Please note:
page- and linebreaks are temporary;
overlong lines, orphans/widows, and split examples will be dealt with after your corrections
have been incorporated;
there is no need to mark split examples or bad page breaks.

Typesetter’s notes/queries
/1/ Article adapted to IRAL style sheet. There is a huge number of minor changes and
corrections, please carefully check the file for errors.
The development of lexical bundle accuracy and production in English second language speakers

SCOTT CROSSLEY AND THOMAS LEE SALSbury

Abstract

Six adult, second language (L2) English learners were observed over a period of one year to explore the development of lexical bundles (i.e., bigrams) in naturally produced, oral English. Total bigrams produced by the L2 learners over the year of observation that were shared with native speakers were compared using a frequency index to explore L2 learners’ accuracy of use. The results of the study support the notion that bigram accuracy increases as a function of time spent learning English. This finding lends credence to the notion that the production of lexical bundles by L2 learners begins to develop in parallel with the frequency of lexical bundles used by native speakers. Like native speakers, L2 learners also begin to produce lexical bundles that serve pragmatic and syntactic functions. The study provides evidence for lexical growth in both form and function and strengthens notions about deeper levels of L2 lexical knowledge.

1. Introduction

This paper analyzes the development of lexical bundle accuracy in the spoken word production of second language (L2) learners. The study is based on a frequency approach and uses a longitudinal methodology along with both quantitative and qualitative analyses to examine the spoken output of six L2 learners over a yearlong data collection period. We demonstrate that L2 learners’ accuracy in using lexical bundles (defined in this study as bigrams) develops with time spent studying English. In addition, L2 learners begin to produce lexical bundles that functioned similarly to those lexical bundles produced by L1 speakers. These findings have important implications for the development of lexical proficiency, connectionist models, and implicit learning.

This study provides a new approach for measuring L2 lexical proficiency. The approach focuses on the frequency of lexical bundle production and can thus be catego-
2 Scott Crossley and Thomas Lee Salsbury

rized as a breadth of knowledge measure (i.e., a measure that examines how many words or word combinations a learner knows). However, the approach also has connection to depth of knowledge measures (i.e., measures that refer to how well a learner knows words and their properties) because it focuses on the connections between words (i.e., which word combinations are frequent and thus allowable). Until recently, approaches that have analyzed the development of L2 depth of knowledge have been relatively rare (e.g., Crossley et al. 2008, Crossley et al. 2009, 2010a, 2010b, in press-a, in press-b; Meara 2002; Salsbury et al. in press; Schmitt 1998) with most studies focusing on word association models (Fitzpatrick 2007a, 2007b; Meara 2006; Wilks and Meara 2007). The majority of studies that have examined lexical proficiency have focused on breadth of knowledge measures (i.e., word frequency and lexical diversity; Polio 2001). These studies, while important, generally touch only on surface level linguistic features. Studies that examine both depth of knowledge and breadth of knowledge features of the lexicon are needed in order to provide researchers and language teachers with an awareness of how lexical features contribute to an overall understanding of lexical proficiency and how individual lexical features help explain lexical acquisition.

The lexical proficiency of L2 learners is a crucial area of study both in practice and in theory. In practice, the inaccurate production of lexical items is a key factor in global errors that inhibit communication (R. Ellis et al. 1994; R. Ellis 1995; de la Fuente 2002). In addition, lexical growth is strongly related to academic achievement (Daller et al. 2003). This is in addition to theoretical viewpoints which posit that lexical proficiency studies are important because they provide insight into how L2 learners process, store, and produce lexical items cognitively. Such explorations into lexical development are critical because L2 lexical acquisition is enigmatic (Schmitt 1998) and an overall theory of lexical acquisition is lacking (Meara 1996; Meara 1997; Nation 1990).

In this study, we analyze the development of lexical bundle accuracy through a bigram analysis (two word combinations) to examine L2 lexical proficiency growth. While many studies have focused on word combination errors (e.g., Bahns and Eldaw 1993; Farghal and Obiedat 1995) and written learner production (e.g., Chi et al. 1994; Granger 1998; Lorenz 1999), few if any have examined lexical bundles in spoken production. In addition, most studies that have looked at word combination errors and written learner production have used cross-sectional methodologies. Few, if any have used longitudinal methodologies. Thus, this study will address an apparent research gap in L2 oral production data by analyzing the development of spoken lexical bundle accuracy using a 12-month longitudinal study. Such an approach will allow us to demonstrate not only growth in lexical bundle accuracy, but also make links to breadth and depth of knowledge features of lexical knowledge.
Lexical bundles fall under the general term phraseology. Phraseology is the study of multi-word lexical units (McCarthy 2006) in both written and spoken language that are easily accessible and form a crucial part of native-like fluency in language (Cowie 1998; Hyland 2008). Other terms that are encompassed by phraseology and similar to lexical bundles include formulaic sequences (Wray and Perkins 2000; Wray 2002), prefabricated patterns (Granger 1998; Weinert 1995), and the common term collocations.

