The Asian Journal of Applied Linguistics
Vol. 3 No. 1, 2016, pp. 114-124

AJAL

Interpreting error patterns in a longitudinal primary school corpus of writing

Lubna Alsagoff
English Language and Literature, National Institute of Education, Nanyang Technological University, Singapore

This paper examines the error patterns in a small longitudinal corpus of Singaporean primary school children's writing and argues that corpus data, although highly effective in uncovering patterns that might otherwise go unnoticed, need to be meaningfully interpreted for them to be useful in contributing to improvements in pedagogy. Such interpretations require that researchers take into account a variety of other factors in addition to textual variables, including the sociolinguistic context and the particular pedagogic practices of the local educational system.

Keywords: learner corpus; grammar; writing; spelling; punctuation; Singapore

Introduction

Linguistic corpora have proven eminently useful in offering an empirical basis for investigating the patterns of language use, often uncovering linguistic features that might otherwise have gone unnoticed (Biber & Reppen, 1998). However, while such electronically compiled collections of text may provide much needed objective data, the insights that they yield are very dependent on the researchers' interpretations of such data. Baker (2006) cautions that corpus data do not interpret themselves, and it is up to the researcher to make sense of the patterns of language that such data might reveal. Conrad (2010) offers a similar cautionary note in her discussion of learner corpora, asserting that they should only be seen as identifying patterns of language use that still require interpretation as well as further investigation. Urzúa (2015), likewise, points out that interpreting learner corpus data cannot be undertaken solely on the basis of frequency data about the text variables, but crucially involves knowledge about the learners as well as contextual information about educational settings, and should preferably involve teachers. Only with such context-specific insights can learner corpora be expected to result in successful pedagogical applications (Hasko, 2013).

As such, in presenting the findings from a project involving the grammatical analysis of data from a longitudinal written corpus of Singaporean primary-school children's writing[1], this paper highlights the ways in which various aspects of the project focused on achieving outcomes that would be of optimal use to Singapore's educational policy and classroom practices, from the way in which the coding was carried out to the attention paid to the qualitative contextualised interpretations of quantitative analyses. The paper begins with a review of the research literature regarding learner corpora, with particular focus on how such corpora can lead to the improvement of language pedagogy and practice. Then a description of how the learner corpus in the present study was developed is presented. This is followed by a discussion

ISSN 2308-6262
http://caes.hku.hk/ajal

of the ten most frequent error types found in the learner corpus with an examination of their sources and possible explanations.

Learner corpora

Learner corpora, collections of texts produced by learners, (Granger, 2002; Sinclair, 2004) are often used to investigate common problem in language learning by uncovering patterns in the use of linguistic features , for instance, adverbs (Pérez-Paredes & Sánchez-Tornel, 2014), linking adverbials (Conrad, 2004), modal verbs (Römer, 2005), tense and aspect (Granger, 1999; Römer, 2005), time expressions (Mindt, 1997), verb-noun collocations (Nesselhauf, 2005). Thus, inherent in their design and construction is the fundamental belief that learner corpora are developed to contribute to the improvement of language education.

Nesselhauf (2004) discusses the three ways in which learner corpora can result in pedagogical applications. Firstly, learner corpora can inform the preparation of teaching materials, for example, corpora containing learner writing help identify "features of learner language … [that can help] focus teaching methods and contents more precisely so as to speed acquisition" (Stewart, Bernardini, & Aston, 2004, p. 3). Secondly, and more ambitiously, corpora can contribute to data-driven language learning (Johns, 1994). Thirdly, corpora can provide general theoretical understandings of language development that help inform policy-making and curriculum development.

More typically, learner corpora are compiled to study the language development of EFL or ESL learners (Granger, 2012). In such contexts, learner corpora are examined through the use of computer-aided error analysis (Granger, 2002; Thewissen, 2013) that typically compares the language output of EFL learners against a similar corpus of native speakers of the target language or the educated speech of adult speakers of the language, or contrastive interlanguage analysis (Granger, 1996, 2015) that typically compares the speakers' L1 and L2 language outputs. In multilingual postcolonial contexts such as Singapore, this comparative approach is not as straightforward because the norm against which learner language is compared may not be easily determined. In particular, it is not always clear whether speakers can be classified as ESL learners given that for many speakers and learners, English is their predominant language, both at home and at work[2].

