MSU Working Papers in Second Language

Studies Volume 1, Number 1

2009

A Corpus-Based Study of Adverbial Connectors in Learner Text

Mark Shea
PhD Student, Second Language Studies Program
Michigan State University
sheamark@msu.edu

Introduction

Many teachers of English as a second or foreign language, particularly those who have spent time teaching writing, have anecdotal evidence that their students tend to misuse and overuse adverbial connectors (e.g., *however* and *first*), and two decades of work on writing pedagogy has consistently identified these connectors as a source of difficulty for second language writing (e.g., Crewe, 1990; Hinkel, 2004). Computer text-processing capabilities have provided an opportunity to investigate large samples of learner writing through corpus-based research, and such research has compared learner-created texts with those written by native speakers in an attempt to corroborate these intuitions (e.g., Chen, 2006; Granger & Tyson, 1996). The results of this line of research indicate that there are differences between NS and NNS adverbial connector use in writing, but a consistent pattern has not been observed across these studies. Although many studies found that particular connectors are used more or less often in NNS writing than in NS writing (e.g., Granger & Tyson, 1996), no overall pattern of over- or underuse was found.

Even when a study presents clear results, whether it be that a population of learners overuses connectors in general or misuses a particular connector relative to NS samples, the practical significance of these findings is not clear. For example, Granger and Tyson (1996) found a pattern of NNS overuse for particular connectors, but the connector that exhibited the largest difference between NSs and NNSs, *moreover*, had approximately six more tokens per 10,000 words. In corpus linguistics, the term "overuse" has a technical meaning, devoid of the connotations signaling a problem which is associated with its everyday usage. And although research into the characteristics of learner text is valid in its own right, research which aims to inform pedagogy should look to move beyond designations of the relative frequencies of lexical items in texts and begin to address the question of whether these differences have a substantive effect on a learner's writing.

This point is related to what the goal of writing instruction should be; is the goal to make learners' writing as similar to NSs' as possible, or is it to help them write effectively? Corpus-based studies of adverbial connectors in learner writing have been conducted for over fifteen years now, and it is perhaps time to move forward. To do so, it would be desirable to know whether the use of these connectors correlates with measures of the quality of a text, or, considered another way, whether the use of these connectors correlates with indicators of development in learner writing. In addition to practical concerns, there should also be theoretical motivation for the study of adverbial connectors. The field of text linguistics, particularly as related to the creation of cohesion and coherence within a text, suggests that conjunction, represented

in part by the use of adverbial connectors, is one way to create a more unified text.

The present study, then, represents an initial step toward the expansion of corpus-based research on the use of conjunctive adverbials in learner writing. These connectors were extracted from a corpus of timed learner writing. The number of T-units per total number of words in each text was also calculated in order to provide objective measures of complexity and writing development. In addition, the texts were coded using an analytic scale rating vocabulary, content, language use, organization, and mechanics. This method had the advantage of allowing rater perceptions of relevant aspects of the text, such as organization and content scores, to be isolated while also providing an indication of the overall text quality through the total score for all traits.

The combination of textual data from a large sample of learner writing with information on how those texts are perceived offers the opportunity to determine whether the use of these overt cohesive devices does contribute to the perceived quality of a text. If the use of adverbial connectors correlates with measures of text quality, this suggests that they may indeed help to increase the quality of texts. If no correlation is found, then this validates intuitions that such words do not benefit student writing (e.g., Crewe, 1990; Hinkel, 2004) and lends support to claims by these writing researchers that other methods of increasing cohesion in writing should be emphasized in writing pedagogy.

Review of Literature
Cohesion

The concept of text cohesion was first developed in Halliday and Hasan's (1976) seminal work on the topic. Examining what quality makes a series of sentences cohere into a single text, Halliday and Hasan identified five cohesive relations that can signal relationships between texts. Three, reference, ellipsis, and substitution, make use of syntactic operations and closed-class words, creating cohesion through the fact that their presence in a sentence presupposes the existence of an element in another sentence; the use of a pronoun, for example, presupposes the existence of its referent elsewhere in the text.[1] The fourth type of cohesive relation is conjunction, which makes use of elements such as coordinating and subordinating conjunctions, as well as conjunctive adverbials to make explicit connections between propositions. Conjunction has a larger lexical element than the preceding three, in that it makes use of a wider set of lexical items, but is also applied with a degree of systematicity that indicates the incorporation of grammatical aspects as well. The fifth type of cohesive relation is lexical cohesion, the repetition of lexical items or use of synonymous items throughout various sections of a text.

A division can be made between Halliday and Hasan's (1976) first three categories of cohesive relation and the second two, as the first three deal with closed-class word categories such as personal and demonstrative pronouns. The conjunctive cohesive devices, while still a closed set, represent a much wider variety of single and multiword expressions, and lexical cohesion can be created by many types of words or phrases. In another sense, the conjunctive cohesive relation stands apart from the other four, in that it does not connect to a second element elsewhere in the text but rather makes explicit a relationship between two propositions. Seen from these two perspectives, conjunctive adverbials stand apart from the others, in terms of their status as a relatively large set of fixed elements and their function within a text.

Corpus Studies of L2 Conjunctive Adverbial Use

Following Halliday and Hasan's (1976) work, a number of studies were done in the 1980s examining the use of cohesive devices in learner writing. These initial studies were inconclusive, but they were done without the benefit of computers capable of processing large amounts of text and so had extremely small sample sizes. For example, Connor (1984, as cited in Tyson & Granger, 1996) used only six learner essays for her analysis of cohesion in ESL writing. In the 1990s researchers began examining conjunctive adverbials using learner corpora. Milton and Tsang (1993) published one of the first of these studies, using a corpus gathered from English learners in Hong Kong. Of the 25 logical connectors included in their analysis, they found that 20 were "overused", contributing to an overall pattern of overuse.

However, Milton and Tsang (1993) was by the authors' admission a first step, and there were several aspects of the analysis which bring the reliability of the results into question. First, in determining whether the learner papers overused adverbial connectors, the student texts were not compared to similar NS texts; Milton and Tsang attempted to compensate by drawing from various NS corpora, but none of the NS texts were produced by students or represented similar text genres. Second, a set of 25 logical connectors were selected, whereas the potential number of connector types is far higher; for example, Granger and Tyson (1996) included 108 types in a similar study. Finally, working from an untagged corpus, Milton and Tsang did not identify whether a token appeared in the correct context; that is, as a connector or otherwise. This is of particular importance as the study included connector types such as *and* and *also*, which could easily have appeared in numerous contexts in which they were not

serving as cohesive ties between sentences.

Granger and Tyson (1996) performed a similar study using the International Corpus of Learner English (ICLE). The study compared French first language (L1) English learners' use of conjunctive adverbials, rather than the full range of connectors which includes coordinating and subordinating conjunctions, with a corpus of NS writing. This study was well controlled in terms of the source texts; the NS corpus was the Louvain Corpus of Native Essay Writing (LOCNESS), created as a complement to the texts in the ICLE. Granger and Tyson hypothesized that a general pattern of overuse would be found in the NNS writing, but the results of the analysis were not conclusive. An analysis of individual connectors indicated that particular classes of connector, such as appositional (e.g., *in fact*) and additive (*moreover*), were overused, whereas contrastive (*however*) and argument advancing (*therefore*) connectors were underused. A comparison of the French writing with writing by German learners of English suggested that some, but not all, of the usage patterns were the result of L1 discourse conventions and translation equivalents. Altenberg and Tapper (1996) compared their analysis of the writing of Swedish L1 English learners to the results reported by Granger and Tyson and found that the Swedish writers used far fewer conjunctive adverbials overall due to transfer effects, but with a usage pattern similar to the French L1 learners.

Narita, Sato, and Sugiura (2004) used the Japanese subcorpora of the ICLE and the US subcorpora of the LOCNESS to investigate the usage of connectors in the second language (L2) English writing of Japanese L1 learners. They found a statistically significant overuse of connectors, but it is again questionable as to how relevant this finding is because the

difference represents a total difference of 14 tokens per 10,000 words. Narita et al. also found, similarly to Granger and Tyson, that some connectors were used more by the learners while others were used less often. A comparison of Narita et al.'s results with those obtained from analyses of the French, Swedish, and Chinese subcorpora revealed limited similarities; for example, *moreover* was overused by all four groups, but was the only connector which appeared on all four lists.

Chen (2006) compared the adverbial use in a corpus of 23 final papers written by Taiwanese MA TESOL students to that in research articles in TESOL related journals. Based on the argument that conjunctive adverbials connect sentences rather than words, Chen calculated the number of conjunctive adverbials both per word and per sentence and found that learners used slightly more connectors than in the NS sample when calculated as a percentage of total words, but not when calculated as a percentage of total sentences. Chen hypothesized that this result is due to the fact that the NS authors' more complex sentences require the additional structure provided by adverbial connectors.

Chen's (2006) NNS sample of 23 texts was made up of varying numbers of five different text genres, and the comparison corpus consisted of 10 published journal articles. Chen acknowledged the small and heterogeneous sample weakened her results, but the study also illustrated a deeper question of what implications a finding of "overuse" might have for informing writing pedagogy and language development. Chen pointed out that per 10,000 words, the learners used only 6 more conjunctive adverbials than the NS sample, though particular adverbials demonstrated more dramatic differences. One question, then, is what should be considered a meaningful difference. Also, if

adverbial use in the NNS sample differed from that in the NS sample only when measured per word and not per sentence, then it is not clear that the difference truly lies in the two groups' use of adverbials. Rather, the NNSs wrote shorter sentences which they connected in ways similar to NS writers; as their writing fluency increases, it might catch up to their conjunctive adverbial use.

Further bearing on the question of what constitutes a meaningful finding of overuse of adverbial connectors, Bolton, Nelson, and Hung's (2002) study compared the use of connectors in the Hong Kong and Great Britain subsets of the International Corpus of English (ICE) to a sample of published academic writing. The results of the analysis demonstrated a tendency for both groups of students, NNSs and NSs, to overuse connectors when compared to published academic writing. In light of such results, the question of what represents significant overuse becomes a central one to interpreting the findings of corpus-based research.

Research Questions

The present study addressed the question of learner use of conjunctive adverbials using an analysis that has the potential to provide results more directly relevant to writing pedagogy. Instead of comparing NNS patterns of adverbial use to those found in an NS corpus, a correlation analysis was performed between NNS's usage of conjunctive adverbials and the scores given to the texts by two raters. Rather than describing a pattern of usage in terms of a comparison corpus, the results of this analysis will provide insight into the effect that the use of these connectives has on readers' perceptions of learners' writing.

Method
The Corpus

The corpus was compiled using 30-minute timed essays written by NNS students in the Intensive English Program at a large Midwestern university. The learners were in their second to fourth semesters, at varying points in the intermediate band of proficiency as determined by their placement within that program. The essays were written on a variety of topics, all designed to elicit descriptive writing. For the purposes of a previous study, the essays had been coded by two raters for language use, vocabulary, organization, content, and mechanics, and the subscores were combined for a total score. The raters were both English NSs with extensive teaching experience. One was a Faculty member in a MA TESOL graduate program, and the other was a PhD candidate. The means of the two raters' organization, content, and total scores were used in the analysis; the two subscores were considered those most likely to reflect cohesion within the essay, whereas the total score allowed all the subscores to be counted in the analysis. For the measures of organization, content, and total scores the respective Pearson's coefficients were r = .71; .83; .88 with $p <$ .01 for all, indicating relatively good interrater reliability.

In the present study, 100 essays with a total of 33,147 words were included in the analysis. These 100 essays are part of a larger corpus and were selected at random; the choice to work with a subset of the larger corpus was made to make the initial analysis more manageable in terms of person-hours. The larger corpus contains multiple entries from each writer; the subset does not control for this fact. This means that the corpus as it is reported is not balanced; certain writers contributed more. As work on the corpus progresses and the rest of the texts are entered, the contributions of each writer will balance. As the corpus stands currently, the results presented here should be considered indicative of future findings but interpreted with caution.

Procedure

Because of the difficulties in using sentences to analyze learner writing (c.f., Bolton et al., 2002; Chen, 2006), the corpus was manually divided into T-units according to the guidelines described by Polio (1997, Appendix C). As it was not predictable which of the extensive set of connector types might appear in the corpus, all tokens were extracted by hand. The list of connectors provided by Celce-Murcia and Larsen-Freeman (1999, p. 530) was used as a starting point, and other connectors were added to the list when encountered in the text. Adverbial connectors were identified if they were (a) adverbs that (b) modified the T-unit as a whole and (c) modified it in relation to a preceding or following T-unit. For example, *consequently* and *unfortunately*, would each fit the first two criteria, but whereas *consequently* indicates that one proposition is caused by another, *unfortunately* comments only on the proposition to which it is attached and so would not be included. Although included in some early studies on the topic (e.g., Milton & Tsang, 1993), coordinating and subordinating conjunctions are generally not included in corpus analyses of connector usage and were not included in the study. In addition, connectors were only counted if they were used to create links between T-units. For example, in the case of: *I went to the store. I also stopped at the gas station*, the connector *also* would be coded as a conjunctive adverbial, whereas in the following example, it would not: *I stopped by the store and also the gas station.* It was decided to use a more conservative criterion for counting connectors, as interT-unit

bonds seem more relevant to analyses presented in the theoretical literature on cohesion as well as to anecdotal reports of learner use of connectors. Table 1 presents the descriptive statistics for the essays comprising the corpus. As can be seen from these statistics, there was a great deal of variation in the amount and complexity of the writing. Quintero-Wolf, Inagaki, and Kim (1998, as cited in Reynolds, 2001) have indicated that number of words per text and words per T-unit are the two most reliable measures of writing development (particularly total number of words when dealing with timed writing). This variation in indicators of development provides an opportunity to investigate the data for correlations with connector use; although such an analysis cannot determine causation, it can at least determine whether connector use co-occurs with established measures of writing development.

Analysis

Table 2 presents a list of the most common adverbial connectors extracted from the corpus (see Appendix A for a full list), along with raw frequency, frequency per 1,000 words, and the percentage of all connectors each represents.

