
	Spoken					
Dialogue	1771	39,32	3209	77,35	2441	64,69
Monologue	1261	44,30	1245	44,69	1092	44,07
	Written					
Printed	1016	27,05	522	18,72	847	28,03
Non-printed	336	27,81	348	31,96	461	45,86

TABLE 5. Raw and normalized frequencies of the progressive in the sub-sections of the Asian varieties corpora
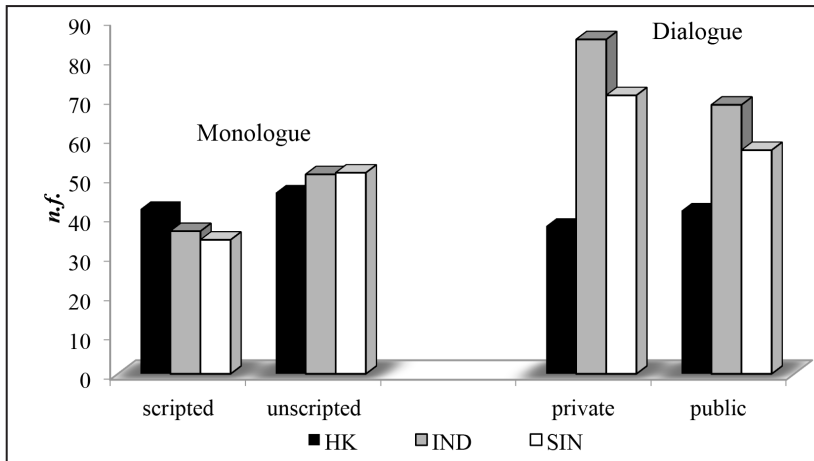
FIGURE 3. Distribution of the progressive within the speech samples

The written component of ICE can also be further divided into printed and non-printed material, the former being more formal in terms of register. The three Asian varieties analysed have higher frequencies of the progressive in the non-printed section: 45.86, 31.96 and 27.81 for SingE, IndE and HKE, respectively.

However, while in IndE and SingE the proportion of the progressive in the printed samples is approximately half of the findings in the non-printed, in HKE the distribution between each part is almost equal. The progressive predominates particularly in *social letters, creative writing* and *press reports* in all the corpus components, which stand out for being more colloquial registers or having adopted colloquial features (Kranich 2010: 102-103).

3.3. Distribution of the progressive by tense and voice

This section discusses the distribution of the progressive across the English verbal paradigm, considering the following verbal constructions: present and past simple, present and past perfect, modal and *to* infinitive, in the active and the passive voices, all of them across speech and writing. As shown in Figure 4, the progressive is clearly associated with the present simple[8] tense in all the varieties, both in the active and in the passive voice, though with slight differences. Thus, in the active voice (Fig. 4.a) IndE has the highest frequencies of the progressive aspect in the present simple, whereas HKE presents the lowest values, not only in the present but also in all the verbal constructions under scrutiny.
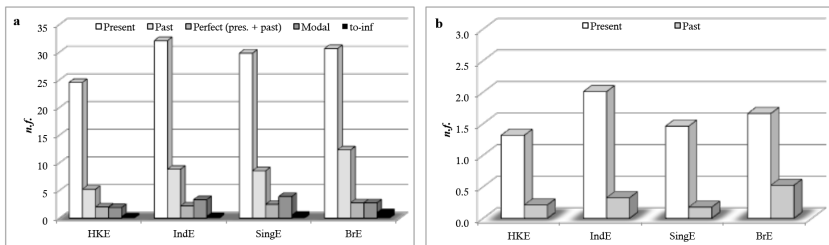
FIGURE 4. Distribution of the progressive aspect across the verbal paradigm, active (a) and passive voice (b). Note that Present and Past refer only to simple progressive, Perfect includes both present and past, and within Modal the few cases retrieved of modal perfect progressive are included[9]

Concomitantly, the statistical analysis confirms that HKE differs significantly from IndE, SingE and BrE at a level of *p<0.001* in the use of the progressive in the present and past simple and modal construction. Conversely, IndE and SingE where not significantly different for any verbal construction at a *p<0.001*, and both varieties differed from BrE for the past simple progressive, the modal construction, and the *to*-infinitive progressive only in the case of SingE.