In this study, we use the term lexical bundles to refer to groups of two or more words that repeatedly appear in language as fixed items more frequently than expected by chance and much more frequently than phrasal verbs and idioms (Biber et al. 1999; Hyland 2008). It is the frequency of these multi-word lexical units that provides meaning in word combinations (Sinclair 1991). Lexical bundles are so frequent in fact, that they occur in over 80% of the words produced in spoken English (based on the London Lund Corpus). Two-word lexical bundles (i.e., bigrams) are the most common lexical bundles (Altenberg 1998) and may occur more frequently than high frequency words (McCarthy and Carter 1997), especially in spoken language (Altenberg 1998). Word combinations are therefore important lexically because of their repetition. Conversely, from a second language perspective, a higher frequency of significantly infrequent word combinations in second language production might signal abnormal word patterns production.

A distinction needs to be drawn between lexical bundles and other phraseological terms, especially the general term collocation. This distinction, however, is not always unequivocal. Lexical bundles, unlike collocations, are semantically transparent and regular in their formation. Most importantly, as already mentioned, they are identified not by their structure, but rather on the basis of their frequency (Biber et al. 1999; Hyland 2008). Although similar to lexical bundles, the general term collocation often refers to specialized word patterns that are consistently found together (e.g., bitter collocates with disappointment while sour does not; Shirato and Stapleton 2007). Thus, collocations do not refer to all types of co-occurrence, but to specific types of word combinations. These combinations are less rigid than idioms and are generally syntactically regular in that they follow syntactic combinations (e.g., adjective + noun collocations). For instance, common collocations for the word bank that follow the adjective + noun pattern include central bank, bank teller, and bank card. However, there is no real agreement upon how fixed or free a combination of words must be to classify as a collocation and, thus, there is no truly agreed upon definition of what a collocation is (Howarth 1996, 1998a, 1998b; Nesselhauf 2003). Generally though, collocations are commonly classified as either 1) free collocations, where words combine without restrictions (e.g., make a
decision, take a decision, and reach a decision) or 2). collocations where one word might be free, but the other might not (e.g., take a picture).

1.2. Phraseology and L2 lexical studies

The majority of studies that have focussed on phraseological features have focussed on collocations, but have defined collocations quite broadly to include lexical bundles. Thus, we use the more general term, multi-word lexical units (McCarthy 2006), in reference to research on the second language acquisition of such forms. Studies into multi-word lexical items provide a strong foundation for our current analysis of lexical bundles and, as such, we provide a summary of these studies in order to ground our lexical bundle analysis within current phraseological research.

It has long been argued that knowledge of multi-word lexical units is important for native speakers’ communicative competence (Moon 1992; Bahns and Eldaw 1993; Fontenelle 1994; Herbst 1996; Lennon 1996), but it has also been argued that such combinations are one of the most important types of lexical units that L2 learners must acquire (Farghal and Obiedat 1995; Nesselhauf and Tschichold 2002). This notion is premised on the important role of the frequency of these forms and the concept that L2 learners have more difficulty approximating the meaning of multi-word phenomena than single words (Nesselhauf and Tschichold 2002). Recent scholarship in second language research has, in fact, argued that knowledge of multi-word lexical units is not just important, but a central mechanism necessary for language learning (Hoey 2005; Lewis 1997, 2000; Nattinger and DeCarrico 1992; Richards and Rogers 2001) and also one of the most difficult to acquire (Shei and Pain 2000). Multi-word lexical units are also essential for a general understanding of lexical proficiency and word knowledge (Nation 1990).

As noted by Leśniewska and Witalisz (2007), the production of multi-word lexical units by L2 learners has not been thoroughly examined, particularly for spoken discourse. There have, however, been multiple studies that have compared the L2 written knowledge of these forms to the written knowledge of these forms produced by native speakers. These studies have demonstrated that significant differences exist between the two groups (Aston 1995; Fillmore 1979; Kjellmer 1991; Pawley and Syder 1983) and that the inaccurate use of multi-word lexical units is a strong indicator of non-native speech (Granger 2002; McArthur 1992; McCarthy 1990; Nattinger 1980; Shei and Pain 2000). Specifically, research has demonstrated that more competent L2 learners use more multi-word lexical units and use them more accurately and with more variety than less proficient L2 learners (Al-Zahrani 1998; Zhang 1993). Knowledge of these forms by L2 learners might be predictive of L2 speaking profi-
The development of lexical bundle accuracy and production