While direct comparisons to the results of studies using different corpora as data are difficult, there are contrasts between these and previous findings. *However* and *therefore*, for example, were found to be underused by the learners in Granger and Tyson's (1996) study. They were among the most common in the present study and used with greater frequency than in Granger and Tyson's NS sample.

Kolmogorov-Smirnov tests indicated that the data were normally distributed, so

Table 1 *Descriptive Statistics for Corpus Essays (N = 100)*

	Minimum	Maximum	Mean	*SD*
Words per essay	139	645	331.5	91.5
Words per T-unit	7.2	21	10.9	2.1
Conjunctive adverbials	0	16	4.3	3.4
Conj. Adverbials per word^2	.00	.04	.01	.01
Connectors per T-unit	.00	.56	.14	.11

Table 2 *Most Common Adverbial Connectors in Corpus*

Connector	Raw Frequency	Frequency/1000 Words	% of All Connectors
however	78	25.27	0.17
also	61	19.76	0.13
first of all	28	9.07	0.06
for example	28	9.07	0.06
therefore	27	8.75	0.06
finally	24	7.77	0.05
in addition	22	7.13	0.05
secondly	20	6.48	0.04
in conclusion	16	5.18	0.03

Pearson's correlation was used to investigate the relationships between rater scores and the number of cohesive adverbials in the text. Because the raw frequency of adverbial connectors would likely vary as a function of total words, the variables *connectors per number of words per text* (CON/W) and *connectors per number of T-units per text* (CON/T) were included in addition to number of words. As discussed above, Bolton et al. (2002) argued that the sentence, rather than the word, should be the basic unit of analysis when analyzing cohesive ties. Chen (2006) found different results for the two types of analysis, so both were included in the present study, although sentences were replaced by T-units as a more suitable unit of analysis for learner writing.

The mean of the two raters' scores for the subscores (language use, vocabulary, organization, and content), as well as the total score, were included in the analysis. It was predicted that, as the organization and content subscores might reflect the same aspects of writing proficiency that drive the use of connectors, significant correlations might be found between those subscores and measures of connector use. Additionally, number of words and number of words per T-unit were included as text-based measures of writing development.

Results

Table 3 reports the correlations of adverbial use with text-internal (words/T-unit) and text-external (rater evaluations) measures of writing. The results of the correlation analysis indicated that there were few significant relationships between CON/W or CON/T and text or rater-based measures of writing development. The raw number of adverbials in a text correlated significantly with the organization subscore, though accounted for only 5% of the variance. This would seem to support the current view of the role of connectors: Connector usage is related to text organization, but the relationship is not a particularly strong one (e.g., Hinkel, 2004). It is also possible that the relationship between number of connectors and the organization subscore is an artifact of the stronger correlation between number of words and each subscore.

A second measure of connector use that correlated with a text development measure was CON/T and words per T-unit, which is relatively uninteresting as it tells us little beyond the fact that a writer who uses more words per T-unit is likely to use more of a particular type of word per T-unit.

Table 3 *Correlations of Adverbial Use and Text-Internal and External Measures of Writing Development*

	1	2	3	4	5	6	7	8	9
1. Raw Connector	-	.92**	.90**	.29**	.12	.20	.24*	.07	.02
2. CON/W		-	.97**	-.04	.05	-.02	.09	-.07	-.11
3. CON/T			-	.00	.25*	.01	.13	-.03	-.05
4. Total Words				-	.27**	.68**	.56**	.47**	.42**
5. Words/T-unit					-	.26*	.32**	.32**	.35**
6. Content						-	.89**	.79**	.79**
7. Organization							-	.78**	.81**
8. Vocabulary								-	.86**
9. Language									-

*p < .05, **p < .001

Discussion

The results of this study indicate that in timed writings on descriptive topics, intermediate learners of English do not use conjunctive adverbials in a pattern that relates to their writing development as measured by text-internal or external measures. It suggests that the use of connective adverbials should perhaps not be a priority in writing instruction. This does not mean that learners never misuse these adverbials, for they certainly do. Nor does it mean that they use conjunctive adverbials in similar distributions as NS writers; previous research, though inconsistent, generally shows different patterns of use. What these findings do suggest is that how learners use these connectors may have little to do with how their writing is perceived by readers. The findings also suggest that the use of conjunctive adverbials does not follow the general trend of language development, as usage did not correlate with two reliable measures of writing development: total words written and words per T-unit.

Moving forward, there are two main directions that future research on cohesion in L2 writing can take; the first is to develop a more sophisticated analysis of the use of conjunctive adverbials, and the second is to move beyond conjunctive adverbials to examine other sources of cohesion in writing. Both directions will briefly be discussed in the following sections.

Analyzing Adverbials

The present study reported the number and type of connectors used in learner writing and whether the number of connectors differed in relation to the level of the writing. It did not examine whether texts displaying different levels of development also displayed different patterns regarding the use of particular classes of connectors, which some research has suggested may be a relevant way to approach the topic. The present study also did not take the misuse of conjunctive adverbials into account. It is possible that extracting this information from the corpus would, in fact, provide some insight into how learners deploy connectors in their writing. It could be that better-developed texts employ connectors which express more complicated relationships between propositions, such as causality or dismissal, than less well-developed texts, which may, for example, rely more heavily on sequential connectors such as *first* and *thirdly*, which express simpler relationships between propositions.

If it were the case that variables such as class of adverbial used or rate of misuse correlated positively or negatively with measures of writing development, such a finding would potentially be of interest to the field. However, there are several reasons why the implications of such correlations would be less clear than it might seem. In the case of classes of connectors, the results of many of these studies, and much of the criticism of existing pedagogical methods for presenting them to students, focus on the fact that individual connectors carry slightly different semantic meanings, collocational restrictions, and register. Pedagogically minded writers criticize the common practice in textbooks of presenting a variety of connectors grouped in boxes, categorized by a particular function, such as *contrastive*, but with an example provided for only one or two of the items (e.g., Hinkel 2004). It could be argued that research which seeks to inform pedagogy should avoid recycling a practice which treats a variety of connectors as members of a single, unified class.

The second possibility, examining learner misuse of connectors, carries its own complications. Chief among these is the difficulty in identifying what should be classified as a mistake. Milton and Tsang (1993), for example, classified problematic usages as misuses and overuses. Misuses,

such as using an adverbial of causality when one is not warranted, are relatively straightforward to identify. Overuses, such as using connectors when they are redundant and "do not contribute to meaning," would clearly be much harder to objectively code. Yet it is the overuse, rather than the misuse, that gets the focus in much of the pedagogical literature's discussion of adverbial usage (e.g. Hinkel, 2004).

Despite the difficulties described, knowing how effectively adverbial connectors are used in a text would certainly present a clearer picture of their role in reader perceptions of NNS writing, and future research should seek to develop a methodological framework to examine these questions. It may also be the case that adverbial connectors are distributed differently with texts written by learners of different levels. For example, a less-skilled writer might use adverbials as paragraph transitions while a more-skilled writer uses them to connect propositions within paragraphs. The use of corpus data to create a more detailed picture of how these connectors are used is another potentially informative direction for future research.

Conclusion

An analysis of connective adverbials used in NNS English learners' writing revealed little correlation between the use of these cohesive devices and raters' perceptions of the texts. Additionally, the use of connective adverbials did not correlate with measures of writing development. This finding supports arguments in the pedagogical literature that these connective adverbials do not contribute to the overall quality of a text (e.g., Hinkel 2004). Future research will seek to develop a more detailed picture of how connective adverbials are deployed in texts.

References

Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). Harlow: Addison Wesley Longman Limited.

Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from the International Corpus in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics, 7*(2), 165-182.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*. United States: Heinle & Heinle Publishers.

Chen, C. W. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics, 11*(1), 113-130.

Connor, U. (1985). A study of cohesion and coherence in English as a second language students' writing. *Papers in Linguistics, 17*(3), 301-316.

Crewe, W. J. (1990). The illogic of logical connectives. *ELT Journal, 44*(4), 316-325.

Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes, 15*(1), 17-27.

Halliday, M. A. K., & Hassan, R. (1976). *Cohesion in English*. New York: Longman.

Hasan, R. (1984). Coherence and cohesive harmony. In J. Flood (Ed.), *Understanding reading comprehension* (pp. 181-219). Newark: International Reading Association.

Hinkel, E. (2004). *Teaching acdemic ESL writing: Practical techniques in vocabulary and grammar*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hooey, M. (1991). *Patterns of lexis in text*. New York: Oxford University Press.

Milton, J., & Tsang, E. S. C. (1993). A corpus-based study of logical connectors in EFL students' writing: Directions for future research. In R. Pemberton & E. S. C. Tsang (Eds.), *Lexis in studies* (pp. 215-246). Hong Kong: Hong Kong University Press.

Narita, M., Sato, C., & Sugiura, M. (2004). *Connector usage in the English essay writing of Japanese EFL learners.* Paper presented at the Fourth International Conference on Language Resources and Evaluation.

Reynolds, D. (2001). Language in the balance: lexical repetition as a function of topic, cultural background, and writing development. *Language Learning, 51*(3), 437-436.

APPENDIX A
CONNECTORS INCLUDED IN THE ANALYSIS

Connector	Raw Frequency	Freq./1000 Words	% of All Connectors
however	78	25.27	0.17
also	61	19.76	0.13
first of all	28	9.07	0.06
for example	28	9.07	0.06
therefore	27	8.75	0.06
finally	24	7.77	0.05
in addition	22	7.13	0.05
secondly	20	6.48	0.04
in conclusion	16	5.18	0.03
actually	13	4.21	0.03
first	11	3.56	0.02
second	10	3.24	0.02
then (temp)	10	3.24	0.02
In my opinion	10	3.24	0.02
especially	10	3.24	0.02
as a result	8	2.59	0.02
thirdly	8	2.59	0.02
first (temp)	6	1.94	0.01
furthermore	6	1.94	0.01
on the other hand	6	1.94	0.01
third	6	1.94	0.01
anyway	5	1.62	0.01
lastly	5	1.62	0.01
moreover	5	1.62	0.01
thus	5	1.62	0.01
at first (meaning first)	4	1.30	0.01
in fact	4	1.30	0.01
first of all (temporal)	3	0.97	0.01
then	3	0.97	0.01
at the same time	2	0.65	0.00
besides	2	0.65	0.00
firstly	2	0.65	0.00
last (temporal)	2	0.65	0.00
next	2	0.65	0.00
according to this	1	0.32	0.00
fourth	1	0.32	0.00
at the same time (temporal)	1	0.32	0.00
consequently	1	0.32	0.00
like (for example)	1	0.32	0.00
for instance	1	0.32	0.00
for that reason	1	0.32	0.00
in the meantime	1	0.32	0.00

instead	1	0.32	0.00
last	1	0.32	0.00
on the contrary	1	0.32	0.00
second (temp)	1	0.32	0.00
secondly (tem)	1	0.32	0.00
to conclude	1	0.32	0.00
after all	1	0.32	0.00

No Instances Found

additionally	in consequence	nonetheless
all in all	in contrast	otherwise
anyhow	in fact	overall
as a consequence	in other words	rather
at any rate	in short	similarly
at least	in sum	that is
by contrast	in that case	that is to say
conversely	in turn	thereby
despite this	in turn	to sum up
fifth	initially	to summarize
further	in summary	
hence	later	
in any case	likewise	
in any event	meanwhile	
in brief	nevertheless	

Endnotes

[1] The use of 1st and 2nd person pronouns does not presuppose the existence of a referent within the text, but rather a referent in "the world." Thus, 1st and 2nd person pronouns are regularly excluded from counts of cohesive elements within texts.

[2] A reviewer points out that many adverbials are multiword constructions, making the calculation of connectors per word problematic. Although the present study uses this statistic as per previous studies (e.g., Chen, 2006), the question does seem to require a more elegant solution.

Accounting for L2-English Learners' Article Choices

Benjamin White
PhD Student, Second Language Studies Program
Michigan State University
whitebe5@msu.edu

Introduction

Teachers of English as a Second Language (ESL) are well aware that their students often fail to achieve native-like use of English articles. Even learners at higher levels of proficiency continue to make errors (Kharma, 1981; Master, 1997). They may substitute one article for another (e.g., *the* for *a*) or may omit articles entirely (e.g., *book* instead of *the book*). Research on ESL writing has found inaccurate article use to be one of the most frequent errors committed (Bardovi-Harlig & Bofman, 1989; Bitchener, Young, & Cameron, 2005). Learning to use articles appropriately is especially challenging for native (L1) speakers of article-less languages (Master, 1987). Whereas Spanish or German speakers may be learning the nuances that distinguish the uses of articles in English from those in their L1s, Korean and Russian speakers must learn this system without reference to comparable linguistic items in their own languages.

That second language (L2) English learners make errors with articles is unsurprising given the complexity of the English article system, which requires a number of factors be considered. Celce-Murcia and Larsen-Freeman (1999) stressed the importance of noun classification for article choice. As seen in Figure 1, appropriate article decisions are dependent on classifications of common versus proper[1], count versus noncount, and singular versus plural.