Sociolinguistically, multilingual postcolonial contexts such as Singapore, which are characterised as Outer Circle countries (Kachru, 1992), exhibit a great deal of heterogeneity in terms of language behaviour, partly because of the varied linguistic backgrounds of the community of speakers, but also because there usually exist two different linguistic norms (Kirkpatrick & Sussex, 2012). On the one hand, there is a local variety that is widely used among speakers of the community and is a *de facto* endonormative standard; and on the other there is an exonormative standard, usually an exogenous variety of English such as British or American English, that is considered as the more desirable standard, which is taught in schools because of its established international prestige and currency in global economic, technological and financial marketplaces. Consequently, despite theoretically influential descriptions that characterise Singapore English as having achieved endonormative stabilisation (e.g., Schneider, 2007) with an established local variety of English that serves a wide range of linguistic domains, there is still a strong institutional push by the Singapore authorities towards the use of British English as the standard in schools which is clearly done to ensure that English continues to pivotally serve Singapore's economic interests and maintain its position as a global nation (Alsagoff, 2010).

It is therefore pertinent, given Singapore's complex sociolinguistic landscape, to note O'Keeffe, McCarthy, and Carter's (2007) concerns that learner corpus research has predominantly adopted a comparative approach which privileges the native speaker (Cook, 1999). Such an approach may be understandable and applicable, even in contexts such as Singapore where there are emergent local norms, if comparisons with the standard language of native speakers are seen as a means of allowing researchers to discover and address learner difficulties (Nesselhauf, 2004). Precedence for this reliance on a comparative approach also lies in the fact that native-speaker corpora have been used in the production of teaching and reference materials that provide attested examples of native speaker language use. In addition, it is often a lack of availability and documentation of local standards that makes it necessary to rely on British or American English as standards of reference internationally[3]. The project described here adopted a comparative approach, albeit one mediated by local norms.

Compiling the corpus

The corpus data discussed in this paper is a subset of the data from a longitudinal corpus developed at the National Institute of Education in Singapore that sought to provide insights into the language development of Singapore primary school children in the area of grammar, at the word, sentence and text levels. This corpus was compiled from a collection of 2,139 handwritten essays of 351 pupils who participated in a research study conducted by the Ministry of Education, Singapore (MOE), between 2007 and 2012, as part of a programme to evaluate the effects of a then newly-launched reading and literacy programme for Singapore primary schools[4]. The MOE study sampled students from twenty government-funded co-educational neighbourhood schools, generally seen as representative of the average student population in Singapore, and which did not include high performing schools. Students were sampled randomly from within each of these twenty schools but excluded those with identified learning difficulties.

In this study, the 351 pupils wrote one essay at the start of Primary 1, and then one at the end of each year of their six years of primary school as part of a set of tasks in the MOE study. Each participant thus contributed seven pieces of writing to the corpus. The student compositions were narratives written in response to picture stimuli, comprising three illustrations, presented in the form of a task familiar to Singapore students[5]. The writing task was similar in format across all years except Primary 4, in which the students had the option of writing on any topic they chose. The majority of the Primary 4 texts were narratives with only a very small number of students writing personal recounts and informational texts, which were excluded from the error-coded corpus. The students in the primary grades 1 and 2 were given 30 minutes to write, while the rest of the students were given 45 minutes to complete their writing tasks. Word count rose from an average of 108.6 at Primary 2 to 300.3 at Primary 6.

The research team chose to narrow its focus to the error coding of the Primary 2 to Primary 6 essays written by 233 students, comprising 271,300 words[6]. This subset represents a true longitudinal corpus of students who participated from the start of Primary 1 to the end of their primary school education at Primary 6. Although small, this corpus is valuable because true longitudinal corpora are difficult to compile and therefore rare (e.g., Bernadini, 2004; Thewissen, 2013).

To address concerns that learner corpora often privilege the native-speaker orientation, the error tagging employed in the study attempted to provide a locally-mediated norm through the use of a reference grammar text written by a Singaporean

(Alsagoff, 2007) which has been one of the key standard references in the grammar content-knowledge and grammar pedagogical content knowledge courses at the National Institute of Education (Singapore's national teacher education institute) since 2007. Although the reference text showed no difference from those detailing the grammar of standard British English, it did, however, provide focus on areas of grammar which were of importance to the English language syllabus in Singapore. In coding the errors, it is also worth noting that there were no instances where there was dispute over grammaticality judgements among the four coders (three Singaporean and one Indian) attesting to Gupta's (2012) observation that in large part, grammatical variations across varieties of English around the globe are marginal.