1.3. Frequency effects and phraseology

N. Ellis (2002) argues that the production and comprehension of multi-word lexical units (e.g., lexical bundles) is a function of their frequency of occurrence in language. In such a model, the frequency of multi-word lexical units helps determine lexical acquisition with the underlying assumption that each repetition of a frequent multi-word lexical unit strengthens the connections between the multi-word form and its meaning categorization. Thus, as learners are exposed to frequent multi-word lexical units, they begin to develop stronger links between that form and its meaning, creating a formula (or a lexical chunk). The more formulas a learner has, the less working memory demand are placed on the learner because the learner can rely on memorized chunks and does not need to compute the meaning between the words (Bresnan 1999; Wray and Perkins 2000). Because they are memorized and stored holistically, multi-word lexical units are accessed much more quickly, freeing up time and processing space for other aspects of discourse (Vogel Sosa and MacFarlane 2002). Multi-word lexical units are thus not viewed as being comprised of individual words,

1. The notion of holistic storage of multi-word lexical units is a theoretical position that is not supported by all researchers (e.g., Schmitt and Carter 2004; Wray 2002).
but as lexical units that can be retrieved whole from long-term memory (Pawley and Syder 1983). Theories that support a chunking mechanism are supported by word frequency studies that demonstrate high frequency words are named more quickly than low frequency words (Balota and Chumbley 1984), are processed more quickly in reading tasks (Crossley et al. 2008; Kirsner 1994), are judged more quickly to be words in lexical decision tasks (Forster 1976), and have faster response latencies (Glanzer and Ehrenreich 1979; Whaley 1978).

According to frequency models, phraseological learning results from the “gradual strengthening of associations between co-occurring” words and is probabilistic (N. Ellis 2002: 173). Such a premise aligns phraseological learning with theories of lexical networks. Lexical networks belie the notion that lexical acquisition is the simple memorization of a word and its definition and instead supports the notion that lexical acquisition is based on the interconnections between words (Haastrup and Henriksen 2000). In lexical network models, words interrelate with other words to form clusters of words that act categorically such as lexical bundles. These clusters connect to other clusters and other words, until entire lexicons are developed based on interconnections (Ferrer i Cancho and Solé 2001; Ferrer i Cancho et al. 2004; Haastrup and Henriksen 2000). Connections between words allow newly acquired words and phrases to be more easily assimilated within these networks because new words are not learned in isolation, but through links to already learned words. As learners progress lexically, they build lexical networks that are strengthened by the frequency of lexical bundles.

2. Method

The focus of this paper is a longitudinal study of the development of lexical production accuracy. We focus specifically on the longitudinal production of spoken bigrams and their frequency accuracy because most past multi-word lexical unit studies have concentrated on L2 multi-word errors or written production using cross-sectional methods. We argue that our approach is advantageous because it can provide evidence for growth in lexical bundle accuracy, depth and breadth of knowledge lexical development, as well as links to frequency models and connectionist approaches to lexical acquisition.

We selected a longitudinal approach to data collection and analysis because such an approach better focuses on the developmental nature of lexical growth and lexical networks (Haastrup and Henriksen 2000) and has been proven successful in recent lexical proficiency studies (Crossley et al. 2008; Crossley et al. 2009, 2010a, 2010b). In our approach, we are not so much interested in whether errors result within the selected lexical bundles, but rather how the frequency of the lexical bundles change over the course of a year and how
these frequencies compare to the lexical bundle frequencies of native speakers of English. We hypothesize that as L2 learners acquire lexical proficiency, the frequency of their lexical bundle production will more closely reflect that of native speakers.

We are also interested in whether or not L2 learners begin to produce lexical bundles, over time, that are common in natural, spoken discourse (cf. Biber et al. 2004; Crossley and Louwerse 2007). Therefore, in this paper we will report on two analyses. The first analysis will examine the frequency of shared lexical bundles spontaneously produced in naturalistic discourse over a 12-month period from six beginning-level L2 learners and quantitatively compare these frequencies to a normalized L1 standard. This approach is necessary to demonstrate that accuracy in the use of spoken lexical bundles develops as time spent learning English increases. The second analysis will qualitatively analyze the lexical bundles shared between the individual learners and the L1 corpus to examine if the lexical bundles produced are commonly associated with spoken discourse. This analysis will be key to understanding and classifying the actual lexical bundles produced by the L2 learners.