Not only must learners determine the count and number status of a noun, they must also distinguish definiteness and indefiniteness as applied to noun phrase (NP) reference. Regarding reference, Bickerton (1981) proposed two universals: a semantic universal that differentiates a
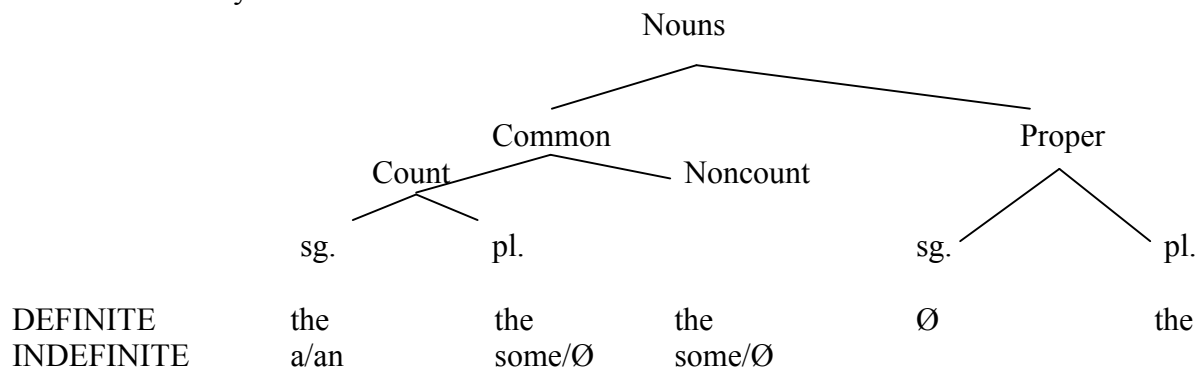

	sg.	pl.	sg.	sg.	pl.
DEFINITE	the	the	the	Ø	the
INDEFINITE	a/an	some/Ø	some/Ø		

Figure 1. Celce-Murcia and Larsen-Freeman's (1999) noun classification for article use (p. 272).

specific referent [+SR] from a nonspecific referent [-SR], and a discourse universal that differentiates a referent that is assumed known to the hearer [+HK] and a referent assumed unknown to the hearer [-HK]. There are thus four possibilities for NP reference as indicated in Huebner's (1983) semantic wheel (see Figure 2).

English divides the above semantic fields in the following ways: *the* for [+SR,+HK]; *a* or Ø for [+SR,-HK]; and *a* or Ø for [-SR,-HK]. Underscoring the complexity of the article system, generic reference (or [-SR,+HK]) may take *the*, *a*, or Ø². Researchers have found it necessary to include a fifth category of idiomatic use (Butler, 2002; Thomas, 1989) as there remain article uses (e.g., *all of a sudden* and *living hand to mouth*) which Bickerton's universals seem unable to account for. The idiomatic category, like generics, contains all three article choices: *the*, *a*, and Ø.

This paper investigates how advanced ESL learners with article-less L1s interpret (and misinterpret) English articles. It contributes to the body of previous research on article acquisition by placing particular focus on how learners' determination of countability influences their article choices within different semantic contexts.

Literature Review

Much research on the acquisition and use of English articles by L1 speakers of article-less languages has focused on the semantic context of NP reference. In particular, studies have investigated learners' sensitivity to the features of definiteness and specificity. Fewer studies have considered the role of countability on learners' use of articles. This section briefly reviews some of the major findings from research undertaken with these foci (semantic context and noun countability).
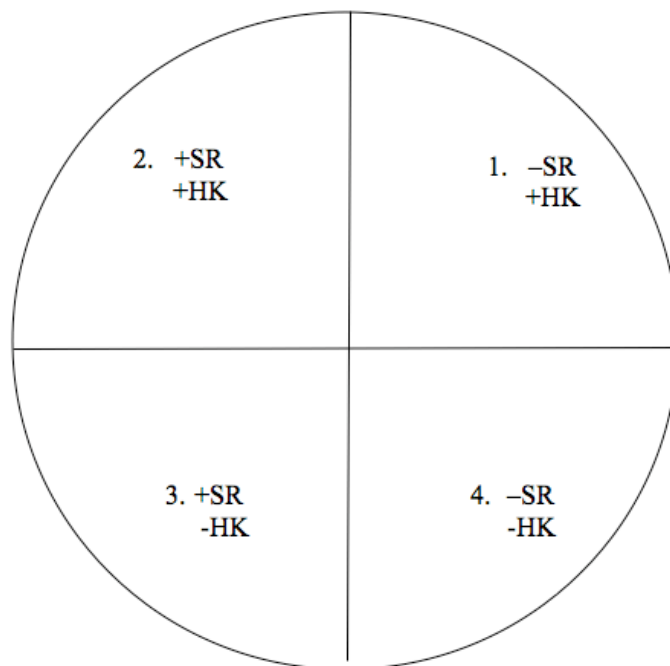

Figure 2.
Huebner's (1983) semantic wheel (p. 133).

For the analysis of NP reference, a number of studies have made use of Bickerton's (1981) semantic universals (HK and SR). A common finding is *the*-flooding by learners at lower levels of English proficiency (Chaudron & Parker, 1990; Huebner, 1983; Master, 1987; Young, 1996). That is, learners overgeneralize use of the definite article, especially in [-HK, +SR] contexts. As proficiency increases, the choice of *the* in this semantic context becomes less likely. Whereas Huebner (1983) and Master (1987) argued that participants in their studies were correctly associating *the* with definiteness (i.e., [+HK]), other researchers suggested that learners may incorrectly associate *the* with specificity (i.e., [+SR]) (Chaudron & Parker, 1990; Lu, 2001; Parrish, 1987; Tarone & Parrish, 1988; Thomas, 1989). Examining the metalinguistic knowledge

behind nonnative speakers' (NNS) article use, Butler (2002) found that L1-Japanese speakers at various levels of English proficiency exhibited problems correctly detecting both HK and SR in explanations for article choices on a gap-fill task.

That accuracy rates for indefinite NP contexts (i.e., [-HK]) lag behind those for definite contexts has been interpreted to mean definiteness is encoded before indefiniteness (Chaudron & Parker, 1990). As Lardiere (2005) pointed out, article choices in indefinite contexts are more complex decisions because they involve the determination of number and countability (see Figure 1). Learners appear to first use Ø in [-HK] contexts and only later begin to use *a* as they gain an understanding of countability (Chaudron & Parker, 1990; Master, 1987).

Recent studies have replaced the classification scheme of HK and SR with one of definiteness and specificity. Ionin, Ko, and Wexler (2004; henceforth referred to as IKW) defined the latter terms as follows: *definiteness* occurs when "the speaker and hearer presuppose the existence of a unique individual in the set denoted by the NP" and *specificity* occurs when "the speaker intends to refer to a unique individual in the set denoted by the NP and considers this individual to possess some noteworthy property" (p. 5). It has been claimed that article systems, across languages, encode either definiteness or specificity (Lyons, 1999). In English, it is definiteness that is marked. The following sentences illustrate this.

(1) Let's take *the new car* and leave the old one behind.
(2) I am planning to buy *a new car*. It's candy apple red and gets decent gas mileage.
(3) I am planning to buy *a new car*. Any car will do.

Because the hearer is assumed to know the unique car to which the speaker is referring, *new car* is definite in (1) and is preceded by the definite article. However, such a presupposition of uniqueness fails for the hearer in (2) and (3), resulting in use of the indefinite article. Note that *new car* in (1) and (2) is specific, but it is nonspecific in (3). It is, thus, possible to classify the above italicized NPs as: [+definite, +specific] in (1); [-definite, +specific] in (2); and [-definite, -specific] in (3). L1 speakers of article-less languages must determine the proper encoding feature for articles in the L2. IKW proposed the Fluctuation Hypothesis (FH), whereby learners will sometimes rely on definiteness and at other times on specificity until they settle on the correct feature.

To test this hypothesis, IKW asked L1 Russian and L1 Korean speakers to produce short written texts and to complete a multiple-choice elicitation task. On the elicitation task, the adult participants read short dialogues and chose between *a*, *the*, and Ø in order to complete a target sentence toward the end of each dialogue. These choices always occurred before singular nouns. The accuracy rates for target items with [+definite, -specific] and [-definite, +specific] contexts were significantly lower than for [+definite, +specific] and [-definite, -specific] contexts. The authors accounted for this result by suggesting that some of the learners had yet to settle on the feature of definiteness. Fluctuation between definiteness and specificity in article decisions has been identified for L1 Russian speakers in Ionin, Zubizarreta, and Maldonado (2008) and L1 Japanese speakers in Snape (2005). Zdorenko and Paradis (2008) also found variation between *the* and *a* in [-definite, +specific] contexts for children whose L1s lack articles as well as children whose L1s contain articles.

With the exception of Snape (2005), these recent studies investigating the Fluctuation Hypothesis have not considered article choices for uncountable nouns. Snape compared results of L1 Japanese and L1 Spanish speakers on an elicitation task similar to the one used in IKW. In addition to singular nouns, the task included plural nouns and noncount mass nouns in target NPs within [+definite, +specific], [-definite, +specific], and [-definite, -specific] contexts. As predicted by the FH, in the [-definite, +specific] context, the Japanese participants exhibited variation between *the* and *a* before singular nouns and variation between *the* and Ø before plural nouns. For uncountable mass nouns in this semantic context, however, Snape found variation between *a* and Ø. This result could be interpreted to mean that the learners correctly based their article decisions on [-definiteness] and that the variation was the result of learners' uncertainty over count status. Yet, why would learners fluctuate between definiteness and specificity for article choices before singular and plural nouns but not before noncount nouns?

Interestingly, the Japanese participants in Snape (2005) showed variation between *the* and Ø in decisions before mass nouns within the [+definite, +specific] context. Snape accounted for this result as a sign of L1 interference, given that Japanese does not distinguish between count and mass nouns. If such interference leads to omission of articles (i.e., the choice of Ø) in definite contexts, then the same explanation could hold for omission in indefinite contexts. It is possible that uncertainty over noun countability could at times distract learners from considering the semantic context of a NP. The question of count status may take precedence: if a noun is deemed to be uncountable, no article is chosen. Thus, Ø choices in the [-definite, +specific] context may have had nothing to

do with sensitivity to [-definiteness]. Rather, the variation between Ø and *a* in this context and between Ø and *the* in the [+definite, +specific] context could have been the result of participants' fluctuation between exclusive attention to countability and attention to semantic context (as well as countability for choices of *a*).

Semantic context of the NP has been emphasized in efforts to explain omission, identified as the most frequent article error both for beginning-level learners (Master, 1987; Parrish, 1987; Thomas, 1989) and advanced learners (Leung, 2007; Robertson, 2000; White, 2003). For example, Jarvis (2002) considered use of zero article by the Finnish speakers in his study to reflect "the L1 Finnish convention of avoiding (what Finns perceive to be) redundant markers of definiteness and indefiniteness" when such features are salient in the discourse (p. 416). Utilizing variable rule analysis, Young (1996) found that Czech and Slovak speakers were likely to omit articles before NPs that appeared as rhemes, which were coded as [-HK], in the final position of an utterance. Young accounted for this result as a transfer effect of the pragmatic use of word order in Czech and Slovak. Working with Chinese speakers, Robertson (2000) found omission to occur in cases where the definiteness or indefiniteness of a NP was "recoverable from the context" (p. 135). Each of these explanations requires learners to be sensitive to definiteness (and indefiniteness).

Alternative accounts for article omission have been put forth. To name just two, Master (1997) claimed learners are responding to input (i.e., omission is the reflection of the high frequency of Ø in English), and Trenkic (2008) proposed that omission may be the result of NNSs misinterpreting articles as adjectives, which are omitted when understood to be pragmatically redundant. With a few

exceptions (Butler, 2002; Master, 1987; Snape, 2005; Young, 1996), most of the studies cited above did not give detailed consideration to countability and how it may influence learners' choice of article.

Research which has addressed countability has shown that inaccurate identification of count status influences L2 learners' production of articles (Butler, 2002; Hiki, 1991; Yoon, 1993). In fact, Master (1987) asserted that countability "appeared to cause the most persistent difficulty in article acquisition" (p.181). For L1 Japanese speakers with high English proficiency, Yoon (1993) found a correlation between intuitive countability judgments made for nouns presented with no context and article choices made later on a gap-fill task. No such correlation was found for native speakers (NS), who were more flexible when choosing articles before nouns in context. To illustrate, if *life* had been judged to be uncountable given no context, NSs were able to supply *a* when the context implied count status (e.g., *It's a wonderful life*), whereas NNSs were more likely to choose Ø infelicitously (e.g., *It's Ø wonderful life*). In an analogous finding, Butler (2002) reported that lower-level learners treated countability as "a fixed or static entity" (p. 466) and were unaware of the potential for nouns to be countable in one context and uncountable in another. In addition, participants across proficiency levels "often cited the difficulty of determining countability for indivisible entities" (p. 471) such as nouns like *environment* or *culture*. Similarly, Hiki (1991) found that L2 English speakers made the greatest number of countability misjudgments with abstract nouns on an editing task and were least accurate in article use with abstract nouns on a composition writing task.

Studies that have investigated countability have not focused on how it may

influence article choice within definite contexts. For example, Hiki (1991) and Yoon (1993) analyzed article choices for indefinite contexts only. Although Butler (2002) discussed learners' misdetection of countability, the semantic contexts for those count misdetections were not reported. Although Young (1996) found singular number status to influence the likelihood of NPs marked with *a* and noncount status to influence the likelihood of NPs marked with Ø, it is unclear if these influences were maintained across indefinite and definite contexts.

In summary, much research on L2 English article acquisition has considered the role of semantic context, including the features of definiteness (or HK) and specificity (or SR). There has been less consideration of the role countability may play in NNSs' uses of articles. From previous research, it is evident that learners: (a) may choose the definite article based on the specificity of the NP, (b) often exhibit higher rates of accuracy for article choices within definite contexts than indefinite contexts, and (c) struggle to determine the count status of nouns in context. Given a lack of attention to the potential influence of countability on article choice across semantic contexts, it remains unclear if there is a link between learners' determination of count status and article omission in definite contexts.

The Present Study

The current study investigates the influence of both countability and semantic context on NNSs' article choices. To do this, a forced-choice elicitation task was created. Target NPs were comprised of three noun types (imaginable count, abstract count, and noncount) within six semantic contexts (the same used in IKW). Participants with article-less L1s completed the elicitation task, rated the confidence of their article

choices, and wrote explanations for those choices. The data from these three tasks were analyzed in order to answer the following research questions (RQs).

Research Questions

1. How do noun type and semantic context influence learners' choice of the definite article, choice of the indefinite article, and choice of zero article?
2. How do participants rate confidence for article choices with imaginable count, abstract count, and noncount nouns?
3. When participants express awareness of countability, what patterns can be identified in their article choices?