In keeping with the aims of the project to generate findings and outcomes that would be useful to the improvement of pedagogical practices in the teaching of grammar in Singapore schools, the project team developed an error coding scheme that was in alignment with the Singapore MOE's English Language Syllabus 2010 (MOE, 2010). This was an important consideration given the purpose of the project was to lead to outcomes that could be used not only to inform language education policy decisions and future syllabus revisions, but also to enable the data from the corpus to be used by teachers in developing lesson materials. The error coding scheme was a hierarchical one with errors first categorised into major grammatical classes such as noun, verb and adjective, and then subcategorised into more specific error types. The coding scheme was comprehensive and comprised 110 different types of errors. The use of such a fine level of categorisation was to ensure that grammatical categories referred to in the English language syllabus were captured. In addition, differentiations in error types were also made in order to be able to code specific features of local Singapore English usage. Although the large number of error categories meant that coder reliability had to be carefully monitored, the research team felt it was important that the error categories captured details important to the local context. It was also felt that the errors could be regrouped into broader categories at a later stage where necessary.

The error coding scheme was piloted and refined through three rounds of coding. Training sessions for the coders were conducted on samples from the essay sets of 25 students (125 essays in total), and error tagging was checked in the major grammatical categories of noun, verb and sentence to ensure that coder reliability above 80% was achieved. To facilitate the team of four coders in the error tagging and cross-checking of error codes, a web-based coding system was also developed to provide easy access to the corpus so that the coders could not only work on their error annotations but could also check each other's coding in round-robin fashion and which also offered a means of comparing closely related error categories. Coding was checked by one other coder, and random audits of the coding were also made by the principal investigator. Issues and disagreements were raised and resolved in regular monthly meetings. Thus, rather than relying solely on statistical measures of inter-coder reliability, a more rigorous system of checks and counterchecks was used to ensure that errors were consistently tagged according to the error coding scheme.

Interpreting the findings

Of the 110 categories available to the coders, 82 different error types were found. This paper focuses on the ten most frequent error categories (Table 1) which, account for 21,664 out of a total of 35,926 errors (approximately 60%) . Interestingly, the error patterns were largely consistent in the student writing from Primary 2 to Primary 6.

Table 1: The ten most frequent error types in the learner corpus

Error Code	Error Description	Frequency	% of total errors
VDT.2	Inconsistent or incorrect marking of tense (past-present) on the verb at the text level.	4310	11.99
TSU.1	Misspelling of words (where the target word can be ascertained).	3773	10.50
TPU.3	Errors relating to the use of full stops.	2242	6.24
TPU.2	Errors relating to the use of commas.	2044	5.69
LLU.2	Non-conventional collocation of words.	1998	5.56
VDT.1	Tense (past-present) of the verb group of the embedded clause conflicts with that of the main clause verb.	1844	5.13
TPU.1	Errors related to the use of capital letters	1588	4.42
VFT.1	Incorrect marking of verb group for aspect	1492	4.15
TPU.5	Punctuation errors in relation to direct speech	1241	3.45
SEZ.3	Incorrect complex clause structure	1132	3.15
Total		21664	60.26

Across the corpus, verb errors (VDT.1, VDT.2, and VFT.1) accounted for three out of the ten most frequent errors and 21.27% of the total number of errors. All three of these error types relate to the use of tense and aspect. The second most common group comprises four punctuation errors (TPU.1, TPU.2, TPU.3, and TPU.5) and accounts for 19.8% of the total number of errors. The third group of errors (TSU.1) relates to spelling (10.5%). The remaining two errors in the top ten (LLU.2 and SEZ.3) relate to collocation (5.56%) and clause structure (3.15%) respectively. These error rates were calculated using a method commonly used in corpus based error analysis research which counts the number of errors of a specific type in relation to the total tokens in the corpus. However, as Thewissen (2013) suggests, a potential occasion analysis might be more suitable, especially for some types of errors because "it makes more sense to count errors in relation to the number of times a learner could potentially have committed such an error" (p.81).