2.1. Bigrams

Since lexical bundles can be defined as any multiword unit, it is important to distinguish the length of the lexical bundles we will analyze in this paper. While longer lexical bundles can provide more pragmatic and contextual information, they create sparse data problems. However, bigrams, or two word clusters, do not provide sparse data problems and they provide access to both paradigmatic and syntagmatic features of language (Crossley and Louwerse 2007). Bigrams are also commonly used in natural language processing and computational linguistics (Jurafsky and Martin 2002) and are simple to compute. Thus, for this study we analyzed all bigrams shared between the native speakers and the L2 learners that occurred more than once. As Sinclair (1991) discusses, any word combination that occurs more than once is an unlikely event because of the low frequency of most words.

2.2. Participant Selection

The participants in this study were six L2 English learners who were interviewed by native speakers of English every 2 weeks (not including program and university breaks) over a 1-year period. All six participants were students in an intensive English program at a large American university; each learner started in the first level of the 6-level program and participated in the research project while continuing intensive English studies. Additionally, institutional
TOEFL scores were collected from the learners at two month intervals. The ages of the participants ranged from 18 to 29 years old. All had studied English through the secondary level of school in their countries of origin. The six learners in the study were given pseudonyms: Marta (Spanish L1), Takako (Japanese L1), Eun Hui (Korean L1), Faisal (Arabic L1), Kamal (Arabic L1) and Jalil (Arabic L1).

2.3. Corpus

As stated above, each learner was interviewed by a native speaker about every two weeks. In each interview session general discussion topics were prepared in advance. These included emotion and topic cards (e.g., lonely, excited, love, travel) from which participants could choose; questions related to photographs; questions about studying English; questions on attitudes towards the participants’ host community; and items related to imagination, wishes and hopes for the future (e.g., Imagine you have the ability to change something in your country – what do you change? Imagine you are the top person in your field – describe your life). The discussion topics often prompted related topics that participants introduced spontaneously into the discussion. Overall, the sessions were characterized by naturally occurring discourse. Each session lasted between 30 to 45 minutes. The sessions were recorded and later transcribed. This method of longitudinal data collection has been successfully employed in similar studies that examine the development of L2 lexical proficiency (Crossley et al. 2008; Crossley et al. 2009, 2010a, 2010b; Salsbury et al. in press).

The spoken data collected from the six learners was transcribed and forms the basis for the L2 corpus used in this analysis. A total of 99 transcripts were collected. Descriptive data for the corpora of each learner is presented in Table 1. The descriptive information includes the number of transcripts for each learner and average word production and bigram production per transcript. In preparation for the analysis of the learner corpus, transcriptions of each elicitation session were cleaned to eliminate interjections and all punctuation except periods and question marks. In addition, all common contractions were broken into individual words (e.g., I’m to I am, I’d to I would, I’ve to I have) and common phonological blends (e.g., gonna, wanna) were also broken into individual words (e.g., going to, want to). Each elicitation session was saved as a single text file containing the oral production of only the learner in focus, not the interviewer. The text file was manually and electronically checked for spelling errors.

---

2. Two learners (Faisal and Kamal) missed one TOEFL test each, so only five TOEFL scores were collected from these learners.
The development of lexical bundle accuracy and production

Table 1. Descriptive statistics for longitudinal language data

<table>
<thead>
<tr>
<th>Learner</th>
<th>Number of meetings/transcripts</th>
<th>Average number of words per utterance</th>
<th>Average number of utterances per text</th>
<th>Average number of words per text</th>
<th>Average number of bigrams produced per quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eun Hui</td>
<td>18</td>
<td>21.31</td>
<td>52.17</td>
<td>1120.67</td>
<td>591.00</td>
</tr>
<tr>
<td>Faisal</td>
<td>13</td>
<td>33.42</td>
<td>71.08</td>
<td>1870.07</td>
<td>731.50</td>
</tr>
<tr>
<td>Takako</td>
<td>18</td>
<td>19.40</td>
<td>51.00</td>
<td>1470.72</td>
<td>1032.50</td>
</tr>
<tr>
<td>Kamal</td>
<td>15</td>
<td>23.75</td>
<td>50.27</td>
<td>1216.20</td>
<td>588.75</td>
</tr>
<tr>
<td>Jalil</td>
<td>17</td>
<td>38.82</td>
<td>61.76</td>
<td>2359.77</td>
<td>1183.75</td>
</tr>
<tr>
<td>Marta</td>
<td>18</td>
<td>33.31</td>
<td>63.61</td>
<td>1912.00</td>
<td>1076.25</td>
</tr>
</tbody>
</table>

The learner corpus was compared to a corpus of natural speech taken from native speakers of American English. The corpus used in this study was the Santa Barbara corpus (Du Bois et al. 2000). The corpus is a collection of natural speech recordings taken from native speakers in the United States engaged in casual conversation in natural settings and is thus similar to the situations found in the L2 learner corpus. The Santa Barbara corpus consists of 200,000 tokens. In a similar fashion to that of the L2 corpus, the Santa Barbara corpus was cleaned to eliminate all interjections, and all common contractions were divided into individual words.