Hypotheses

For RQ1, it was predicted that noun type would not influence choice of the definite article, as advanced learners should be aware that *the* can appear with each of the three noun types. Noun type, however, was predicted to influence choices of indefinite and zero article. Following Young (1996), imaginable count nouns should favor *a* and disfavor Ø; noncount nouns should favor Ø and disfavor *a*. Due to the anticipated challenge of determining countability for abstract count nouns, no prediction was made on how this noun type would influence the choice of *a* and Ø. Based on learners' high rates of accuracy in [+definite, +specific] and [-definite, -specific] contexts (IKW; Ionin et al., 2008; Snape, 2005), the former context was hypothesized to favor *the* and disfavor *a* and Ø, while the latter context was predicted to favor *a* and Ø and disfavor *the*. Given fluctuation between definite and indefinite articles for [+definite, -specific] and [-definite, +specific] contexts in IKW, these contexts were not expected to favor any particular article choice.

For RQ2, given previous findings on learners' difficulties determining countability for nouns in context (Butler, 2002; Hiki, 1991; Yoon, 1993), uncertainty over count status was predicted for both noncount and abstract count nouns. This uncertainty was expected to translate into lower confidence ratings for article choices before these noun types. Because imaginable count nouns would likely be construed as individuals and not as "indivisible entities," participants were expected to determine these nouns as countable more easily, and thus be more confident in their article choices.

For RQ3, following Yoon (1993), it was expected that nouns (correctly or incorrectly) identified as countable in indefinite contexts would match choices of *a*, and nouns identified as uncountable in indefinite contexts would match choices of Ø. Given the above analysis of Snape's (2005) results (i.e., that learners may at times focus on countability to the exclusion of semantic context, associating noncount status with no article), the choice of Ø was also hypothesized to appear frequently for nouns identified as uncountable in definite contexts.

Method
Participants

Participants were enrolled in four sections of an advanced ESL writing course at Michigan State University. Students had been placed into the course based upon their results on the Michigan State University English Language Test, which measures listening, reading, and writing skills. In order to restrict the target group to learners with article-less L1s, data from five L1 Arabic students were excluded from the present analysis because Arabic has a definite prefix and an indefinite suffix (Lyons, 1999). Remaining participants included 41 learners from five different L1 backgrounds (see Table 1). In order to create an accuracy baseline for the test instrument,

Table 1 *Participating Students*

L1	*N*
Korean	33
Chinese[a]	4
Japanese	2
Thai	1
Turkish[b]	1
Gender	*N*
Female	19
Male	22
Average age	24
	(Range: 18-49)
Average years of	7
English study	(Range: 2-15)

[a]Two participants listed Chinese as their L1 and two others listed Cantonese.
[b]See White (2003) for discussion of lack of definite determiners and realizations of specificity in Turkish.

20 NSs of American English also participated by completing the forced-choice elicitation test.

Test Instruments

The test instrument, adapted from IKW, was constructed to elicit article decisions for target items which were controlled for noun type and semantic context (see Appendix). Through target items, participants were exposed to the issue of countability and the semantic features of definiteness and specificity. A forced-choice elicitation task was considered the best way to do this, given that learners are known to avoid problematic structures and uses in free production (Butler 2002; Kharma 1981; Mizuno 1985). In efforts to minimize the amount of class time required for the study, the test was limited to 18 short dialogues, which appeared in randomized order across four separate test forms. Each dialogue contained one target item, which always appeared near the end of the dialogue and in object position. Participants chose between

a, *the*, and no article (−). In order to emphasize semantic over phonological features, no choice was required between *a* and *an*.

Each of the three noun types contained six target items. Noun Type 1 (imaginable count) was comprised of the following nouns: *book*, *creator*, *owner*, *girl*, *play*, *student*. These appeared in dialogues taken directly from IKW and were all in singular form. The term imaginable (see Trenkic, 2008, p.16) is here understood as a feature which signifies that the referent of a noun is clearly discernable as a discrete object or entity. For example, the edges of a book, the outline and movement of a person, the characters and plot of a play enable these objects and entities to be visualized as individuals. Such individuation is more difficult with the abstract items (*life, imagination, style, mood, atmosphere, environment*) comprising Noun Type 2. Of the six nouns in this type, five[3] hold below-average concreteness ratings within the MRC Psycholinguistic Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). As with most abstract nouns, depending on the context, these particular nouns may appear as countable or uncountable (Celce-Murcia & Larsen-Freeman, 1999). By including them as singular count nouns on the elicitation task, an opportunity was created in which to examine the article choices of participants who might mistakenly identify these nouns as uncountable. Given the tendency for abstract nouns to appear more often as uncountable (Biber, Johansson, Leech, Conrad, & Finegan, 1999) and previous research findings (Butler, 2002; Hiki,1991), such misdetections were expected. Finally, for Noun Type 3 (noncount), the items (*bread, water, sand, money, furniture, equipment*) were chosen from a recent grammar book (Maurer, 2006, p. 118), which lists each one as an uncountable noun.

The semantic contexts, controlled for definiteness and specificity, were the following:

Context1   [+definite, +specific]
            Previous Mention
Context2   [+definite, +specific]
            Explicit Speaker Knowledge
Context3   [+definite, -specific]
            Denial of Speaker Knowledge
Context4   [-definite, +specific]
            Explicit Speaker Knowledge
Context5   [-definite, -specific]
            First Mention
Context6   [-definite, -specific]
            Denial of Speaker Knowledge

IKW's definitions were used to operationalize definiteness and specificity (see Literature Review above). It should be noted that Contexts 1 and 2 shared, as did Contexts 5 and 6, the same features. As in IKW, Contexts 1 and 5 were included since ESL grammar books often present the fundamental rule: use *a/an* for first mention and *the* for subsequent mention (Celce-Murcia & Larsen-Freeman, 1999).

Explanation sheets required participants to note their confidence levels for the 18 article choices. Confidence for each decision was expressed by checking off one of four possibilities:

___   My answer is definitely right. I am completely confident.

___   My answer is probably right. I am pretty sure about my choice.

___   My answer might be right.

___   My answer is a complete guess.

Below the confidence rating, space was provided in which to write an explanation for the particular article choice. Such a format, like Butler's (2002) interviews, sought to encourage participants to reflect upon their own article choices and to express their strategies for article use.

Procedure

The research project was introduced by the researcher to each of the four sections of the ESL writing course.[4] Before execution of the study, each of the section instructors admitted to spending no class time on articles. What little attention had been paid was limited to error correction of students' writing. The project was carried out over four class periods, as may be seen in Table 2. Participating teachers were given a script with specific instructions for each day.

Table 2 *Project Calendar*

Day 1	Consent forms
Day 2	Background questionnaires Forced-choice elicitation test Explanation sheets for homework
Day 3	Explanation sheets collected
Day 4	Review of NS-completed test

On Day 2, pens[5] and background questionnaires, which sought information on age, gender, country of origin, native language, and years of English study, were distributed. After finishing the questionnaires, participants were given 10 minutes in which to complete the forced-choice elicitation test. This period of time was determined through piloting of the test with international graduate students who were proficient and fluent in English. The average time it took these individuals to complete the activity was doubled for the learners in the study, as it was felt that reading the dialogues too quickly could yield faulty textual representations which could influence article decisions.[6] Teachers reported that all students finished in the allotted time.

Explanation sheets, which sought confidence ratings and explanations for each article choice (see description of task above), were distributed to the students, who retained their elicitation tests. Participants were instructed to complete the sheets for homework.[7] On Day 3, explanation sheets and tests were collected by the teachers. On Day 4, NS-completed test dialogues were distributed. Each article choice was coded for definiteness/indefiniteness and for countability in the indefinite contexts. The coded dialogues were meant to afford participants a learning opportunity. Teachers were given the option of discussing the dialogues in class or assigning them for outside reading. Because the study was carried out during regular class time, providing explanations for the article choices made by NSs was considered a step toward ecological validity.

Data Analysis

To compare the role of noun type and semantic context in predicting article choice (RQ1), variable rule analysis (VARBRUL) was employed. This type of probabilistic multivariable analysis, common in sociolinguistics (see Tagliamonte, 2006), utilizes logistic regression in order to compare factor groups and to weight individual factors within those groups. Although the present data were collected through an experimental forced-choice elicitation task and not through more natural language production (e.g., narratives told in an interview or writing produced in an essay), they meet the three key VARBRUL requirements set out by Sankoff (1988): choice, instability, and recurrence. In the present study, (a) article decisions were choices (participants could select from the options of *the*, *a*, or *Ø*), (b) these decisions were unpredictable in that learners did not converge on the same choices nor on the choices made by NSs, and (c) decisions

occurred repeatedly across the elicitation task.

The decision to use VARBRUL, specifically Goldvarb 3.0, was made in recognition of Preston's (1996) belief "that the discovery and weighting of influencing factors is the most valuable area of interaction between variation linguistics and SLA" (p. 25). The design of the elicitation task is well suited for the comparison of factor groups. It is possible to compare noun type (as one factor group) with semantic context (as a second factor group) in terms of their influence on participants' article choices. Within each of these groups, individual factors can be weighted against one another. Thus for a particular article choice, it is possible to determine the strength with which the factors favor or disfavor that choice. As these factors are known, it is not suitable to apply an exploratory factor analysis, which identifies latent factors. Nor is a confirmatory factor analysis suitable given an absence of previous research identifying loadings for the various factors. Unlike MANOVA which requires subsequent statistical analysis in the form of univariate ANOVA or discriminant analysis in order to arrive at meaningful results, VARBRUL may be carried out in one step. Furthermore, VARBRUL does not require data to be normally distributed.

To address RQ1, three step-up/step-down VARBRUL runs were made: one for definite article choice, another for indefinite article choice, and a third for zero article choice. For the first run, each choice of *the* was coded as 1 and each choice of *a* or *Ø* was coded as 0. For the second run, *a* was coded as 1 and *the* and *Ø* as 0. For the third run, *Ø* was coded as 1 and *the* and *a* as 0. Two factor groups, one representing noun type and the other semantic context, were included in the analysis. The first group was comprised of three factors: imaginable

count, abstract count, and noncount. The second group was comprised of six factors: the six semantic contexts listed above.

To address RQ2, participants' confidence ratings, which had been marked on the explanation sheets, were coded on a scale of 1-4. The following values were assigned:

4   My answer is definitely right. I am completely confident.
3   My answer is probably right. I am pretty sure about my choice.
2   My answer might be right.
1   My answer is a complete guess.

For each of the 18 items, a confidence score was computed by totaling the values for participant responses on that item and dividing by the total number of responses. In order to check for a significant effect of noun type on confidence, a one-way analysis of variance (ANOVA)[8] was performed. The three noun types were the factors, and item confidence score was the dependent variable. Additionally, based on an individual's ratings across the 18 items on the task, confidence hierarchies (by noun type) were established for each participant.

In order to address RQ3, participants' explanations were coded for reference to countability. This included cases where any variation on the word *countable* (e.g., *count, countable, uncountable, uncount, noncount,* etc.) was written as well as cases where the number of the target item was mentioned (e.g., Participant 12 wrote "number of girl is one" for item 10, and Participant 46 wrote "there is only one owner of his store" for item 7). Explanations were further coded for correct and incorrect labeling of countability. There was the potential for the following labeling errors: [imaginable count → noncount], [abstract count → noncount], [noncount → count]. Additionally, all explanations that

referred to countability were examined for reference to other factors and were coded for such reference (e.g., hearer knowledge, previous mention, etc.). Finally, article choices for both mislabeled and correctly labeled countability were tabulated across definite and indefinite contexts within each of the three noun types.

Table 3 *The Influence of Semantic Context on Definite Article Choice*

Corrected mean	.391	
Log likelihood	-367.418	
Total N	738	
	Factor weight (%)	*N*
Semantic context		
[+definte,+specific] PM	.84 (77%)	123
[+definite,+specific] ESK	.82 (74%)	123
[+definite,-specific] DSK	.72 (63%)	123
[-definite,-specific] FM	.23 (16%)	123
[-definite,+specific] ESK	.20 (14%)	123
[-definite,-specific] DSK	.18 (12%)	123
Range	.66	

Results
RQ1: How do noun type and semantic context influence learners' choice of the definite article, choice of the indefinite article, and choice of zero article?

When definite article choice was the dependent variable (see Table 3), only the factor group of semantic context was a significant predictor ($p < 0.001$); noun type was not significant. Within a factor group, individual factor weights closer to one encourage application and weights closer to zero discourage application (Young, 1996). From Table 3, it is evident that [+definite] contexts favored choice of *the*, whereas [-definite] contexts disfavored that choice.

Table 4 *The Influence of Noun Type and Semantic Context on Indefinite Article Choice*

	Factor weight (%)	N
Corrected mean	.296	
Log likelihood	-363.221	
Total *N*	738	
Semantic context		
[-definite,+specific] ESK	.83 (65%)	123
[-definite,-specific] DSK	.76 (55%)	123
[-definite,-specific] FM	.64 (43%)	123
[+definite,-specific] DSK	.33 (20%)	123
[+definite,+specific] ESK	.28 (17%)	123
[+definite,+specific] PM	.16 (10%)	123
Range	*.67*	
Noun type		
Imaginable	.75 (55%)	246
Abstract	.45 (31%)	246
Noncount	.29 (20%)	246
Range	*.46*	

Table 5 *The Influence of Noun Type and Semantic Context on Zero Article Choice*

	Factor weight (%)	N
Corrected mean	.146	
Log likelihood	-312.117	
Total *N*	738	
Semantic context		
[-definite,-specific] FM	.76 (41%)	123
[-definite,-specific] DSK	.67 (33%)	123
[-definite,+specific] ESK	.51 (21%)	123
[+definite,-specific] DSK	.44 (17%)	123
[+definte,+specific] PM	.35 (13%)	123
[+definite,+specific] ESK	.26 (9%)	123
Range	*.50*	
Noun type		
Noncount	.76 (37%)	246
Abstract	.64 (26%)	246
Imaginable	.15 (4%)	246
Range	*.61*	

When indefinite article choice was the dependent variable (see Table 4), both semantic context and noun type were significant predictors ($p < 0.001$). The strength of factor groups relative to each other may be identified through a comparison of the ranges within each factor group: the higher the range, the more influence that factor group exhibits on variation within the dependent variable (Tagliamonte, 2006). For choice of the indefinite article, semantic context (range = .67) was a stronger influence than was noun type (range = .46). Contexts of [-definite] favored choice of *a*, while [+definite] contexts disfavored this choice. For noun type, imaginable count nouns favored *a*, whereas noncount nouns disfavored and abstract count nouns slightly disfavored this choice.