The past simple progressive in the active voice shows the highest frequency in BrE, a figure that is significantly different from all the Asian varieties (*p<0.001*). In turn,

while IndE and SingE show similar values for the past simple progressive, 8.84 and 8.57, respectively, HKE again stands out with the lowest figure (5.23). Considering the proportion of present and past simple uses in the active voice, the data show that HKE is more reluctant to use the past simple progressive, given that the ratio present/past is significantly higher in HKE than in BrE, 4.69 and 2.48, respectively. On the other hand, IndE and SingE have similar ratios, 3.63 the former and 3.48 the latter. Regarding the differences between spoken and written texts in this matter, BrE exceeds the Asian varieties in both sections, while IndE shows less use of the past simple progressive in the written samples than in the spoken ones. HKE and SingE maintain a similar proportion for the spoken and the written progressive (Figure 5).
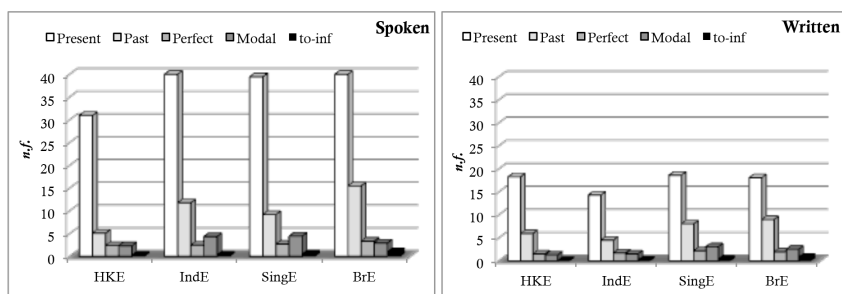
FIGURE 5. Distribution of the progressive for the spoken and written section of the corpora (active and passive voices considered)

The perfect progressive only occurred in the active voice as no structure with the form *have/has been being* + past participle was retrieved. The frequency of the perfect progressive forms, both present and past, show a similarly distribution among the corpora. Thus, BrE has a *n.f.* of 2.79, SingE of 2.51, IndE of 2.22 and finally HKE with 2.16. In the same vein, values for the spoken and the written sections show a similar distribution among the varieties. However, while BrE presents the highest frequency in the spoken language, SingE and IndE slightly exceed the rest of corpora in the written texts (Figure 4). These results agree with previous studies, which in general find a small proportion of perfect progressive, and no significant difference among the varieties analysed (Collins 2008: 233, Rautionaho 2014: 121). Some examples of perfect progressive are shown below:

(1) *Since* the start of the Industrial Revolution air pollution *has been creating* biological deserts around industrial centres (ICE-GB, W2A-030)

(2) We *have been trying* to resolve problems (ICE-IND, S1B-036)

(3) Eighteen year old Noi *has been working* here *for just over a year* (ICE-SIN, S2B-025)

(4) *I've been eating* seafood *every* day (ICE-HK, S1B-045)

(5) Uh this is different uh from uh what we'*ve been doing all the time* in the past (ICE-HK, S2B-048)

The perfect progressive expresses a situation in progress that started somewhere in the past but that is still unfinished in the present. It focuses particularly on the duration of the event, which seems "to invite the use of temporal adverbials" (Kranich 2010: 140). According with the examples presented above, the co-occurrence with temporal adverbials is common but not strictly necessary, as in example (2). In other cases, rather than being used to emphasize the duration of the event, as is the case in (1) and (3), the adverbial indicates repetition, as in (4) and (5). The question of perfect progressive co-occurring with adverbials has been addressed by Rauthionaho (2014). Her analysis based on conversation samples concludes that SingE seems less prone to use adverbials of time modifying the perfect progressive than other OC and IC varieties (Rauthionaho 2014: 122-125). It would be interesting to extend the analysis to other type of spoken samples as well as written texts, in order to test Smitterberg's hypothesis, which postulates that in those varieties where the progressive is more integrated the need for a temporal marker is lower (2005: 188).