3. Analysis one

For each learner in this study, the number of bigrams with a frequency of two or above that were shared between the learner’s discourse and that of native speakers as found in the Santa Barbara corpus was calculated for each quarter of learning. We included both high and low frequency bigrams because we hypothesized that L2 learners would not produce bigrams in similar distributions as L1 speakers, especially in the first quarter of learning. Thus, frequent L1 bigrams might be infrequent in the L2 data and infrequent L1 bigrams might be frequent in the L2 data. The bigrams shared between the L1 speakers and the L2 learners were then normalized for frequency using Biber’s normalization formula (i.e. the incidences of a specific bigram was divided by the number of words in the transcript and multiplied by 1000; 1988) and compared. The same was done for a conflated analysis that grouped all the learner data together based on quarters of learning. The normalized frequency bigram counts from the L2 learners (both individual and conflated) were first correlated to the normalized frequencies taken from the Santa Barbara corpus to examine whether
similarities existed and whether growth in accuracy occurred across quarters. Next, the difference between the first quarter correlation and the fourth quarter correlation was calculated using z-score conversion to test whether learners demonstrated significant growth in their accuracy of bigram use from the first quarter to the fourth quarter. We argue that significant growth in bigram accuracy is best measured by the L2 learners’ production of native-like lexical bundles at native-like frequencies.

3.1. Eun Hui

Eun Hui’s discourse and that of native speakers of English shared 42 bigrams. These bigrams appeared in the L1 corpus and appeared in quarters one, two, three, and four of Eun Hui’s transcripts. Correlations between Eun Hui’s normalized bigram frequency and that of native speakers showed significant, but low correlations in the 2nd and 4th quarter (details presented in Table 2). The data did not support the notion that bi-gram accuracy increased. Comparing the correlation between the first quarter and the fourth quarter demonstrated no significant differences ($z_{diff} = .24, p = .40$).

Table 2. Correlations between Eun Hui and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.70**</td>
<td>0.36*</td>
<td>0.57*</td>
<td>0.27</td>
</tr>
<tr>
<td>2nd Quarter</td>
<td>0.72**</td>
<td>0.72**</td>
<td>0.30*</td>
<td>0.30*</td>
</tr>
<tr>
<td>3rd Quarter</td>
<td>0.73**</td>
<td>0.18</td>
<td>0.18</td>
<td>0.31*</td>
</tr>
<tr>
<td>4th Quarter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Correlations significant at $p < .05^* p < .001^{**}$

3.2. Faisal

Faisal’s discourse and that of native speakers of English shared 44 bigrams. These bigrams appeared in the L1 corpus and appeared in quarters one, two, three, and four of Faisal’s transcripts. Correlations between Faisal’s normalized bigram frequency and that of native speakers was significant in all four quarters with correlations increasing from $r = .41$ in the first quarter to $r = .63$ in the fourth quarter (details presented in Table 3). The data supported the notion that bigram accuracy increased with time spent learning English. Comparing the correlation between the first quarter and the fourth quarter demonstrated no significant differences ($z_{diff} = 1.31, p = .08$).
The development of lexical bundle accuracy and production

Table 3. Correlations between Faisal and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.58**</td>
<td>0.31*</td>
<td>0.37*</td>
<td>0.41*</td>
</tr>
<tr>
<td>2nd quarter</td>
<td>0.63**</td>
<td>0.26</td>
<td>0.34*</td>
<td></td>
</tr>
<tr>
<td>3rd quarter</td>
<td>0.50**</td>
<td>0.53**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th quarter</td>
<td></td>
<td></td>
<td></td>
<td>0.63**</td>
</tr>
</tbody>
</table>

Correlations significant at $p < .05^*, p < .001^{**}$

3.3. Takako

Takako’s discourse and that of native speakers of English shared 102 bigrams. These bigrams appeared in the L1 corpus and appeared in quarters one, two, three, and four of Takako’s transcripts. Correlations between Takako’s normalized bigram frequency and that of native speakers was significant in all four quarters with correlations increasing from $r = .62$ in the first quarter to $r = .81$ in the fourth quarter (details presented in Table 4). The increase in $r$ scores supported the notion that bigram accuracy increased over time. Comparing the correlation between the first quarter and the fourth quarter demonstrated significant differences ($z_{\text{diff}} = 3.04, p < .001$).