When zero article choice was the dependent variable (see Table 5), both factor groups were again significant predictors ($p < 0.001$). This time, however, noun type (range = .61) showed a stronger influence than did semantic context (range = .50). Noncount and abstract count nouns favored choice of zero article, whereas imaginable count nouns strongly disfavored this choice. Although contexts of [-definite, -specific] favored and [+definite, +specific] contexts disfavored Ø, both [-definite, +specific] and [+definite, -specific] showed little influence in terms of favoring or disfavoring this choice.

RQ2: How do participants rate confidence for article choices with singular imaginable, singular abstract, and noncount nouns?

As one explanation sheet lacked both confidence ratings and article explanations, it was discarded from analyses addressing RQ2 and RQ3. Thus, data from 40 participants were considered in the analysis of confidence. Confidence scores for each item are found in Table 6. These individual item ratings resulted in the following average confidence scores by noun type: 2.96 for imaginable count, 2.31 for abstract count, and 2.66 for noncount. Results of the one-way ANOVA showed a significant effect of noun type on confidence ratings, $F(2,15) = 10.04$, $p < .01$, $\omega = .71$.

Post hoc tests[9] revealed a significant difference ($p < .01$) between ratings for imaginable nouns and abstract nouns; no other significant differences ($p > .05$) between noun types were found. On an individual level, nearly half the participants rated themselves least confident with abstract nouns and most confident with imaginable nouns. A full list of confidence hierarchies is presented in Table 7.

Table 6 *Item Confidence Scores*

Test item	Noun type	Score	SD
#1	Imaginable	3.38	(.85)
#2	Abstract	2.55	(1.00)
#3	Noncount	2.90	(1.03)
#4	Imaginable	2.76	(.83)
#5	Abstract	2.45	(1.06)
#6	Noncount	2.64	(.99)
#7	Imaginable	2.60	(.98)
#8	Abstract	2.34	(1.05)
#9	Noncount	2.30	(1.00)
#10	Imaginable	2.95	(1.05)
#11	Abstract	2.33	(1.01)
#12	Noncount	2.64	(.93)
#13	Imaginable	2.74	(1.00)
#14	Abstract	2.00	(.78)
#15	Noncount	2.85	(.90)
#16	Imaginable	3.35	(.86)
#17	Abstract	2.18	(.96)
#18	Noncount	2.63	(.94)

Table 7 *Frequency of Confidence Hierarchies by Noun Type*

Hierarchy	N
I > NC > A	19
NC > I > A	5
I = NC > A	4
I > NC = A	4
I > A > NC	3
A > I > NC	2
NC > A > I	1
I = A > NC	1
I = NC = A	1

Note. I = Imaginable Count,
NC = Noncount, A = Abstract Count.

These results confirm the prediction that confidence would be higher for imaginable count nouns and lower for abstract count nouns. Although confidence was also predicted to be lower for noncount nouns, average confidence scores were not significantly different from the other two noun types.

RQ3: When participants express awareness of countability, what patterns can be identified in their article choices?

A majority of participants who completed the explanation sheets (34 out of 40) mentioned countability, and reference to countability appeared in 230 of 720 explanations. Two types of countability mislabeling occurred in the explanations. There were 13 instances where participants referred to noncount target items as countable and 45 instances where participants referred to abstract count items as uncountable. Furthermore, there were three cases (two for an abstract count noun and one for a noncount noun) where participants wrote that they were unsure if the target noun was countable or uncountable. No imaginable count items were identified as uncountable.

Of the 230 explanations that mentioned countability, 122 made exclusive

reference to the countability status of the target item (i.e., no other factors were noted in the written explanations). For instance, Participant 4 simply wrote "Atmosphere is uncountable noun" for Item 17, and Participant 17 wrote "I think style is not able to count" for Item 11. The other 108 explanations made reference to the following additional factors: first mention (n = 24), previous mention (n = 7), speaker knowledge (n = 5), hearer knowledge (n = 13), speaker and hearer knowledge (n = 17), specific referent (n = 14), reader knowledge[10] (n = 25), proper noun (n = 2), and preceding adjective (n = 1). Table 8 presents article choice totals for target items where count status was accurately labeled, while Table 9 presents the same for target items where count status was mislabeled.

Table 8 *Article Choices where Countability Labeled Correctly*

Context	Imaginable count		Abstract count		Noncount	
	[+definite]	[-definite]	[+definite]	[-definite]	[+definite]	[-definite]
Article choice	*N*	*N*	*N*	*N*	*N*	*N*
the	**14**	1	**7**	1	**14**	6
a	9	**36**	1	**11**	0	0
Ø	0	1	0	1	22	**45**
Total	23	38	8	13	36	51

Note. Bold font reflects accurate article choices.

Table 9 *Article Choices where Countability Labeled Incorrectly*

Context	noncount → count		abstract count → noncount	
	[+definite]	[-definite]	[+definite]	[-definite]
Article choice	*N*	*N*	*N*	*N*
the	**2**	0	**6**	4
a	1	10	1	**1**
Ø	0	**0**	12	21
Total	3	10	19	26

Note. Bold font reflects accurate article choices.

The numbers in these tables support the hypotheses made for RQ3. For indefinite contexts, it was predicted that nouns identified as countable would match choices of the indefinite article and nouns identified as uncountable would match choices of zero article. Results show that when count status was correctly labeled countable (for imaginable count and abstract count nouns), *a* was chosen 92% of the time (47 of 51 choices) and when correctly labeled uncountable, Ø was chosen 88% of the time (45 of 51 choices). When participants mistakenly identified noncount nouns as count, they chose *a* at a rate of 100% (10 of 10 choices); when they misidentified count nouns as noncount, they chose Ø at a rate of 81% (21 of 26 choices). The final prediction was that nouns identified as uncountable would also match choices of zero article in definite contexts. For these contexts, of the 36 nouns correctly labeled uncountable, 22 (or 61%) were with choices of Ø; of the 19 incorrectly labeled uncountable, 12 (63%) were with choices of Ø.

Discussion

VARBRUL results for choice of the definite article (see Table 3) indicate that this choice was favored for [+definite]

contexts and was not influenced by noun type. Thus, it appears participants were sensitive to the appropriate feature (i.e., definiteness) in their choice of *the*. For choice of indefinite article and zero article, noun type and semantic context were significant predictors (see Tables 4 and 5). Based on the VARBRUL results, *a* was clearly favored in [-definite] contexts and disfavored in [+definite] contexts. This again implies that participants were sensitive to the appropriate feature (i.e., indefiniteness) in the semantic context. Participants appear to have been less sensitive to semantic context in their choice of zero article. Although [-definite] contexts mostly favored Ø and [+definite] contexts mostly disfavored Ø, individual factor weights for semantic context produced a smaller range for Ø-application (.50) than for *a*-application (.67) and *the*-application (.66). Given this and the range of .61 for noun type on choice of zero article, it appears participants were more sensitive to noun type (and the issue of countability) than to features of definiteness and indefiniteness when choosing Ø.

An interesting finding in Tables 4 and 5 is the patterning of noun types. For both indefinite article and zero article choice, abstract count nouns pattern with noncount nouns. That is, both abstract nouns and noncount nouns disfavor *a* and favor Ø. The participants have it right for the noncount nouns, but they have it wrong for the abstract nouns, which – based on the design of the elicitation task – are countable and should be patterning exactly like imaginable count nouns. Instead of choosing *a* for singular abstract nouns in indefinite contexts, participants were inclined to choose zero article. These results suggest that the ESL learners in this study had trouble classifying abstract nouns as countable. Perhaps, the participants were sensitive to the tendency of abstract nouns to

be uncountable (Biber, et al., 1999) and were resistant to considering them countable.

Confidence ratings reveal participants' greater uncertainty in article decisions before abstract count nouns. Such uncertainty could arise from learners' awareness that abstract nouns may appear as countable or uncountable, depending on context. Lower confidence for these items aligns with previous observations that learners find the determination of countability particularly challenging for abstract nouns (Butler, 2002; Hiki, 1991).

From the analysis of participants' explanations for article choices, a clear pattern emerges. The determination of a noun as countable leads to the choice of the definite or indefinite article, whereas the determination of a noun as uncountable often leads to the choice of zero article. Consider that for the 95 cases where target items were identified as countable on the explanation sheets, Ø was chosen only twice. For the 132 target items identified as uncountable, Ø was chosen in 100 of these instances. In indefinite contexts, of course, zero article is the correct choice. Accordingly, it is possible that for the 66 choices of Ø for [-definite] items, participants were considering the role of countability and semantic context. Yet, it is equally possible that they only considered countability, if in fact they were following a strategy of noncount→Ø. For definite contexts (where the correct choice is *the*), Ø was chosen a majority of the time (34 of 55 choices). In these instances, it is possible that participants considered neither definiteness nor specificity of the target item context. Instead, they immediately sought to identify the count status of the target noun. Once that status was determined to be noncount, zero article was chosen.

Such a strategy would help account for the stronger influence exhibited by the

noun-type factor group than the semantic-context factor group on choice of Ø (see Table 5). Further support that participants were employing this strategy is found in the fact that the majority of the explanations which mentioned countability made reference solely to countability. Obviously, the participants in these instances could have considered other factors and simply not written those factors down on the explanation sheets. Rather than look at all references to countability, it is perhaps more constructive to examine those for target items in definite contexts where noncount status was assigned and where inaccurate article choice was made. In the 13 cases where inaccurate article choices coincided with incorrect labels of uncountable, 11 explanations mentioned only countability. In the 21 cases where inaccurate article choices coincided with accurate noncount labels, 20 explanations mentioned only that the noun was uncountable. That *the* was not chosen in these cases is a clear sign that participants failed to recognize the feature of [+definite]. Although not all participants exhibited this pattern, 18 (i.e., nearly half) did.

From the results discussed above, it is possible to posit that a major problem for some ESL learners is that they consider the count status of the noun without considering definiteness in the semantic context. Once they perceive a noun to be uncountable, they resort to the choice of Ø. This presents an alternative explanation for some cases of article omission. Disproportionate attention to countability leads to insensitivity (rather than sensitivity) to semantic context.

Pedagogical Applications

Based upon the results of the current study, the researcher proposes efforts at article instruction emphasize the following: (a) determination of definiteness before determination of count status, (b) consideration of discourse context to determine definiteness, and (c) conceptualization of countability.

Given that the question of count status may at times distract learners from considering the semantic context of a NP, learners should be encouraged to contemplate definiteness before countability. A definite context (with common nouns) will yield the definite article irrespective of the count status of the NP. Thus, teachers might guide their students to consider countability only after indefinite contexts have been determined.

To help learners more effectively determine the definiteness or indefiniteness of a NP, teachers may want to emphasize the discourse context, especially the perspectives of the interlocutors (i.e., the speaker and the hearer). Learners can be guided to consider the speaker's presumptions about the hearer's knowledge. Participants' explanations in the current study that referred to reader knowledge (see Endnote 10) exemplify the need to place focus on the discourse participants. Trenkic (2008) suggested that learners with article-less L1s may understand the definite article as signifying that which is identifiable and the indefinite article as signifying that which is unidentifiable. The crucial question is how identification is determined. If learners seek identification outside the discourse situation (e.g., learners make article decisions based on their own ability to identify NPs rather than on the relevant interlocutors' abilities), article errors may arise. Hence, pedagogical treatments that equate identifiability with the definite article (Master 1995, 1997; Yule, 1998) must make it clear that determination of identifiability is found in the speaker's presuppositions about the hearer.

Finally, focusing on what it means conceptually for a noun to be count or nouncount (Wierzbicka, 1988) can move learners away from inflexible classifications

of a word as either countable or uncountable and toward a richer and more detailed interpretation of texts and discourse. Yule (1998) proposed exposing learners to the concept of individuation. When a noun is preceded by Ø, its referent lacks clear boundaries and thus resists individuation. When a noun is preceded by *a*, its referent possesses boundaries and thus may be construed as an individual entity. The key in all this seems to be flexibility in conceptualization and sensitivity to context. Activities could lead students to consider the same noun across different contexts – where in one it is countable and in the next it is not. This could easily be done with the abstract nouns in the current study (e.g., *It's a wonderful life* vs. *Life is beautiful)*. Perhaps in implementing some of the above recommendations, teachers may improve their students' understanding of English articles and thereby challenge the observation made by Gass and Selinker (2008) that "the English article system… appears to be virtually impermeable to instruction" (p. 383).

Limitations and Future Studies

Among the shortcomings of the current project was the limited amount of class time available for implementation of the study. Additional class time would allow for more target items to be included on the elicitation test. This would enable a more confident interpretation of the quantitative results. It is possible that participants' article decisions for individual dialogues were influenced not solely by semantic features and noun types, but also by previous exposure to lexical chunks. For example, Participant 18 wrote in explanation of Item 14 that she often heard her roommate say "I'm not in a mood."[11] A greater number of target items across individual noun types and semantic contexts would allow for more

certain conclusions regarding the influence of countability, definiteness, and specificity.

Another limitation may be found in the fact that participants wrote explanations for article choices in the target language. Instead, as in Butler (2002), explanations could be sought through interviews in the L1; another alternative is for participants to complete the elicitation task while following a think-aloud protocol (e.g., Leow & Morgan-Short, 2004). These measures would allow the researcher to have a better window onto learners' thinking processes. As mentioned above, participants rated their confidence and wrote explanations for article choices at home. If researchers were interested in establishing the extent of learners' metalinguistic knowledge, participants could be required to write explanations in class. Additionally, as pointed out by an anonymous reviewer, participants could be asked to rate their confidence levels immediately after choosing articles. That confidence was marked at home may have led participants to reflect in their ratings the confidence of explanations rather than of article choices per se. Asking participants to mark confidence directly after an article choice would most likely result in more intuitive ratings.