The combination of modal + progressive appears more frequently in SingE and IndE than in BrE and HKE, particularly in the spoken language (Figure 4), where SingE has a frequency of 4.46, IndE of 4.37, BrE of 2.94 and HKE of 2.34. This construction can be further unfolded according to which modal verb is more or less frequently combined with the progressive. Within the modals, the construction *will* be +*-ing* form is considered to be one of the 'special' uses of the progressive, (Leech et al. 2009: 139), as it can be applied to an event in progress set in the future (examples 6 and 7), or referring to a future situation not in progress (examples 8 and 9). See examples below:

(6) By the time you arrive, *I 'll probably be having* my final exams (ICE-HKE, W1B-010)

(7) *They all will be doing* development work and then when users have problem they wo n't call us directly […] (ICE-SIN, S1A-045)

(8) Matthew has just finished his exams and *will be leaving* for a short trip, either Thailand or Indonesia. (ICE-SIN, W1B-008)

(9) So uh and why I am asking it now and not when the final printout is out because uh *I'll be asking* for funds in April or May (ICE-IND, S1B-071)

The analysis of the data shows that this combination, also called the futurate progressive, represents between the 40% and the 60 % of all the modal progressives, and that this proportion changes among the varieties and field (spoken or written). Thus, as a whole the construction *will* + progressive was favoured in the Asian varieties in comparison with BrE. In SingE the modal *will* represents 64.5 % of the total modal progressives, followed by IndE with 62.15%, HKE with 55.27% and BrE with just 40.60%. In addition, all the Asian varieties show a higher percentage of *will* + progressive in the spoken than in the written samples, while BrE presents the opposite situation.

In turn, the use of the *to*-infinitive progressive, e.g. *to be asking*, presents the lowest values in all the corpora, appearing more frequently in BrE than in the Asian varieties, both in speech and writing. Thus, for the whole BrE corpus the percentage of *to*-infinitve progressive is 2.11% out of all verbal constructions analysed, followed by SingE with 1.38%, HKE with 1.00% and IndE with the lowest value, 0.85%. However, in BrE this construction is more common in speech than in writing, the difference between the native and the non-native varieties is larger in the written section, by a factor of 1.9 in comparison with the speech. To sum up, BrE uses more *to*-infinitive progressive constructions than the Asian varieties, particularly in written texts. The non-finite progressive is normally used to express an action that is repeated with a progressive meaning, as is shown in examples 10, 11, and 12. However, in other cases it is used to present an action in progress with the emphasis in the action as in example 13.

(10) to treat an individual who is diagnosed *to be suffering* from deviance (ICE-HK, W1A-012)

(11) Because we don't want *to be saying* you know the same thing (ICE-HK, S1A-053)

(12) He seemed *to be following* Gary Schofield everywhere (ICE-GB, S2A-004)

(13) And they are supposed *to be writing* in their English medium uh in exams or whatever projects (ICE-SIN, S1A-071)

The progressive appears in the passive voice less frequently in our data, not only in terms of its total frequency but also in terms of the overall distribution of tenses. Thus, only 2 passive uses have been identified, namely: present and past. Once again the progressive in the passive voice shows higher frequencies in IndE, followed by BrE and SingE with very similar values for the present passive, and finally HKE presenting the lowest frequencies in both present and past progressive forms (Fig. 3.b). The statistical analysis proves that in the passive voice HKE shows differences with respect to IndE in the use of the present progressive, and with respect to BrE past (*p<0.001*), whereas no difference is found between HKE and

SingE. For IndE no statistical difference appears if compared with BrE, and only a slight variation is found from SingE in the present tense ($p<0.05$). Finally, SingE and BrE only differs in their past use of the passive progressive at a level of $p<0.001$.

A further analysis of the passive constructions shows that SingE has the highest values for the present vs. past ratio, while BrE presents the lowest. In fact, the three Asian varieties present similar ratios: 8.33 for SingE, 6.23 for HKE, 6.20 for IndE, while BrE retrieves only 3.18. These results agree with the fact that Outer Circle (OC) varieties try to avoid more complex syntactic structures as in the case of the combination of past tense, passive voice and progressive aspect. As Rautionhao (2014: 106) proposed "the fact that present tense progressives are more frequent in OC varieties may be regarded as indication that the progressive is not as well established in OC varieties as it is in IC [Inner Circle] varieties". Thus, according to our data HKE presents the highest proportion of the present simple progressive in relation with all the progressive forms of the paradigm, with a 72.34%. SingE and IndE have similar proportions, 63.30% and 65.15%, confirming the closeness of these varieties, while in BrE the percentage of present simple progressive decreases to 62.32%. Consequently, it seems that the progressive is more uniformly distributed among the verbal paradigm in the native variety (BrE in our case) than in the non-native Englishes, as HKE, IndE and SingE.