Table 4. Correlations between Takako and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.82</td>
<td>0.80</td>
<td>0.72</td>
<td>0.62</td>
</tr>
<tr>
<td>2nd quarter</td>
<td>0.92</td>
<td>0.87</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>3rd quarter</td>
<td></td>
<td>0.89</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>4th quarter</td>
<td></td>
<td></td>
<td></td>
<td>0.81</td>
</tr>
</tbody>
</table>

All correlations significant at $p < .001$

3.4. Kamal

Kamal’s discourse and that of native speakers of English shared 61 bigrams. These bigrams appeared in the L1 corpus and appeared in quarters one, two, three, and four of Kamal’s transcripts. Correlations between Kamal’s normalized bigram frequency and that of native speakers was significant in all four quarters with correlations increasing from $r = .69$ in the first quarter to $r = .71$ in the fourth quarter (details presented in Table 5). The correlations did not support growth in bigram accuracy because the $r$ values appeared stable. Comparing the correlation between the first quarter and the fourth quarter demon-
Scott Crossley and Thomas Lee Salsbury

strated no significant differences ($z_{\text{diff}} = .21, p = .41$).

Table 5. Correlations between Kamal and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.78</td>
<td>0.62</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>2nd quarter</td>
<td>0.84</td>
<td>0.83</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>3rd quarter</td>
<td>0.83</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th quarter</td>
<td></td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All correlations significant at $p < .001$

3.5. Jalil

Jalil’s discourse and that of native speakers of English shared 102 bigrams. These bigrams appeared in the L1 corpus and appeared in quarters one, two, three, and four of Jalil’s transcripts. Correlations between Jalil’s normalized bigram frequency and that of native speakers was significant in all four quarters with correlations increasing from $r = .48$ in the first quarter to $r = .86$ in the fourth quarter (details presented in Table 6). The data supported the notion that bigram use became more accurate as time spent learning English increased. Comparing the correlation between the first quarter and the fourth quarter demonstrated significant differences ($z_{\text{diff}} = 6.38, p < .001$).

Table 6. Correlations between Jalil and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.62</td>
<td>0.42</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>2nd quarter</td>
<td>0.68</td>
<td>0.66</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>3rd quarter</td>
<td>0.81</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th quarter</td>
<td></td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All correlations significant at $p < .001$

3.6. Marta

Marta’s discourse and that of native speakers of English shared 118 bigrams. These bigrams appeared in the L1 corpus and appeared in quarters one, two, three, and four of Marta’s transcripts. Correlations between Marta’s normalized bigram frequency and that of native speakers was significant in all four quarters with correlations increasing from $r = .36$ in the first quarter to $r = .71$ in the fourth quarter (details presented in Table 7). The data supported growth
The development of lexical bundle accuracy and production

in accuracy regarding bigram production. Comparing the correlation between the first quarter and the fourth quarter demonstrated significant differences ($z - \text{diff} = 3.87, p < .001$).

Table 7. Correlations between Marta and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.81</td>
<td>0.74</td>
<td>0.65</td>
<td>0.36</td>
</tr>
<tr>
<td>2nd quarter</td>
<td>0.86</td>
<td>0.79</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>3rd quarter</td>
<td>0.86</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th quarter</td>
<td></td>
<td></td>
<td>0.71</td>
<td></td>
</tr>
</tbody>
</table>

All correlations significant at $p < .001$

3.7. All learners

The correlation results for all learners’ bigram frequency demonstrated significant correlations with the native speaker corpus for all quarters of learning with correlations increasing from $r = .60$ in the first quarter to $r = .81$ in the fourth quarter (details presented in Table 8). The data supported the notion that bigram accuracy increased as a result of time spent learning English. Comparing the correlations between the first quarter and the fourth quarter demonstrated significant differences ($z - \text{diff} = -7.47, p < .001$).