An invaluable perspective on learners' understanding of the article system may be found in their actual article use. Thus, in future studies, it would be worthwhile to gather data through oral interviews or written compositions. Through VARBRUL analysis, the article choices participants freely produce could then be analyzed across noun types and semantic contexts. It would be interesting to see if, as in IKW, results of freely produced article use matched the patterns found in results on the forced-choice elicitation task employed here.

Conclusion

Having examined article choices made by advanced-level L2 English learners with article-less L1s, the present study has established that noun type and semantic context influenced choice of *a*, and choice of Ø. Only semantic context was a significant predictor of *the*. Analysis of participants' confidence ratings and their explanations for article choices has revealed misdetections of count status for abstract count nouns and noncount nouns. Furthermore, the learners appear to have based some article choices on countability alone, choosing Ø for target nouns they believed to be uncountable. This suggests the question of countability is a factor not only for inappropriate omission of *a* but also for omission of *the*.

References

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, *11*, 17-34.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Longman.

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma Publishers.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing, 14*, 191–205.

Butler, Y. G. (2002). Second language learners' theories on the use of English articles: An analysis of the metalinguistic knowledge used by Japanese students in acquiring the English article system. *Studies in Second Language Acquisition, 24*, 451-480.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston: Heinle & Heinle.

Chaudron, C., & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, *12*, 43-64.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.

Hiki, M. (1991). *A study of learners' judgment on noun countability*. Unpublished doctoral dissertation, Indiana University.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma Publishers.

Ionin, T., & Wexler, K. (2003). The certain uses of *the* in L2-English. In J. M. Liceras et al. (Eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*, (pp. 150-160). Somerville, MA: Cascadilla Press.

Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition, 12*, 2-69.

Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, *118*, 554-576.

Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, *24*, 387-418.

Kharma, N. (1981). Analysis of the errors committed by Arab university students in the use of the English definite/indefinite articles. *IRAL, 19*, 333-345.

Lardiere, D. (2005). On morphological competence. In L. Dekydtspotter et al. (Eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference*, (pp. 178-192). Somerville, MA: Cascadilla Press.

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, *26*, 35-57.

Leung, Y-K. I. (2007). Second language (L2) English and third language (L3) French article acquisition by native speakers of Cantonese. *International Journal of Multilingualism*, *4*, 117-149.

Liu, D., & Gleason, J. I. (2002). Acquisition of the article *the* by nonnative speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition, 24*, 1-26.

Lu, C. F.-C. (2001). The acquisition of English articles by Chinese learners. *Working Papers in SLS*, *20*, 43-78. University of Hawaii.

Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.

Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation, UCLA.

Master, P. (1995). Consciousness raising and article pedagogy. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language*, (pp. 183-204). Norwood, NJ: Ablex.

Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System, 25*, 215-232.

Master, P. (2002). Information structure and English article pedagogy. *System, 30*, 331-348.

Maurer, J. (2006). *Focus on grammar: An integrated skills approach* (3rd ed.). White Plains, NY: Pearson Longman.

Mizuno, H. (1985). A psycholinguistic approach to the article system in English. *JACET Bulletin*, *6*, 1-29.

MRC Psycholinguistic Database. Retrieved September 3, 2007, from http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm

Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, *37*, 361-383.

Preston, D. (1996). Variationist perspectives on second language acquisition. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 1-45). Amsterdam: John Benjamins.

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research, 16*, 135-172.

Sankoff, D. (1988). Variable rules. In U. Ammon, N. Dittmar, & K. J. Mattheier (Eds.), *Sociolinguistics: An international handbook of the science of language and society* (pp. 984-997). Berlin, Walter de Gruyter.

Snape, N. (2005). The uses of articles in L2 English by Japanese and Spanish learners. *Essex Graduate Student Papers in Languages and Linguistics*, *7*, 1-23.

Tagliamonte, S. A. (2006). *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.

Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning 38*, 21-44.

Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, *10*, 335-355.

Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, *11*, 1-18.

Wierzbicka, A. (1988). *The semantics of grammar*. Amsterdam: Benjamins.

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, *6*, 129-141.

Yamada, J., & Matsuura, N. (1982). The use of the English article among Japanese students. *RELC Journal, 13*, 50-63.

Yoon, K. K. (1993). Challenging prototype descriptions: Perceptions of noun countability and indefiniteness vs.zero article use. *IRAL*, *31*, 269-289.

Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 135-175). Amsterdam: John Benjamins.

Yule, G. (1998). *Explaining English grammar*. Oxford: Oxford University Press.

Zdorenko T., & Paradis, J. (2008). The acquisition of articles in child second language English: Fluctuation, transfer or both? *Second Language Research*, *24*, 227-250.

APPENDIX
ELICIATION TEST ITEMS

_____Unlike the random order which was used on the actual test forms, dialogues are presented here within the six semantic contexts. Please note for each context below, the noun type is always imaginable count in the first dialogue, abstract count in the second dialogue, and noncount in the third dialogue.

Context 1: [+ Definite, + Specific], Previous Mention

1. Vicky: Where were you yesterday? I tried to call you, but you weren't home.
   Rachel: I went to a bookstore.
   Vicky: Oh, what did you get?
   Rachel: I got lots of things – several magazines, two red pens, and an interesting book. I really liked (a, the, –) book.

2. Molly: What should we do for a vacation this year?
   Tom: Let's go to Las Vegas. It's supposed to have an amazing environment.
   Molly: Good idea. But I don't care about (an, the, –) environment. I want to spend time at the casinos.

3. Sarah: Look, I just bought some fresh bread.
   Claudia: Hey – that looks really good. And I'm hungry.
   Sarah: Let's eat then. What should we have with (a, the, –) bread?

Context 2: [+ Definite, + Specific], Explicit Speaker Knowledge

4. Kathy: My daughter Jeannie loves the new comic strip about Super Mouse.
   Elise: Well, she is in luck! Tomorrow, I'm having lunch with (a, the, –) creator of this comic strip – he is an old friend of mine. So I can get his autograph for Jeannie!

5. Andrew: The lectures in our history class have been very interesting.
   Nora: Yes, they have. I'm learning a lot about India.
   Andrew: So am I. But, unfortunately, I couldn't come to class last week. What did the professor talk about?
   Nora: She talked about (a, the, –) life of Gandhi. It was a great lecture.

6. Paul: Hi, Sheila! What are you doing in the park?
   Sheila: I'm just walking my dog around the little lake here.
   Paul: Then where is your dog?
   Sheila: Oh, there he is. He is swimming in (a, the, –) water and chasing those birds.

Context 3: [+ Definite, - Specific], Denial of Speaker Knowledge

7. Bill: I'm looking for Erik. Is he home?

Rick: Yes, but he's on the phone. It's an important business matter. He is talking to (an, the, –) owner of his company! I don't know who that person is – but I know that this conversation is important to Erik.

8. Mike: Have you seen that new building downtown?
Angela: No, I haven't. Is it impressive?
Mike: I don't know. But people seem to like it. They are talking about (an, the, –) imagination of the architect. However, because I haven't seen the building, I don't really know what to think.

9. Rose: I haven't seen your sister in a while. How is she doing?
Alex: Oh, she's great. She's in Hawaii right now.
Rose: Wow! Hawaii – I've never been there.
Alex: Neither have I, but I would love to go. My sister says she spends all her time at Waikiki Beach. She swims and lies on (a, the, –) sand. She's really enjoying herself.

Context 4: [- Definite, + Specific], Explicit Speaker Knowledge

10. Gary: I heard that you just started college. How do you like it?
Melissa: It's great! My classes are very interesting.
Gary: That's wonderful. And do you have fun outside of class?
Melissa: Yes. In fact, today I'm having dinner with (a, the, –) girl from my class – her name is Angela, and she is really nice!

11. Christina: Have you been to the restaurant next to our office recently?
Rob: The last time I ate there was 2 years ago.
Christina: It is better than it was. You should go again – they are changing everything. And they are creating (a, the, –) different style.

12. John: Hi, William! I haven't seen you in a while. What are you doing in the market?
William: Oh, just a little shopping. I am buying (a, the, –) special equipment for cooking a turkey. This recipe I am following tells me exactly what I need.

Context 5: [- Definite, - Specific], First Mention

13. Tom: How was your trip to New York?
Susan: Great! I went to many museums, and ate in lots of wonderful restaurants. I also visited many friends. And I saw (a, the, –) play.

14. Judy: Last Saturday, I didn't have anywhere to go, and it was raining. I was bored.
Samantha: So what did you do?
Judy: First, I cleaned my apartment. Then I read a book. And then I was in (a, the, –) good mood.

15. Mary: I heard your little brother got a lot of presents when he was in the hospital.
Roger: Yes, he did. He got books from my Mom, a video game from my Dad, the neighbors gave him a coloring book, and my rich uncle gave him (a, the, –) money.
Mary: Really? How much?
Roger: One hundred dollars!

Context 6: [- Definite, - Specific], Denial of Speaker Knowledge

16. Professor Clark: I'm looking for Professor Anne Peterson.
Secretary: I'm afraid she is busy. She has office hours right now.
Professor Clark: What is she doing?
Secretary: She is meeting with (a, the, –) student, but I don't know who it is.

17. Karen: Have you been to that new store?
Anne: No I haven't, but I have heard good things about it.
Karen: Like what?
Anne: People say it is developing (an, the, –) interesting atmosphere, but I am not sure what they mean.

18. Clara: I need to find your roommate Jonathan right away.
Chris: He's not here – he left a few hours ago.
Clara: Where did he go?
Chris: He is moving (a, the, –) furniture – but I don't know whose. A friend of his is moving into a new apartment.

Endnotes

[1] It may be noted that although the chart predicts *Ø Mr. Ortiz is playing tonight*, it does not predict *The Mr. Ortiz that you picked to win the MVP last year is playing tonight* or *A Mr. Ortiz is playing tonight*.

[2] Examples from Butler (2002) include: *Pass me the pen* for referential definite [+SR,+HK]; *I saw a strange man standing at the gate*, and *I keep sending Ø messages to him* for referential indefinites [+SR,-HK]; *I'm going to buy a new bicycle*, and *Ø Foreigners would come up with a better solution for this matter* for nonreferentials [-SR,-HK]; *A cat likes mice*, *The whale is a mammal*, and *Ø Language is a great invention of humankind* for generics [-SR,+HK] (pp. 478-479)

[3] *Environment* does not have a rating in the database.

[4] It should be noted that the researcher was one of the instructors.

[5] The decision to hand out pens, rather than to allow students to use pencils with erasers, was made in an effort to track any changes in article decisions. In fact, 65 out of a total of 738 item responses showed signs of being altered. Given the current design, however, there is no way to tell if these changes were made during the original administration or later at home.

[6] In fact, one of the NSs did just this. When the researcher asked her about an article choice that differed from that of the other NSs, she reread the dialogue and admitted that she had misread one of the function words, thereby altering her interpretation of the target DP. Upon second reading, her article choice corresponded to that made by the other NSs.

[7] The decision to have participants mark confidence ratings and write explanations at home was made to reduce the amount of class time required for the study. Unlimited time (at home) to complete the explanation task was considered acceptable because the goal was for participants to reflect on their article choices. There was the possibility that students might consult textbooks in search of rules to justify decisions made on the test. Again, this was considered acceptable. The focus of the project is on how learners attempt to make sense of the article system. Searching for rules in a grammar book is exactly what an advanced ESL student might do when writing an essay outside of class and considering where (or where not) to put an article.

[8] This parametric test was chosen only after Kolmogorov-Smirnov and Levene's tests revealed that item confidence scores met the assumptions of normal distribution and homogeneity of variance.

[9] Tukey's Honestly Significant Difference test was used for post hoc tests.

[10] The reader knowledge category was comprised of explanations that made reference to knowledge outside the perspective of the speaker or hearer. For example for Item 1, Participant 28 wrote "Book is countable, and the object is already obvious." Although it is possible that "obvious" could refer to the knowledge of the speaker or hearer within the dialogue, the phrasing of this explanation suggests reference is being made to the knowledge of the reader of the dialogue (i.e., the participant himself).

[11] One wonders if the participant's roommate regularly said "I'm not in *the* mood." The participant may have misidentified the unstressed article *the* as *a*.

Review of *ChinesePod* (2009). *Ken Carroll.* chinesepod.com

Reviewed by *Ching-Ni Hsieh*
PhD Candidate, Second Language Studies Program
Michigan State University
hsiehc12@msu.edu

ChinesePod is a language learning website that provides daily language course podcasts to teach spoken Mandarin Chinese to adult learners of Chinese as a second language (L2).

Learners who are interested in using the podcasts to learn Chinese need to obtain a monthly subscription. The prices for subscription vary, depending on which one of the five subscription packages (Basic, Premium, Praxis, Guided, Executive) learners select. Prices range between $5 and $199 a month. New podcasts are uploaded on the website daily and automatically downloaded to subscribers' portable media players such as iPod or personal computers. Learners can also download the podcasts onto any MP3-compatible software or device. Additionally, individual tutoring with a native Chinese speaker via daily phone calls is available to Guided and Executive subscribers.

Each podcast provided by *ChinesePod* is co-hosted by one native speaker of Chinese and one native speaker of English who introduce lesson dialogues and explain new vocabulary and key sentence structures. The podcasts posted online so far include more than 1,200 lessons ranging from beginning to advanced proficiency levels. The daily podcasts are accompanied by text expansion exercises to consolidate learning outcomes. As the material developers suggest on the website, the goal of teaching Chinese through a variety of topics in podcast format is to enable listeners to learn Chinese in a convenient and motivating manner and to cater to an individual learner's study plan and needs.

A major strength of *ChinesePod* lies in its support for social interaction and communication. Building on the sociocultural approach to second language acquisition which emphasizes language learning processes as they occur in social interactive settings (Lantolf, 2000; Swain, 2000), the founders of *ChinesePod* aim to create an open community where language learners can interact with each other through postings, comments, and a *Question & Answer* section. The *Community* section on the website provides an excellent channel for Chinese learners across the world to communicate with each other. This "community of practice" type of learning culture makes language learning more effective and engaging (Wenger, 1998). Subscribers who have common language learning goals or who are from diverse language learning backgrounds can share in sociocultural practices that may emerge and evolve as they strive toward their Chinese language learning goals. This social interaction process is believed to facilitate language acquisition (Wenger, 1998).