The following sections will discuss the above groupings of errors and examine how interpretations of these errors are dependent on understanding Singapore's sociolinguistic context as well as its specific pedagogic practices and culture.

Verb errors

Verbs contribute a very large percentage of the number of errors in the corpus. This finding is resonant with past studies of Singapore classrooms; Sobrielo's study (1968), for example, also revealed verb errors to be the most frequent in Singapore secondary student writing. It is also pertinent to note that the major endogenous languages spoken in Singapore, namely Mandarin Chinese and Malay, do not mark verbs for tense morphologically, and neither does the local Singapore English vernacular, Singlish. As

such, verb forms have been a key focus in many sociolinguistically-oriented studies of English in Singapore (e.g., Ho, 2003). In the present corpus, the inconsistent use of tense forms at the text level (VDT.2) comprises almost 12% of the total number of errors. An additional 5.13% are related to inconsistencies in the tense of embedded clauses (VDT.1). These errors, if examined more closely, reveal a consistent pattern, and this suggests that the students are using tense in a systematic manner, although one that departs from what is expected in standard English. In particular, an examination distinguishes two possible patterns that might explain the use of the present tense in sentences like the following:

> The name is Alexa. (PP144, P5)
>
> He wanted to take the cookie jar but it is too high. (PP11, P3)
>
> As usual, I went to school and it seems that nobody remembered. (PP3, P6)

For many of the sentences, the switch from the expected past tense to the present is used to indicate that factual information is being reported. For example, students might see someone's name or the fact that the shelf is too high as factual information, particularly since the narratives are based on a description of the picture stimuli provided. This use of the present tense is consistent with the traditional grammar rule that most students are still taught in Singapore classrooms (i.e. that the simple present tense must be used when talking about facts).

Secondly, lexical aspect might provide another reason for the incorrect use of the present tense. Alsagoff, Yap, and Yip (2009), in their investigation of a small Singapore school corpus, show that the lexical aspect hypothesis correctly predicts that learners commit more errors in marking the past tense in non-telic verbs. More recently, Quek (2016) has also demonstrated a clear statistical correlation between telicity and the rate of past tense errors. Thus, copular verbs like *be* and *seem* may not be as consistently marked for the past tense as verbs which take a telic interpretation such as *walk*, *eat*, *take*, demonstrating that Singaporean children may be sensitive to the lexical meanings of the verbs in relation to how they denote the way an event unfolds. Such studies point to a need to further interrogate the present corpus through further coding and analysis.

In attempting to develop better strategies to teach tense, it is therefore important to recognise that apart from the expected problems such as students not knowing the actual past tense forms of certain verbs (including spelling errors) or when to use the past tense form, the lack of use of the past tense form in the narrative compositions might reflect that the students are guided by a different set of rules. Teaching students to mark tense and aspect more effectively may therefore require that students explore such internalised rules and be given opportunities to compare these with the rules of standard English in self-discovery learning exercises.

Punctuation and clause structure

The errors relating to punctuation point to the connectedness of grammar to punctuation. While some of the punctuation errors relating to the full stop and comma indicate a lack of awareness of language conventions, e.g. leaving a space between the last letter and the full stop, or not punctuating after an adverbial, for example:

> Wild thoughts went through my mind( ). (PP11, P3)
>
> After school( ) the sky started to be dark. (PP350, P2)

A significant number of these errors instead point to a lack of understanding of clause structure. Thus, sentences containing these errors, for example:

> I was strolling in the garden(,) it was a nice and cloudy day. (PP25, P5)
>
> When Sean was back in school(.) (PP30, P5)

were, in fact, either run-on sentences (i.e. sentences in which two or more independent clauses are incorrectly joined to form a single sentence) or sentence fragments (i.e. sentences which are not complete in their structure and cannot stand as independent clauses). Such sentences were also typically incorrectly marked for punctuation, where either a comma was used instead of a full stop, or where a full stop was used when there was an incomplete sentence[9]. Such punctuation errors therefore clearly indicate a problem with clause structure and are closely tied to clause structure errors (SEZ.3), such as:

> I immediately swim to the grass area just to prevent if they saw me. (PP32, P4)
>
> When I reached home, shocked to see my father's face as red as beetroot. (PP336, P5)

Some of the punctuation errors involving the incorrect use of capital letters may also be linked to a lack of mastery of sentence or clause structure. These are instances when students do not begin a sentence with a capital letter, for example:

> he tried to reach for the cookie jar. (PP350, P3)

The punctuation problems are not simply issues with the mechanics of grammar but belie more serious problems that students have in understanding how to form clauses and sentences in English. Consequently, a more fruitful approach to reducing punctuation errors may lie with the teaching of clause structures. It also suggests that teaching punctuation separately from grammar, as is routinely done in Singapore classrooms, may not be effective. This lack of linkage between punctuation and grammar might also explain why direct speech punctuation appears as one of the most frequent errors[10].