3.4. Distribution in terms of subject person and number of the present active progressive

The distribution of the progressive in terms of subject person and number has been analysed in the present simple active voice (Figure 6). Again, the greatest differences are found between HKE in comparison with IndE and BrE. Thus, HKE and IndE significantly differ at a level $p<0,001$ for 1st person singular and 3rd person plural (both pronoun and NPs). On the other hand, HKE and BrE are different at the same level of probability for 1st, 2nd and 3rd person singular (pronoun). Moreover, SingE presents statistical differences ($p<0,001$) with HKE and IndE for the 2nd person singular. Finally, IndE and SingE have similar values to those of BrE.

All the corpora present the higher frequency of the progressive associated with the 3rd person singular NPs, followed by the 1st person singular, apart from SingE that shows a higher use of the 2nd person. In addition, the lowest frequencies of the progressive are found with the 3rd person plural pronoun. IndE has a higher use of the progressive than the rest of varieties for the 3rd (NPs) and the 1st person singular (7.79 and 5.48, respectively). The variability among the corpora can be analysed considering the standard deviation (*SD)*. The *SD* shows that the spoken samples have more variability, while the written ones are more conservative

regarding the choice of subject person in the progressive constructions when varieties are compared. In turn, the 2nd and the 1st person singular are more irregular in terms of the distribution of the progressive, and the 1st person plural and the 3rd person plural NPs are more stable (Table 6).

The proportion of each subject person occurring in speech and writing shows a heterogeneous distribution among the Asian varieties, proving that the progressive aspect is used differently. Thus, in SingE the first and the third person singular (pronouns) seem to be less frequent in speech than in writing if compared with IndE and BrE; the ratios for spoken v. written being 1.45 and 2.69 for SingE, 2.89 and 6.20for IndE and 2.34 and 3.09 for BrE.
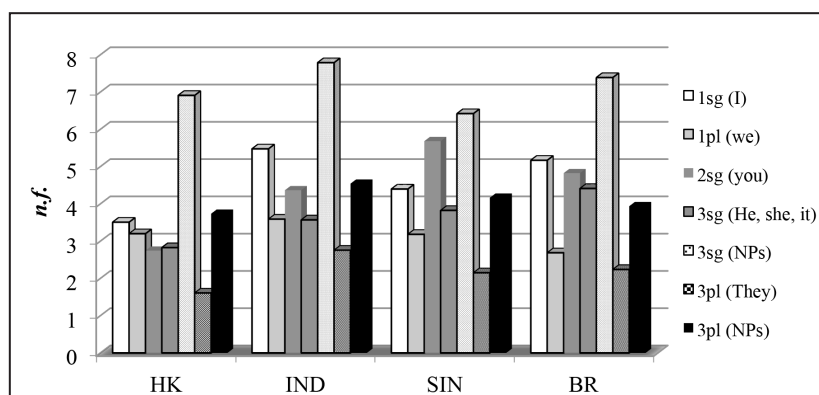
FIGURE 6. Distribution of the progressive by subject person and number in the present simple (active voice)

	n.f. average	*SD*
1sg (I)	4,64	0,88
1pl (we)	3,16	0,37
2sg (you)	4,40	1,23
3sg (He, she, it)	3,66	0,66
3sg (NPs)	7,13	0,59
3pl (They)	2,19	0,47
3pl (NPs)	4,09	0,35

TABLE 6. Average values of *n.f.* and Standard Deviation (*SD)* of the distribution of the progressive in all the corpora