A second major strength of *ChinesePod* involves its focus on the global communicative approach to the instruction of pronunciation. Although a generic introduction of segmental (consonants and vowels) and suprasegmental (tones) aspects

of Mandarin Chinese pronunciation is provided on the *ChinesePod* webpage, they are not taught in isolation in the daily podcast lessons. Rather, as the global communicative approach to teaching pronunciation suggests (Naiman, 1992), the podcasters emphasize the instruction of prosody (such as stress and intonation) rather than individual sounds. This approach can enhance learners' ability to communicate in Chinese in a more linguistically and pragmatically appropriate manner because prosodic features of the Chinese language carry important functional meanings that are associated with specific contexts.

Yet another distinctive feature of *ChinesePod* pertains to its provision for different speech rates and foreign-accented L2 speech input. Research in L2 speech suggests that fluency features such as speech rate and the degree of foreign accent may affect perceptions of comprehensibility by individual listeners (e.g., Munro & Derwing, 2001; Munro, Derwing, & Morton, 2006). Research regarding the effect of speech rate on the perception of L2 speech development (e.g., Munro & Derwing, 2001) suggests that the alteration of pace in the native Chinese podcaster's speech rate may help learners to better comprehend the content of the daily podcasts, particularly for lower-level learners. Additionally, the English-speaking podcasters speak Chinese with a noticeable foreign accent. Nonetheless, their accent does not appear to interfere with the overall intelligibility of their speech in Chinese. Thus, although subscribers are listening to the foreign-accented speech of the native English-speaking podcasters, they may develop their perceptual skills in differentiating different varieties of Chinese pronunciation at the same time. Given that most *ChinesePod* subscribers are learning Chinese for business purposes, it may be beneficial for them to be exposed to different foreign-accented Chinese speech as learners

will inevitably encounter accented speech in the real world. However, it should be noted that lower-level learners could develop inaccurate Chinese pronunciation if they model their pronunciation after that of the English podcasters. Thus, subscribers need to be mindful in this regard while listening to the daily podcasts.

Several drawbacks of *ChinesePod* will now be addressed. First of all, when learners are practicing reading aloud the new vocabulary, no corrective feedback is provided to monitor their pronunciation. Unlike other language learning software, such as *Connected Speech* (Egbert, 2004), in which a speech recognition mechanism is built into the software to provide feedback on suprasegmental features such as intonation and stress, such functions are not available in *ChinesePod*. Perceptual training studies in L2 speech suggest that feedback or visual display of pitch contours on different aspects of speech, such as segmental accuracy or prosody, can help L2 learners better produce phonological features of the target language (Carey, 2004; Hardison, 2004, 2005; Wang & Munro, 2004). Other research also shows that speech-recognition-based language learning software is beneficial to the learning of pronunciation (Egbert, 2004; Hincks, 2003). Therefore, future development of *ChinesePod* should consider building in state-of-the-art speech recognition mechanisms to provide timely feedback for its users.

A second drawback of *ChinesePod* concerns the lack of visual input in the podcast format, which might have facilitated the learning process. *ChinesePod* only provides audio podcasts for listeners; a video podcast is not available. Furthermore, the daily dialogues and vocabulary are presented in a traditional listen-and-repeat fashion. Recent research on the role of gestures and facial cues in the development of L2 perception and production skills has provided positive evidence of their contributions

(Goldin-Meadow, 1999; McCafferty, 2004; Sherman & Nicoladis, 2004; Sueyoshi & Hardison, 2005). Future development of the website should include incorporating short video clips of the situational dialogues acted out by the podcasters themselves or native speakers of Chinese. Such video clips can capture nonverbal cues, including the actors' lip movements, hand gestures, or facial expressions, to enhance comprehension.

ChinesePod is a useful website to supplement the instruction of Chinese in foreign language learning contexts and is particularly useful for English learners of Chinese as the lessons are taught bilingually in Chinese and English. The podcasts cover a full range of proficiency levels and can be used with different levels of Chinese instruction accordingly. The wide variety of topics covered and the large number of available online podcast lessons make *ChinesePod* an optimal resource site for Chinese language teachers. Based on students' interests, needs, and levels, various teaching and learning purposes can be achieved: introducing the Chinese language, learning business Chinese, getting familiar with Chinese popular culture, and the like. Learners and teachers of Chinese will find these podcasts interesting and helpful for enhancing learning motivation and autonomy.

References

Carey, M. (2004). CALL visual feedback for pronunciation of vowels: Kay Sona-Match. *CALICO Journal, 21*, 571-601.

Egbert, J. (2004). Review of *Connected Speech. Language Learning & Technology, 8*(1), 24-28.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences, 3,* 419-429.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology, 8*(1), 34-52.

Hardison, D. M. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal, 22*(2), 1-16.

Hincks, R. (2003). Speech technologies for pronunciation feedback and evaluation. *ReCALL, 15*(1), 3-20.

Lantolf, J. (2000). Second language learning as mediated process. *Language Teaching, 33,* 79-96.

McCafferty, S. G. (2004). Space for cognition: Gesture and second language learning. *International Journal of Applied Linguistics, 14*(1), 148-165.

Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition, 23*, 451-468.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*, 111-131.

Naiman, N. (1992). A communicative approach to pronunciation teaching. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 163-171). Oxford: Oxford University Press.

Sherman, J., & Nicoladis, E. (2004). Gestures by advanced Spanish-English second-language learners. *Gestures, 4*(2), 143-156.

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning, 55*(4), 661-699.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*, 539-552.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Review of *SpeechInAction (2005). Richard Cauldwell & Mike Beiby*. speechinaction.com

Reviewed by *Maren Schierloh*
PhD Student, Second Language Studies Program
Michigan State University
schierl1@msu.edu

The speech of nonnative, late English language learners usually exhibits an accent that results from transferring the phonological rules and segmental or suprasegmental features from their first language (L1) into their English speech. Second language (L2) learners may also create novel pronunciations of English sounds that are not part of their L1 (e.g., Best, 1995; Flege, 1995). Native-like attainment in L2 pronunciation by adult language learners is the exception rather than the rule (Bongearts, 1999), yet most adult learners wish to speak the L2 without an accent.

Advertising with the slogan "Speak like a native speaker – understand fast speech," the authors of the web program *SpeechInAction* claim to make native-like attainment—in terms of speech perception and production—possible, a goal that likely appeals to a large audience of adult foreign language (FL) learners. The target audience of *SpeechInAction* is adult learners of English who have intermediate or higher proficiency. The overall content, speech samples, and exercises are appropriate for such learners. The program is designed to familiarize FL learners with a range of English accents so that they are better able to understand a range of everyday fast, spontaneous English speech. The authors cover a vast variety of dialects of the English language, including accents found in the US, the UK, Ireland, Canada, and New Zealand.[1] *SpeechInAction* features five main

modules, each containing 10 chapters. Completion of an entire module would take about 10 to 20 hours. The five modules are entitled: (a) "Listening and Pronunciation (British and Irish Voices)"; (b) "Listening and Pronunciation (US and Canadian Voices)"; (c) "Listening to Accents of the British Isles"; (d) "Listening to Accents of the USA"; and (e) "Dictations: Intensive Listening" (covering all major accents, including New Zealand). Each chapter is further subdivided into various units that target specific dialects.

One major asset of *SpeechInAction* is its logical and sequential organization. Overall, the program is organized by proficiency level, as some activities target intermediate learners and others target intermediate to advanced learners. Typically, the subsections move from more comprehensive listening activities (e.g., listening to longer dialogues and completing multiple-choice questions) to more specific exercises such as recognizing, distinguishing, and comparing dialects. The sequence of the exercises moves from perception to production, where the user typically compares his/her production with that of native English speakers. Each dialect unit closes with a review that displays a summary of the student's work.

SpeechInAction has a variety of strengths and weaknesses. Strengths include its presentation of a wide variety of authentic speech samples and natural discourse. These samples provide a great

deal of practice for students who aim to determine, distinguish, and ultimately comprehend different dialects in the English language. The dialect units first present a stretch of discourse (either a monologue or a dialogue). If a learner is unable to comprehend certain passages of the discourse, he/she can access both a prose transcript as well as a detailed transcript of each speech unit which can be played separately. The discourse is then followed by multiple-choice exercises for either general or selective listening. In order to answer the multiple-choice questions correctly, the listener must employ both bottom-up and top-down processing strategies. That is, the listener must make use of contextual knowledge as well as information on linguistic features in order to sufficiently comprehend the message. The context is usually given at the beginning of each unit: "In this chapter we work with Catherine, a university lecturer who was born and brought up in Houston, Texas, but has lived and worked in France and England for twenty years. She was 'blown away' by French" (Module 2: Listening to Accents of the USA; Accent 1: Blown Away by French). Some exercises require less contextual information, as when the task is to notice the missing and linked sounds. According to Petersen (1991), pedagogical focus on both top-down and bottom-up processes improves L2 listening comprehension.

Whenever an answer to a question is entered, correction is immediately provided on the next page and the speech sample is then segmented into phrases and into single words. Thus, the learner may develop listening skills such as identifying words, phrases, or clause boundaries to derive meaning from the stream of L2 speech. The learner can work with the accompanying exercises, in which he/she can first record him/herself and then compare his/her

pronunciation with two different dialects. The authors provide distinctive features for which the learner should listen, for example, the differences in the vowels and the consonants /t/ and /r/. The learner can further practice his/her ability to recognize different accents by listening to brief excerpts and matching each to its corresponding accent.

A weakness of *SpeechInAction* is that the listening experience is nonreciprocal or non-communicative, and it provides no visual input. A number of research studies by Hardison (e.g., 2003, 2005a, 2005b) indicated that the presentation of audio and visual cues is beneficial to speech comprehension accuracy by L2 learners. Hardison's recent comprehensive overview (2007) of research studies on the visual element in phonological processing strongly emphasizes the value of facial cues in L2 processing. In addition to facial cues, gestural cues (e.g., hand gestures) appear to facilitate L2 processing. Specifically, L2 learners' listening comprehension appears to be most accurate when the learners can make use of both hand gestures and facial cues (Sueyoshi & Hardison, 2005). Unfortunately, *SpeechInAction* provides auditory input only, and it neglects the substantial role of visual input which is naturally available in face-to-face communication. The program could therefore be improved by including video clips of the speakers.

SpeechInAction encourages its users to record themselves when producing given words or phrases to practice their pronunciation and to enhance their ability to detect and distinguish various dialects. Furthermore, learners are encouraged to self-assess their production on a scale of 1-3 (Yes, No, Try Again) for questions such as "I have pronounced the short vowels accurately/on time" or "I have produced the speech units at original speed." The

advantages of self-assessment coupled with the goal of producing speech at fast, native-like speed are not quite apparent to me. If learners attempt to produce rapid speech, their spoken language may lose intelligibility. Rather than testing whether they have produced rapid speech, learners could work with a partner to judge whether their language is intelligible, which could be of a greater value to their conversational use of English with native speakers. In fact, research studies have demonstrated that L2 learners perceive faster speech to be less accented, even if the speech was less intelligible overall (e.g., Munro & Derwing, 2001). I would suggest that instead of, or in addition to, having learners self-assess their speech production, the software program could allow learners to see their speech on a spectrogram, which would allow for various displays (vowel formants, primary stress, segmental and suprasegmantal features, and prosody). My suggestion is partly based on the findings of Hardison (2004, 2005a), which demonstrated that computer assisted prosody training[2] reliably improves L2 learners' pronunciation. Though the developers do not claim that *SpeechInAction* will enhance learners' L2 prosody, the inclusion of visual prosody displays would serve as a beneficial accessory to enhance learners' ability to detect the variation in the prosody and discourse intonation of different English dialects.

On several occasions, the authors claim that "the better you can imitate an accent or dialect, the better you will be able to understand both the accent itself and the wide range of pronunciations of other accents" (Introduction page). There is no empirical evidence to support this claim, and speakers might be able to comprehend and distinguish a variety of accents while not necessarily being able to imitate the various dialects.

As mentioned above, *SpeechInAction*

is advertised with the slogan "Speak like a native speaker – understand fast speech." Although adult learners will most likely not end up speaking exactly like native speakers (Bongearts, 1999; Flege, Munro, & MacKay, 1995; Munro & Derwing, 2002), training with this software may improve their pronunciation and refine their listening comprehension skills. The developers of *SpeechInAction* also emphasize the benefits of unscripted speech, thereby putting students of English in touch with real speech and a variety of accents.

In addition to the philosophy of motivating and rewarding the learner, the authors strongly emphasize the variety and authenticity of L2 speech samples they use. Although the web program is not interactive, and does not encompass a communicative approach, it does take an inductively and deductively balanced teaching approach. In some subsections, the authors provide an explicit explanation of the differences between two dialects, whereas other subsections direct learners in formulating their own "rules" and perceptional categories based on authentic speech samples. For example, learners have to notice the difference between sounds or cross out letters which are absent or nonaudible in spoken language. The developers seem to value analytical and metacognitive skills, as users have to analyze (among other features) vowel qualities, articulation of consonants, and segmental stress (the sample chapter does not display any analysis of suprasegmantal features). Although the analysis of phonetic features may raise learners' awareness of those in perception and production, overanalyzing language by thinking too much about its forms and consciously lingering on the phonetic domains tend to impede graduation to automaticity (DeKeyser, 1997).

Furthermore, whenever the learner is

recording his/her own language production, he/she is encouraged to compare this production to that of (an) English native speaker(s). The authors obviously value the "native speaker model." Yet, whether the native speaker's speech is the appropriate model in language teaching has been a matter of some debate (Davies, 1991; Medgyes, 1992). As I have pointed out on several occasions, native-like speech production might not be a realistic goal for adult language learners, and it is doubtful that the imitation of native dialects helps learners to become accentless English users.