Spelling errors

Spelling is the second single most frequent error in the corpus (10.5%). Surprisingly, the frequency of spelling errors remained consistently high even at the upper primary levels of Primary 5 and Primary 6. Knowing that spelling was keenly emphasized in English language classrooms and the curriculum in Singapore (Saravanan, 2005) made this finding somewhat puzzling and prompted the research team to explore possible reasons for the spelling errors which fell into two distinct groups. The first group comprised errors that were quite clearly due to the inherently inconsistent spelling rules in English. The lack of a one-to-one sound-spelling correspondence means that homophones might be confused, e.g. *whether* spelt as *weather* (PP57, P4; PP128, P5), and that words such as *noncence* instead of *nonsense* (PP2, P5) and *reciving* instead of *receiving* (PP342, P6) are misspelt. However, apart from these spelling errors, the corpus reveals a significant percentage of spelling errors of a different nature. These appear to align with pronunciation patterns that have been documented of Singapore

English speech (e.g., Deterding & Hvitfeldt, 1994)[7], suggesting that a productive line of investigation may lie in examining the correlation between spelling and pronunciation.

Firstly, there are a number of spellings that indicate a conflation of a number of phonemes that are closely related. For example, the diphthong /oʊ/ and the back vowel /ɒ/ appear to be often confused, resulting in spelling of words such as *alone* as *along* (PP27, P4), *block* as *bloke* (PP46, P2)[8], *lorn* for *loan* (PP342, P5). Spellings such as *worfe* for *worth* (PP346, P6), *breeding* for *bleeding* (PP11, P3), *hungly* for *hungry* (PP14, P3), *lelise* for *realise* (PP345, P5), *feeled* for *filled* (PP16, P6); *dased* for *dashed* (PP336, P6), and *sellter* for *shelter* (PP36, P2) also point to such conflation of many other similar phonemes. Secondly, the spelling errors point to a well-attested area of difficulty in the articulation of consonant clusters, giving rise to a number of spellings where the consonant cluster is either reduced, or where a vowel is inserted, e.g., *clear* spelt as *cear* (PP3, P4), *confidence* spelt as *confidene* (PP11, P4), *gift* spelt as *gife* (PP337, P6), opponent spelt as *opponet* (PP16, P5), *hospital* spelt as *hosipital* (PP20, P3; PP81, P5).

This link between spelling and oral language is one that has been recognized in the research literature (e.g., McCarthy, Hogan, & Catts, 2012). Dockrell and Connelly (2009), for example, discuss evidence that oral language skills influence literacy at many levels, including at the word, sentence, and text levels. More specifically, Graham, Berninger, Abbott, Abbott, and Whitaker (1997) stress the importance of phonological processing on children's spelling development. Research on the language development of children with specific language impairment similarly shows that children with phonological difficulties also have greater problems with their spelling (Briscoe, Bishop, & Norbury, 2001).

These findings clearly suggest that the current practice in Singapore classrooms of having children memorise lists of words for written spelling tests is not effective, and that teachers might need to more actively draw students' attention to the relation between spelling and pronunciation. Spelling lists, which currently are primarily thematically oriented, might, for example, draw students' attention to phonological patterns. It would also be useful for teachers to be made aware that issues with spelling may originate from the specific local pronunciations of certain phonemes or sets of phonemes. In particular, introducing an awareness of Singlish pronunciations or those of the learners' home languages in teacher education courses may prove useful in creating a better understanding of recurrent patterns of misspelling.