Table 8. Correlations between all learners and Santa Barbara shared bigram frequency

<table>
<thead>
<tr>
<th></th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
<th>Santa Barbara</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st quarter</td>
<td>0.90</td>
<td>0.86</td>
<td>0.78</td>
<td>0.60</td>
</tr>
<tr>
<td>2nd quarter</td>
<td>0.91</td>
<td>0.86</td>
<td>0.86</td>
<td>0.65</td>
</tr>
<tr>
<td>3rd quarter</td>
<td>0.93</td>
<td>0.93</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>4th quarter</td>
<td></td>
<td></td>
<td>0.81</td>
<td></td>
</tr>
</tbody>
</table>

All correlations significant at $p < .001$

3.8. Post hoc analysis

We conducted additional correlations between the L2 learners’ time spent studying English (first week through the 52nd week) and the learners’ institutional TOEFL scores. This analysis was conducted to analyze if general linguistic growth occurred in the L2 learners sampled in this study. We were particularly interested in the two learners (Kamal and Eun Hui) who demonstrated no growth trends in their bigram accuracy. Our use of TOEFL scores is premised
on the notion that a single, higher-order factor related to language proficiency underlies individual first order language ability factors such as lexical ability (Oller 1979). Such a position has been supported in the literature (Carroll 1983; Harley et al. 1990; Shin 2005) and specifically in reference to TOEFL scores (Bachman et al. 1995; Bachman and Palmer 1982). We argue that if TOEFL scores reliably explain linguistic growth and they are linked to individual language factors such as lexical proficiency, then we would expect that learners demonstrating growth in lexical bundle accuracy would also demonstrate growth in TOEFL scores. The reverse should also hold true. In this manner, we can investigate a potential explanation for why growth in bigram accuracy was not observed for all learners in this study.

The correlations between time spent learning English and the L2 learners’ TOEFL scores supports this hypothesis. For the three learners that demonstrate significant changes between their first and fourth quarter bigram accuracy (Takako, Jalil, and Marta), significant correlations were reported between TOEFL scores and time spent studying English. The same was found for the one learner (Faisal) who exhibited bigram accuracy growth patterns, but did not have significant differences in correlations between the first and fourth quarters. For the two learners that did not demonstrate significant development in bigram accuracy (Kamal and Eun Hui), no significant correlations were reported for time spent studying English and their TOEFL scores (details presented in Table 9).

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.9. Analysis one discussion

The first analysis appears to demonstrate that as learners study a second language and gain proficiency in English their lexical bundle accuracy as measured through shared bigrams increases. This can clearly be seen in the frequency of bigrams used by Marta, Jalil, Takako, and, to a lesser extent, Faisal. Two learners did not demonstrate growth patterns in their accuracy of bigram use. In the first instance, Eun Hui, bigram frequency seemed generally static.
However, significant correlations with L1 speakers, although small, were noted in the 2nd and 3rd quarter. In the case of the other L2 learner, Kamal, a pattern of growth was also not apparent; though, Kamal’s bigram use did exhibit strong correlations to the frequency of bigrams produced by L1 speakers throughout all quarters of learning. The TOEFL scores of the latter two learners, however, did not demonstrate significant correlations with time spent studying English. When learners were examined as a group, significant increases were noted for all quarters of learning and, more importantly, in bigram accuracy growth between the first and fourth quarter.

These findings are important for a variety of reasons. First they support the notion that L2 learners develop lexical bundle accuracy as a function of increasing English proficiency. In addition, the findings lend credence to the development of both breadth and depth of L2 lexical knowledge. From a breadth of knowledge perspective, the findings in this study demonstrate that learners begin to produce lexical bundles that develop over time to parallel the frequency of production of L1 speakers. Importantly, this finding also links to depth of knowledge lexical features because the learners exhibit trends that demonstrate knowledge about the strength of connections between words. Overall, the data supports the notion that, over time, learners begin to produce common lexical bundles more frequently and uncommon lexical bundles less frequently.

This finding implies that it is the frequency of the lexical bundles that is an important cause of the increased proficiency. The findings thus link to notions of frequency effects in language learning (N. Ellis 2002) and have important implications for connectionist models, lexical formulas (i.e. chunks), and implicit learning. For instance, it is likely that the learners, over time, begin to produce lexical bundles at frequencies that begin to parallel L1 speakers because the frequency of these bundles in natural language allow for L2 learners to implicitly make stronger connections between the words in the lexical bundle. As the connections strengthen, it is likely that lexical bundles are then stored as holistic items that allow for ease of access and production. Such a conclusion contradicts the notion that L2 learners focus solely on individual words while native speakers learn formulaic utterances. Previous studies have posited that L2 learners concentrate on single words because they are more easily controlled and provide a sense of proficiency (Wray 2002). This study, however, questions this notion by demonstrating that L2 learners develop stronger lexical connections between co-occurring words based on frequency of use. The development of stronger lexical connections supports the use of probabilistic knowledge as an implicit strategy used in language learning. These connections are evident in the production and growth of bigram accuracy over the course of a year. The first analysis thus shows that L2 learners do not depend solely on single words, but produce native-like lexical bundles and, with time, increase their accuracy of use.
4. Analysis two