The high proportion of the progressive with the 3rd person singular and plural NPs in all the varieties seems to agree with the diachronic tendency shown by the progressive to co-occur more frequently with non-agentive and/or inanimate subjects (Kranich 2010: 143). However, the agentive and/or animate subjects continue to be the preferred choice for progressive constructions, and of these, animacy seems to weight more than agency (c.f. 146). As the main function of the progressive is to express a dynamic situation, an event in progress, this obviously requires the input of energy supplied usually by an animate and agentive subject. However, other combinations are possible, for instance [- animate, + agentive] or [+ animate, - agentive] in metaphorical uses (Hundt 2004: 50). Although this analysis goes beyond the aim of the present study, an overview of the 3rd person singular NPs group shows that the vast majority correspond with inanimate subjects, while the question of agency is more evenly distributed. The fact that SingE has a lower proportion of the progressive in the 3rd person singular NPs and a higher proportion in the 2nd person seems to indicate a tendency in this variety towards agency and animacy, though this should be further contrasted by analysing each case in particular. The types of possible combinations of the progressive with animate/inanimate and agentive/non-agentive subjects occurring in the 3rd person singular NPs group are illustrated in the following examples:

(15) *Prof. Nadkarni is arranging* accommodation for two days (ICE-IND, W1B-004) [+ animate, + agentive]

(16) If the *government was acting* on its own the clear message to us or to Beijing is that it does not want to take any responsibility for making a decision or any risk of giving the impression that *Hong Kong is acting* too independently or *trespassing* on Beijing's prerogative (ICE-HK, S2B-031) [- animate, + agentive]

(17) Premchand shows how *the daughter-in-law of a man is dying* (ICE-IND, S1A-006) [-animate, -agency]

(18) The hope of owning a bigger flat *is becoming* more and more like an impossible pursuit (ICE-SIN, W1B-022) [- animate, - agentive]

4. Conclusions

This paper deals with the use of the progressive form in some South and Southeast Asian English varieties, namely Hong Kong, India and Singapore, to find out whether the progressive aspect can be identified as a grammatically distinctive feature in those varieties. The data used as source of evidence come from the *International Corpus of English*, which has a selection of comparable corpora for

the varieties surveyed, designed with the same structure and chronology, and of a similar dimension.

Our study has allowed us to draw the following conclusions. First, the distribution of the progressive aspect is not homogenous across the South and Southeast Asian varieties of English, insofar as HKE and IndE are found to present the lowest and the highest occurrence, respectively. The highest frequency found in IndE could be justified, at least partially, as a result of basilectal influence, as has been pointed out by Sharma (2009) in her analysis of the progressive co-occurring with stative verbs. She postulates that the fact that Hindi (the main substrate language for IndE speakers) compulsorily marks all imperfectives is transferred to IndE, producing an overuse of the progressive among Indian speakers of English. Although in terms of frequency of the progressive IndE appears close to SingE and BrE (see Table 5), the functions addressed by the progressive in each variety could be different and they should be further investigated. In particular it would be interesting to investigate further those non-standard uses (i.e. stative verbs occurring in the progressive form in speech and writing), or the preference of some verbal forms (i.e. modal progressive and futurate uses).

In turn, SingE, although with a high frequency of the progressive, both in the spoken and in the written forms, seems to be more constrained than IndE. The basilectal transfer is also contemplated by Sharma (2009) as a plausible cause, considering "the restricted use of imperfective marking in Mandarin (as compared with Hindi)". Certainly the wide range of substrate languages spoken in the area, such as Baba Malay, Bazaar Malay, Cantonese, Hokkien and Mandarin (Deterding 2007: 3), makes it difficult to draw reliable conclusions. On top of that, SingE is presently in an advanced stage of development compared with the other Asian varieties, achieving its *endonormative stabilization* phase according to Schneider's Dynamic Model (2007). This model considers that a certain variety develops through a dynamic process by constructing a particular linguistic identity, which is the result of the encounter between the local/native languages and the transplanted variety. The *endonormative stabilization* supposes that the population shares a common sense of nationhood, which is reflected in the language by the acceptance of linguistic norms, increase of linguistic self-confidence, the emergence of local dictionaries and the development of a local literary creativity (Schneider 2007: 160). In this context, the difference in the use of the progressive in relation with the other varieties could be, at least partially, based on the development of a singular linguistic identity as a result of the dynamic evolution of languages in contact.