Concluding remarks

The discussion of the error patterns demonstrates clearly how learner corpora can yield useful understandings of language pedagogy and practice and lead to the development of better approaches to the teaching of certain problematic areas of grammar. However, the discussion has also highlighted how for the data to yield such meaningful outcomes, interpretations need to be informed by contextual knowledge. Knowing about the sociolinguistic context of Singapore allowed understanding of the patterns of spelling errors, as stemming from Singaporean pronunciation patterns, as did knowledge of how spelling is currently taught in primary classrooms. Examining verb errors again required knowledge of the local context, the researchers needed to be aware not only of the influence of the endogenous languages spoken in Singapore, but also of how the teaching of the simple present tense at the sentential level might interfere with the students' use of the past tense at the textual level in narrative texts. Similarly,

punctuation errors could only be interpreted meaningfully through an understanding of how punctuation is taught in Singapore classrooms.

The study of grammatical error patterns in this study provides only the first step in exploring areas of difficulty of Singaporean school children in learning grammar and writing. It highlights the strong potential of learner corpora, but at the same time, reveals that the value of such corpora can only be realised if researchers are able to interpret the data meaningfully and in contextually appropriate ways. This also entails further investigation into other features that might provide a more holistic picture of language development.

Acknowledgements

This paper benefitted from funding from the Education Research Funding Programme at the National Institute of Education, Nanyang Technological University, Singapore. The author would like to express gratitude to Dr David Gardner and the two anonymous reviewers for their helpful comments and suggestions in the revision of this paper.

Notes

1. This paper presents some of the findings of a project *Investigating the development of the grammar of student writing in Singapore: Analysis of a longitudinal corpus of primary school student essays* for which the author was the Principal Investigator.
2. Further complicating the issue may be the fact that the English many Singaporeans speak at home and also possibly even in workplaces, is not standard Singapore English but the local vernacular, Singlish.
3. Although Singapore English is a much researched and documented variety of English in the world, it must be noted that most of the accounts of Singapore English are of its colloquial vernacular, Singlish, rather than its formal or standard variety.
4. Although the intention of the MOE was to track the same group of 351 students over their six years of education, attrition necessitated the recruitment of additional research participants for the study who were of similar demographic profiles.
5. A narrative is, in the Singapore context, typically a story that describes a series of events that includes a climax. Students are generally taught that narratives contain specific grammatical attributes as well as a particular text structure.
6. The pre-primary and Primary 1 compositions were excluded from a study of errors because these displayed too much variation and vagueness in clause and phrase structures to make error analysis feasible.
7. Generally, such pronunciation patterns are found among Singlish speakers with relatively low levels of competence in standard English, which would also include among them, learners.
8. PP refers to the student number, while P indicates the grade level, e.g., PP46, P2 refers to the Primary 2 composition of student PP46.
9. The coding of the errors was done in a comprehensive manner where all of the errors likely to be identified by teachers would be coded. This is in keeping with the aim of the study to allow the coded corpus to be useful to classroom teachers. Thus, the sentences above would have been coded both as having a clause error (SEZ.3) as well as a punctuation error (either TPU.2 or TPU.3, or both). Note, however, that some clause structure errors do not involve errors in punctuation, as seen in the sentences that follow.
10. Direct speech structures are commonly taught as an important feature of narrative texts in Singapore classrooms, perhaps pointing to why such structures appear so frequently in the corpus of Singapore student essays.

About the author

Lubna Alsagoff is a researcher interested in language variation, linguistic norms, and the interplay of language, identity and culture.

References

Alsagoff, L. (2007). *A visual grammar of English (2nd edition)*. Singapore: Pearson.

Alsagoff, L. (2010). English in Singapore: Culture, capital and identity in linguistic variation. *World Englishes, 29*(3), 336-348.

Alsagoff, L., Yap, D., & Yip, V. (2009). The development of the past tense in Singapore English. In R. Silver, C. C. M. Goh, & L. Alsagoff (Eds.), *Acquisition and development in new English contexts* (pp. 132-146). London: Continuum.

Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum.

Bernadini, S. (2004). Corpora in the classroom: An overview and reflections on future development. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15-38). Philadelphia, USA: John Benjamins.

Biber, D., & Reppen, R. (1998). Comparing native and learner perspectives on English grammar: A study of complement clauses. In S. Granger (Ed.), *Learner English on computer* (pp. 145-158). London: Longman.

Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, 42*(3), 329-340.

Conrad, S. (2004). Corpus linguistics, language variation, and language teaching. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 67-85). Philadelphia, USA: John Benjamins.