By using *SpeechInAction,* the learner gains experience in listening to a variety of speakers who speak different dialects in various speech situations. This exposes the learner to the reality of language variation. The overt emphasis on native-like perception and production seems to be unrealistic and, in most contexts, inappropriate for adult L2 learners and/or L2 teaching. A better objective of the program would be to concentrate on enhancing the learners' intelligibility and comprehensibility rather than focusing primarily on accentedness and "native-like speech rate."

References

Best, C. T. (1995). A direct realist view of cross-language speech perception. In Strange, W. (Ed.), *Speech perception and linguistic experience* Tomonium, MD: York Press.

Bongearts, T. (1999). Ultimate attainment in L2 pronunciation: The care of very advanced late L2 learners. In D. Birdsong, (Ed.), *Second language acquisition and the critical period hypothesis* pp. 133-59). Mahwah, NJ: Erlbaum.

Celce-Murcia, M., Brinton, D. M. & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages.* Cambridge: Cambridge University Press.

Davies, A. (1991). *The native speaker in applied linguistics.* Edinburgh: Edinburgh University Press.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*, 195-221.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange, (Ed.), *Speech perception and linguistic experience* Timonium, MD: York Press.

Hardison, D. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics, 24*, 495-522.

Hardison, D. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning and Technology, 8,* 34-52.

Hardison, D. (2005a). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO, 22,* 1-16.

Hardison, D. (2005b). Second-language spoken word identification: Effects of perceptual training, visual cues, and phonetic environment. *Applied Psycholinguistics, 24,* 579-696.

Hardison, D. (2007). The visual element in phonological processing and learning. In M. C. Pennington, (Ed.), *Phonology in context* (pp. 135-58). Houndmills, UK: Palgrave Macmillan.

Medgyes, P. (1992). Native or non-native: Who is worth more? *ELT Journal, 46,* 340-349.

Munro, M. J. & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition, 23,* 451-468.

Petersen, P. W. (1991). A synthesis of models in interactive listening. In M. Celce-Murcia, (Ed.), *Teaching English as a second or foreign language* (pp. 106-122). Boston: Heinle & Heinle.

Sueyoshi, A. & Hardison, D. (2005). The role of facial cues and gestures in second language listening comprehension. *Language Learning, 48*, 393-410.

Endnotes

[1]Speech in Action does not include
Australian or South African dialects.
[2]Training in which learners can follow along
and compare their pitch contours with that
of native speakers.

Interview with Dr. Judith Kroll

Interviewed by *Emma Trentman*
PhD Student, Second Language Studies Program
Michigan State University
trentma1@msu.edu

Thank you to Judith Kroll, Distinguished Professor of Psychology, Linguistics, and Women's Studies, Penn State University, for agreeing to be interviewed for these working papers. Dr. Kroll specializes in psycholinguistics, bilingualism, and second language acquisition. She came to Michigan State University to give a plenary talk at the Second Language Studies Symposium on April 3, 2009. For more information about Dr. Kroll, please visit her faculty page: http://www.cls.psu.edu/people/faculty/kroll_judith.shtml. This interview was arranged by Emma Trentman.

Emma: How did you become involved in the field of Second Language Studies and Psycholinguistics in particular?

Dr. Kroll: When I completed my Ph.D. in 1977, I was working on a different question, which is how it is that we understand meaning based on language compared to our perceptual experience of just perceiving the world. That question is of course related to the issue of whether bilinguals and second language learners understand meaning in the same way for each of their two languages. In 1986, I was teaching at Mount Holyoke College and I had a student who was in one of my undergraduate classes who had just returned from study abroad in Germany. We had read a paper that had been published in 1984 by Potter, So, Von Eckert, and Feldman, that claimed that learners even at very early stages of acquisition were able to understand the meanings of second language words without mediating them through the first language. The student, Janet Curley, was beginning an honors thesis and she came to me and said that she didn't believe their results. She had just returned from Germany and felt that English was always in her head. I thought this was a wonderful

topic for her honor's thesis. Her results in fact supported her phenomenology, with evidence that learners at low levels of proficiency appeared to rely on the English translations of the German words whereas more proficient learners did not.

A couple of years later I had an opportunity to be on sabbatical in the Netherlands, and the Netherlands is really a candy store for research on psycholinguistics and bilingualism. My research on this topic took on a life of its own after that point. We began with this question about second language learning, and then I started doing work on proficient bilinguals. Since then, the research program has developed in many different directions and with the influence and contribution by the many different students and collaborators who have worked with me. Many of our studies represent different bilingual language pairings, and different countries, and different experiences.

In 1994, I moved from Mount Holyoke to Penn State. By 2001 or 2002, the situation at Penn State developed so that a critical mass of faculty was hired with shared interests in language learning and bilingualism, and we began working

together and became what was then called the Language Science Research Group. In 2006 we became the Center for Language Science (http://www.clas.psu.edu) and together with our graduate students we approach issues of language acquisition in the first and second language and proficient bilingualism, from a variety of perspectives, including psycholinguistics, linguistics, language disorders, and cognitive neuroscience. So that's a fast speedup to the present, and as you could see in the talk I gave this morning, one of the things that I've been really interested in doing is exploiting the power of having different perspectives. So now we're doing ERP research and also acoustic analyses of speech. Working together with colleagues in a cross-disciplinary mode enables us to do research that none of us could do alone, and I find that thrilling.

Emma: How do you see the cross-disciplinary approach as having particular advantages?

Dr. Kroll: I think there's always a tension in being able to develop the depth that's required in each project but it gives you tremendous breadth that you wouldn't otherwise have and new ideas because we each look at things in different ways. There's also a tension in that the language that we use to describe our work is not always the same across disciplines, and I think that to create effective cross-disciplinary research requires a group of people who are really committed to trying to work [through] these differences with each other. If you have this, then terrific things can happen. The other thing that has been wonderful is that we have a fantastic group of graduate students, many of whom are co-advised by different faculty. The analogy that I sometimes use is that we are the pidgin and they are the creole. It really is a

developmental process in which they are taking and running with the input from these different areas and many of them have gone on and really done fantastic things. I think that their training is very different from the training that the faculty had. Coming to your research with a very strong disciplinary bias and then trying to extend it is very different than being immersed in the joint influence of all of these things from the start.

Emma: What advice would you give to graduate students who are interested in psycholinguistics or interested in the same topics as you generally?

Dr. Kroll: I think the best advice is to really try to have an opportunity to be exposed to this breadth of research in terms of shaping your own research program. What I love about this area of research as a cognitive psychologist is that I think it provides a lens for asking almost any interesting cognitive question you want to ask. One of the things that my cognitive colleagues often do is to invent tasks that create competition in artificial ways. The thing that's beautiful about studying bilinguals and second language learners is that that competition is absolutely a natural facet of their experience so you don't have to induce anything special. Research on bilinguals is sometimes messier than what traditional cognitive experimental psychology tolerates, but I think that we've been able to bridge that gap and demonstrate the contribution of research on bilingualism to mainstream psycholinguistics.

Emma: Do you think that the new presidential administration will provide more opportunities for your interests and research?

Dr. Kroll: I hope so. Although I think that there will be more openness to exploring

these issues, I think that we have to be the ones to set the agenda. I don't think that we can expect that it's necessarily going to come top down, but I think there are certainly going to be opportunities that are enabled by Obama's presidency. In some instances I think that there may be less resistance. For example, I can't see how the English Only movement is going to survive in this context. Hopefully there will be a more sophisticated understanding of why bilingual education is important and how it should be crafted.

Interview with Dr. John Norris

Interviewed by *Emma Trentman*
PhD Student, Second Language Studies Program
Michigan State University
trentma1@msu.edu

Thank you to John Norris, Professor of Second Language Studies, University of Hawaiʻi, for agreeing to be interviewed for these working papers. Dr. Norris specializes in assessment and program administration in language education settings. He came to Michigan State University to give a plenary talk at the Second Language Studies Symposium on April 3, 2009. For more information about Dr. Norris, please visit his faculty page: http://www2.hawaii.edu/~jnorris/ This interview was arranged by Emma Trentman.

Emma: How did you first become interested in the field of second language studies?

Dr. Norris: That's such a difficult question, isn't it? I was a soccer player, and that introduced me to a lot of internationalism and intercultural ideas, and I played a lot of soccer in different places around the world, and with different people, and I suppose that was my first interest in language and culture in general. I also did a degree in German studies for my bachelor's and I did a Fulbright in Germany, studying literature, and then I started to pursue a Ph.D. in German literature at the University of North Carolina. All of these experiences coalesced to convince me at some point that what I really wanted to do was something more applied than literature studies, and so I decided at that point to make a change, and I moved to Brazil to play soccer and to teach English. So I managed to do both of those things and at some point during the teaching of English in Brazil, on the basis of essentially no qualifications whatsoever, just being a native speaker of English, I realized that maybe there's a lot more to this than what I thought.

Also, kind of interesting, I was asked there to do a lot of assessment activities like placing students into a curriculum I knew nothing about, simply because I was the native speaker again and so at that point I realized, well, I better go find out more about this stuff. That's when I looked in… back in the '90s, TESOL used to publish a guide to graduate programs in ESL, and I looked in there to find a graduate program that had a beach close to it. I found the University of Hawaiʻi, and indeed they had a graduate program in ESL, and so I said well that's my next step. I handwrote my application and my resume on a piece of paper that I scrounged up out of the back of a book, and I mailed it in, and I got lucky, and was accepted to the University of Hawaiʻi. So then I did my MA and PhD there, and those are the origins of my work in SLS.

Emma: How did you get a specific focus?

Dr. Norris: I guess the specific focus came when I hit the master's program at the University of Hawaiʻi, and I just responded well, I guess, to the variety of courses that they offered, and in particular I found myself increasingly troubled by the ways in

which assessment was being portrayed in the literature used in ESL programs around the world and used in foreign language programs that I was familiar with. So, I think right away I saw kind of a combined social value to doing something about assessment and based on my experiences of what was wrong with it as well as maybe an affinity for the kinds of research that would lend [itself] to that. That was the real beginning of my focus.

I think another thing that happened at that time was that I was trying to complete a language requirement for the University of Hawai'i graduate program, and I had learned Portuguese completely naturalistically while I was in Brazil. So I took a test that was called the Portuguese speaking test from the Center for Applied Linguistics, which is a SOPI [Simulated Oral Proficiency Interview], and I did fine on it, and I got my scores back and I used that to complete this language requirement. But at the same time, I compared the score report with what I thought I knew about my competence in Portuguese and I didn't agree with it entirely, and I thought that's not really me, the descriptions are not really true to what I could do in Portuguese, and so there again I had this other idea that well, maybe there's a lot more going on with assessment beyond just sort of the careful construct representation in a test and the reliable testing of it. But there's more there, there may be misinterpretations, there are uses of tests that may be leading people astray, there are all these things, and so I really became quite fascinated early on with that aspect of testing.

Emma: How does your work fit into the applied nature of applied linguistics?

Dr. Norris: My main interest now is language education and how things we do as applied linguists can help language

education do a good job, increase its impact on society, raise the value of language education in the perceptions of the public, and those kinds of things. Testing and program evaluation I think are quite amenable to accomplishing those kinds of ends and so that's why I've maintained my focus on that, so there's this applied intersection between assessment and the public, assessment and value, assessment and improvement, and that's definitely why I'm trying to continue focusing my efforts in that area. I think it's one area where it can really make a difference if we make even minor changes in the quality of assessments that we're creating or minor changes in how we're using assessments well or not so well.

Emma: How do you think the new presidential administration will affect the field of evaluation and assessment?

Dr. Norris: That's an interesting question. I'm teaching a seminar now called "Accreditation, Accountability, and Assessment," and we deal a lot with these large scale issues, especially accountability testing. Obama himself has certainly said that he wants to change the ways in which large scale standardized tests are used and to, in very general terms, take away a lot of the punitive aspects of that testing and turn it into something that is, in his words, more useful and more informative for educational improvement. I think that's a really good idea, and I hope that he is able and his administration are able to accomplish something along those lines, but I don't know. I think things like [the] No Child Left Behind legislation have so much momentum behind them already, and all of these testing regimes are in place already, that it's going to take a while to reverse what I think are clear damages that have been wrought. We could talk about the directions that those damages have gone, not the least of which is

that fact that language education has suffered tremendously as a result of resources and time and energy being shifted into the teaching of other core academic subjects, and so I'd like to reverse that.

Emma: What do you think are the biggest challenges facing the field today?

Dr. Norris: It's hard to say, there are so many challenges. The major challenges, it also depends on where you look, but I would say a major challenge in the U.S. anyway, and that's where I do most of my work, a major challenge in the U.S. is the perception of the value of language learning and language study and pluriculturalism and plurilingualism on the part of the public and on the part of the people in power, generally speaking. So, reversing that should be a number one priority, and I think some of the relevant professional organizations are doing things along those lines—the Modern Language Association has come out with a statement of the values of studying foreign languages, ACTFL certainly tries to organize and do things along those lines— but it's going to take a lot more 'strategerizing', to borrow a term from George Bush, to turn those value statements into actions. So, what the MLA doesn't understand for example is that you can't just say we're valuable because… you then have to actually put in place a variety of campaigns, a variety of ways of changing your practice, such that the public's perception changes about you and those kinds of efforts. I actually think assessment and program evaluation fit in very nicely as strategies for acting on the value of language education; there are ways to demonstrate, [and] assessment is a really good way to demonstrate what it is that you're accomplishing in language education. Language educators, especially at the college level, have not necessarily paid

attention to that possibility yet, so that's another area where I think we could really gain some traction.

Emma: What advice would you give graduate students interested in assessment or program evaluation?

Dr. Norris: I think, well, sort of the basic advice would be get good training in these things, and then think outside of the box; don't accept the public perceptions about or the received view about assessment and evaluation as something that is just like a metric, or as something that is just mandated and to which people tend to react. I think there is a lot more to those processes that can be quite educational or educative, and so pursue that side of evaluation and assessment, thinking about it as a way of doing different things than just measuring constructs. Instead, we are doing things like making changes in the quality of language education, doing things like transforming language learners' lives—those kinds of changes are possible, too, if we design assessments and evaluations to make that happen. I think there are examples of that work that are emerging in the literature, and so I would encourage people to work in that direction to the extent possible.