It remains an open question why HKE should have such a low proportion of the progressive, having reached the same stage of development as IndE in Schneider's

model and with both of them being OC varieties. It seems difficult to find one particular reason, and we have to recognize that probably many factors intervene in this case, historical, basilectal influence, and SLA features. Our data agree with those of Rautionaho (2014) placing HKE in the lowest position of the frequency of occurrence of the progressive. In her study, she proposes two main reasons for HKE's underuse of the progressive. One is again basilectal influence, since Cantonese, the main substrate language for HKE corpus speakers, has an optional marker for the progressive. Consequently, HKE speakers do not feel the need to mark progressivity. The second possible cause is the occurrence of an inflected form with progressive meaning, i.e. *be* + base form of the main verb, which could replace the 'typical' progressive. However, the low number of VPs found in HKE corpus by this author (10,416) seems to indicate that, in general, verbal constructions are less used by HKE speakers, and for some reason the progressive form in particular even less. The other intriguing question about HKE is the relatively high proportion of the progressive in the written as compared with the spoken section. The low spoken/written ratio indicates that the progressive is particularly restricted in speech, but not in writing, with similar values to those of SingE for instance. Within the written section, HKE has a similar distribution of the progressive contrasting with the other varieties that present a significantly larger proportion of the progressive in non-printed samples, the latter being considered as more colloquial in style.

Secondly, the results show a higher occurrence of the progressive in the spoken domains. The spoken/written ratio ranks the varieties as follows: HKE (1.51) < SingE (1.74) < BrE (2.01) < IndE (2.86). In the case of HKE, this result is justified by the low frequency of the progressive in speech and a relatively high proportion (similar at least to the other varieties) in writing. On the other hand, IndE presents the lowest occurrence of the progressive in writing and the highest in speech. It can be hypothesized that written IndE remains attached to previous formal norms of the native variety, and consequently it is less prone to use the progressive. The distribution of the progressive among the sub-sections of the corpora seems to confirm that this form is associated with a colloquial use. In this vein, dialogues outnumber monologues, and within dialogue, private conversations present a higher frequency than public ones, while the unscripted outnumbers the scripted monologues. The same pattern is found in the writing section, where non-printed samples have a higher proportion of the progressive in comparison with printed registers. The phenomenon of *colloquialization* could have contributed to the expansion of the use of the progressive from the more informal spoken language to written texts (Mair and Hundt 1995: 118). This would explain the cases of SingE and HKE, which have a relatively high proportion of the progressive in written samples.

Thirdly, the distribution of the progressive across the verbal paradigm presents the same pattern among the varieties under scrutiny. Thus, the progressive is ranked as present simple > past simple> modal > perfects > *to*-infinitive in descending order of frequency in the active voice (speech and writing considered). However, while the modal progressive is more frequent in IndE and SingE as compared with BrE, the latter outnumbers the Asian varieties in the use of the *to*-infinitive construction. It seems that the high proportion of the modal progressive in IndE and SingE is due to the extended use of the construction *will* + *be –ing*, which is more frequent than constructions with modal auxiliaries other than 'will' in these varieties. In addition, all the Asian varieties studied present a significantly lower proportion of the past simple progressive than the native variety. This could be explained by second language acquisition features, as L2 speakers could have more difficulties in combining tense and aspect, since the progressive is described as troublesome for L2 learners according to Swan and Smith (2001: ix).

Fourthly, the distribution by subject person and number shows that the progressive form is more likely to occur with singular persons (NPs >1st person > 2nd person), even though this fact should be further validated considering the relative contribution of each person to the whole set of data. Previous studies, such as Leech et al. (2009) also find a rather erratic distribution of the progressive in this respect. In turn, the high proportion of NPs subjects occurring with the progressive shows a tendency in non-native varieties towards the use of inanimate and/or non-agentive subjects, which was also found in native varieties (Kranich 2010: 143).

To conclude, our results agree with previous studies arguing that the distribution of the progressive that emerged is a distinctive feature of non-standard (Outer circle varieties) varieties of English. Thus, while in HKE the progressive appears rather constrained, IndE shows the other side of the coin with an extended use of the phenomenon, SingE remaining in-between. In addition, lexical differences have been found in the use of certain verbs in the progressive, which in the case of IndE may be due to an extended (non-native) usage. Although these differences have been attributed mainly to basilectal transfer, this does not explain the high frequency of the progressive found in the native variety, here BrE. Therefore, other factors may have played an active role, such as the transference of spoken features to the written forms by means of *colloquialization* and the development of new uses of the progressive, such as its extension to non-aspectual meanings or non-standard uses, which should be further investigated.