Conrad, S. (2010). What can a corpus tell us about grammar? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 227-240). Oxon, UK: Routledge.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly, 33*(2), 185-209.

Deterding, D., & Hvitfeldt, R. (1994). The features of Singapore English pronunciation: Implications for teachers. *Teaching and Learning, 15*(1), 98-107.

Dockrell, J. E., & Connelly, V. (2009). The impact of oral language skills on the production of written text. *Teaching and Learning Writing, British Journal of Educational Psychology Monograph Series II, 6*, 45-62. doi: 10.1348/000709909X421919

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89*(1), 170-182.

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast* (pp. 37-51). Lund, Sweden: Lund University Press.

Granger, S. (1999). Uses of tenses by advanced EFL learners: evidence from an error-tagged computer corpus. In H. H. S. Oksefjell (Ed.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 191-202). Amsterdam: Rodopi.

Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch–Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3-33). Philadelphia/Amsterdam: John Benjamins.

Granger, S. (2012). Learner corpora. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 3235-3242). Oxford, UK: Wiley-Blackwell.

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research, 1*(1), 7-24.

Gupta, A. F. (2012). Grammar teaching and standards. In L. Alsagoff, S. L. Mckay, G. Hu, & W. A. Renandya (Eds.), *Principles and practices for teaching English as an international language* (pp. 244-260). New York: Routledge.

Hasko, V. (2013). Capturing the dynamics of second language development via learner corpus research: A very long engagement. *The Modern Language Journal, 97*(1), 1-10.

Ho, M. L. (2003). Past tense marking in Singapore English. In D. Deterding, E. L. Low, & A. Brown (Eds.), *English in Singapore: Research on grammar* (pp. 39-47). Singapore: McGraw-Hill.

Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 293-313). Cambridge: Cambridge University Press.

Kachru, B. B. (1992). *The other tongue: English across cultures*. Champaign, IL: University of Illinois Press.

Kirkpatrick, A., & Sussex, R. (Eds.). (2012). *English as an international language in Asia: Implications for language education*. Dordrecht; New York: Springer.

McCarthy, J. H., Hogan, T. P., & Catts, H. W. (2012). Is weak oral language associated with poor spelling in school-age children with specific language impairment, dyslexia or both? *Clinical Linguistics & Phonetics, 26*(9), 791-805.

Mindt, D. (1997). Corpora and the teaching of English in Germany. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 40-50). London: Longman.

MOE, [Ministry of Education]. (2010). *English language syllabus 2010: Primary & secondary (express/ normal [academic])*. Singapore: Curriculum Planning and Development Division.

Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In S. J. McHardy (Ed.), *How to use corpora in language teaching* (pp. 125-152). Philadelphia, USA: John Benjamins.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: Benjamins.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.

Pérez-Paredes, P., & Sánchez-Tornel, M. (2014). Adverb use and language proficiency in young learners' writing. *International Journal of Corpus Linguistics, 19*(2), 178-200.

Quek, S. Y. S. (2016). *Analysis of patterns of past tense errors in a corpus of Singapore students' narrative writing: Insights on language acquisition.* (Unpublished Ph.D. Thesis), National Institute of Education, Nanyang Technological University, Singapore.

Römer, U. (2005). *Progressives, patterns, pedagogy: A corpus-driven approach to English progressive forms, functions, contexts and didactics*. Amsterdam: John Benjamins.

Saravanan, V. (2005). 'Thinking schools, learning nations' implementation of curriculum review in Singapore. *Educational Research for Policy and Practice, 4*(2-3), 97-113.

Schneider, E. (2007). *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.

Sinclair, J. M. (2004). *How to use corpora in language teaching*. Philadelphia: John Benjamins.

Sobrielo, A. (1968). *An error analysis of the spoken English of Singapore pupils for whom English is a second language*. Singapore: TTC Publications Board, Teachers' Training College.

Stewart, D., Bernardini, S., & Aston, G. (2004). Introduction: Ten years of TaLC. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 1-18). Philadelphia, USA: John Benjamins.

Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *The Modern Language Journal, 97*(1), 77-101.

Urzúa, A. (2015). Corpora, context, and language teachers: Teacher involvement in a local learner corpus project. In V. Cortes & E. Csomay (Eds.), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber* (pp. 99-122). Philadelphia: John Benjamins.