Notes

[1] The origin of the progressive has also been postulated from other perspectives. Thus, it has been observed to derive from the locative construction *be* + a preposition and gerund as in *he wæs on huntende* (Leech *et al.* 2009: 120); or ever as an adjectival derivation, which eventually lost the adjectival dimension to acquire verbal ones, ending up as a verb form (Nuñez-Pertejo 2003).

[2] Another common finding is that the progressive is more frequent by far in spoken than in written registers, and also particularly associated with the present tense and the active voice.

[3] Collins (2008) offers a detailed study of the progressive in terms of speech and writing, the verbal paradigm, the semantic classification of verbs, special uses, grammatical environment, and contraction.

[4] For more information about the ICE project see Greenbaum 1996, Greenbaum and Nelson 1996.

[5] Following the helpful suggestions of the reviewers, the speaker Z <Z>, a native speaker appearing in the HKE, was removed as well as the other extra-corpus material (marked as <x>/</x>) in all the corpora used.

[6] (http://ucrel.lancs.ac.uk/llwizard. html)

[7] "*" indicates a significant difference at the level of *p<0.05* (critical value 3.84), "**" at the level of *p<0.01* (critical value 6.63) and "***" at a level of *p<0.001* (critical value 10.83).

[8] The terminology applied follows the one proposed by Quirk (1985), in which the term "simple" contrasts "perfective" uses, both in the present and in the past tenses. Accordingly, the present simple progressive (i.e. he is examining) signals the difference with the present perfective progressive (i.e. he has been examining), and the past simple progressive (i.e. he was examining) does the same with the past perfective progressive (i.e. he had been examining) (Quirk 1985: 189). Another nomenclature is the one proposed by Biber et al, where the term "progressive aspect present/past tense" equates to the "present/past simple progressive" mentioned before, and the "perfect progressive aspect" refers to the "perfective progressive" uses (1999: 460-461)

[9] Modal perfect progressive only retrieved 8 cases for SingE, 5 for IndE, 4 for BrE and 2 for HKE.

Works cited

AARTS, Bas, Joanne CLOSE and Sean WALLIS. 2010. "Recent changes in the use of the progressive construction in English". In Cappelle Bert and Naoaki Wada (eds.) *Distinctions in English Grammar, Offered to Renaat Declerck*. Tokyo: Kaitakusha: 148-167.

AITCHISON, Jean. 1991. *Language Change: Progress or Decay?* Cambridge: Cambridge U.P.

ANTHONY, Laurence. 2011. "AntConc (3.2.4w)" [Computer software] Tokyo, Japan: Waseda University.

BIBER, Douglas, Stig JOHANSSON, Geoffrey LEECH, Susan CONRAD, and Edward FINEGAN. 1999. *Longman Grammar of Spoken and Written English* (Vol. 2). MIT Press.

COLLINS, Peter. 2008. "The Progressive Aspect in World Englishes: A Corpus-based Study". *Australian Journal of Linguistics* 28 (2): 225-249.

DETERDING, David. 2007. *Singapore English*. Edinburgh: Edinburgh U.P.

ELSNESS, Johan. 1994. "On the progression of the progressive in early Modern English". *ICAME Journal* 18: 5-25.

GREENBAUM, Sidney (ed.) 1996. *Comparing English Worldwide: The International Corpus of English*. Oxford: Clarendon.

GREENBAUM, Sidney and Gerald NELSON. 1996. "The International Corpus of English (ICE) Project". *World Englishes* 15 (1): 3-15.

HUNDT, Marianne. 2004. "Animacy, Agentivity, and the Spread of the Progressive in Modern English". *English Language and Linguistics* 8: 47-69.

KACHRU, Braj B. 1985. "Standards, Codification and Sociolinguistic Realism: the English Language in the Outer Circle". In Quirk, Ralph and Henry Widdowson (eds.) *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge U.P.: 11-36.

—. 2005. *Asian Englishes: beyond the canon*. Hong Kong: Hong Kong U.P.

KACHRU, Braj, Yamuna KACHRU and Cecil NELSON (eds.) 2006. *The Handbook of World Englishes*. Oxford: Blackwell.

KORTMANN, Bernd, and Benedikt SZMRECSANYI. 2004. "Global synopsis – morphological and syntactic variation in English". In Kortmann, Bernd, Kate Burridge, Rajend Mesthrie and Edgard Schneider (eds.) *A Handbook of Varieties of English,* Vol. 2: *Morphology and Syntax*. Berlin, New York: Mouton de Gruyter: 1122-1182.

KRANICH, Svenja. 2010. *The Progressive in Modern English: A Corpus-based Study of Grammaticalization and Related Changes*. Amsterdam: Rodopi.

LEECH, Geoffrey, Marianne HUNDT, Christian MAIR and Nicholas SMITH. 2009. *Change in Contemporary English: a Grammatical Study*. Cambridge: Cambridge U.P.

MAIR, Christian and Marianne HUNDT. 1995. "Why Is the Progressive Becoming More Frequent in English? A Corpus-Based Investigation of Language Change in Progress". *Zeitschrift für Anglistik und Amerikanistik* 43: 111-122.

MAIR, Christian and Geoffrey LEECH. 2006. "Current Changes in English Syntax". In Aarts, Bas and April McMahon (eds.) *The Handbook of English Linguistics*. Oxford: Blackwell: 318-342.

MESTHRIE, Rajend, and Rakesh M. BHATT. 2008. *World Englishes: the study of new linguistic varieties*. Cambridge: Cambridge U.P.

NUÑEZ-PERTEJO, Paloma. 2003. "Adjectival participles or present participles? On the classification of some dubious examples from the Helsinki Corpus". *SEDERI: yearbook of the Spanish and Portuguese Society for English Renaissance Studies* (13): 141-154.

—. 2004. "Some developments in the semantics of the English progressive from Old English to Early Modern English". *Revista Alicantina de Estudios Ingleses* (17): 211-226.

PLATT, John T., Heidi WEBER and Mian Lian HO. 1984. *The New Englishes*. London: Routledge and Kegan.

QUIRK, Randolph, Sidney GREENBAUM, Geoffrey LEECH and Jan SVARTVIK. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

RAUTIONAHO, Paula. 2014. *Variation in the Progressive: A Corpus-based Study into World Englishes*. PhD Dissertation, University of Tampere.

RAYSON, Paul, and Roger GARSIDE. 2000. "Comparing corpora using frequency profiling". In *Proceedings of the Workshop on Comparing Corpora. Association for Computational Linguistics*: 1-6.

RÖMER, Ute. 2005. *Progressives, Patterns, Pedagogy: A Corpus-driven Approach to English Progressive Forms, Functions, Contexts and Didactics* (18). Amsterdam/ Philadelphia: John Benjamins Publishing.

SAILAJA, Pingali. 2009. *Dialects of English. Indian English*. Edinburgh: Edinburgh U.P.

SCHILK, Marco and Marc HAMMEL. 2014. "The progressive in South Asian and Southeast Asian varieties of English–mapping areal

homogeneity and heterogeneity". *Language and Computers* 78 (1): 147-171.

Sᴄʜɴᴇɪᴅᴇʀ, Edgar W. 2007. *Postcolonial English: varieties around the world*. Cambridge: Cambridge U.P.

Sʜᴀʀᴍᴀ, Devyani. 2009. "The typological diversity in New Englishes". *English World Wide* 30 (2): 170-195.

Sᴍɪᴛʜ, Nicholas. 2002. "Ever moving on? The progressive in recent British English". In Peters, Pam, Peter Collins and Adam Smith (eds.) *New Frontiers of Corpus Research: Papers from the Twenty First International Conference on English Language Research on Computerized Corpora*. Sydney 2000. Amsterdam: Rodopi: 317-330.

Sᴍɪᴛᴛᴇʀʙᴇʀɢ, Erik. 2005. *The Progressive in 19th-Century English: A Process of Integration*. Amsterdam and New York: Rodopi.

Sᴡᴀɴ, Michael and Bernard Sᴍɪᴛʜ. 2001. *Learner English. A Teacher's Guide to Interference and Other Problems.* 2nd edition. Cambridge: Cambridge U.P.

Vᴀɴ ʀᴏᴏʏ, Bertus. 2014. "Progressive aspect and stative verbs in Outer Circle varieties". *World Englishes* 11 (2): 157-172.

107