Our second analysis is a qualitative examination of the shared bigrams (between learners and L1 speakers) produced by the L2 learners in the first and fourth quarters. We focus only on the first and fourth quarter in order to compare the greatest gains in change. This analysis examines how the bigrams produced by the L2 learners serve to accomplish pragmatic and syntactic functions and how the frequencies of these bigrams increase or decrease as a result of time. To examine these functions, we categorized the common bigrams that were shared by more than one of the L2 learners and that corresponded to bigrams produced by native speakers. These bigrams were next classified based on pragmatic functions as found in Crossley and Louwerse (2007), Biber et al. (2004), and our own classification based on the apparent functions of the bigrams. The classified bigrams were also analyzed for syntactic properties in a manner similar to Crossley and Louwerse (2007), McCarthy and Carter (1997), and Siyanova and Schmitt (2008) who used qualitative methods to categorize word combinations based on their syntactic categories.

Research that has examined pragmatic functions and spoken lexical bundles has demonstrated that frequent lexical bundles that come to inform spoken registers can be used to discriminate spoken registers from written genres. Crossley and Louwerse (2007) demonstrated that spoken discourse differed from written discourse in its use of 26 bigrams. These bigrams served pragmatic functions and included filler phrases (you know) expression of opinions (I think), and questions (is it). In addition, the bigrams that made up spoken registers include syntactic categorizations such as high indices of coordination, first person pronouns, pro-verbs, phrasal quantity adjectives, and phrasal auxiliaries for the future aspect. Crossley and Louwerse’s findings were similar to those of Biber et al. (2004) in which spoken discourse was distinguished from written discourse based on use of first person pronouns, expression of opinion, and filler phrases.

In the longitudinal data collected for this analysis, similar patterns were noted. Specifically, 4 of the 6 learners begin to produce more filler phrases over time and reached a similar frequency to that of native speakers. While filler bigrams such as I mean and kind of occurred in two learners’ transcripts, only the bigram you know was shared across two or more learners (details presented in Table 10).

In addition to the use of filler bigrams, the L2 learners in this study began to produce bigrams for expressing opinion. These bigrams, shared across all learners, included I think and I like. The former, I think, was produced at a rate similar to that of native speakers, while the latter, I like, was produced at a greater rate than native speakers. However, in both cases, the frequency of bigrams between the first and fourth quarters decreased toward native speaker
The development of lexical bundle accuracy and production levels (details presented in Table 11).

L2 learners in this study also used bigrams related to questioning techniques. These bigrams, usually constructed using what, served to promote exchange of ideas and elicit information. The frequency of these bigrams in both the L2 discourse and native speaker discourse were similar and decreasing trends toward native speaker frequency were noted for the bigrams (details presented in Table 12).

The bigram data for the L2 learners also exhibited frequent uses of coordinating conjunctions, especially when combined with first person pronouns (and second person pronouns to a lesser degree). Most of these bigrams were on par with the production frequency of native speakers and most of the frequencies adjusted from the first quarter to the fourth quarter toward native speaker frequencies (details presented in Table 13).

In addition to the pragmatic bigrams noted by Crossley and Louwerse (2007) and the multi-word units discussed by Biber et al. (2004), the data in this analysis also supported additional categorization of bigrams that accomplished functional expressions. The first such categorization was expression of self. These bigrams (I am, I have, I did, and I want) were used frequently by L2 learners to discuss themselves, their experiences, and their desires. These bigrams were also common in the native speaker discourse, but not to the same degree as

Table 10. Filler phrase bigram

<table>
<thead>
<tr>
<th>Bigram</th>
<th>Number of learners</th>
<th>L2 frequency 1st quarter</th>
<th>L2 frequency 4th quarter</th>
<th>NS frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>You know</td>
<td>4 (4)</td>
<td>0.31</td>
<td>0.97</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 11. Expression of opinion

<table>
<thead>
<tr>
<th>Bigrams</th>
<th>Number of learners</th>
<th>L2 frequency 1st quarter</th>
<th>L2 frequency 4th quarter</th>
<th>NS frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think</td>
<td>6 (6)</td>
<td>0.84</td>
<td>0.38</td>
<td>0.21</td>
</tr>
<tr>
<td>I like</td>
<td>6 (5)</td>
<td>0.47</td>
<td>0.12</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 12. Question bigrams

<table>
<thead>
<tr>
<th>Bigrams</th>
<th>Number of learners</th>
<th>L2 frequency 1st quarter</th>
<th>L2 frequency 4th quarter</th>
<th>L1 frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is</td>
<td>5 (6)</td>
<td>0.13</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>What I</td>
<td>2 (2)</td>
<td>0.12</td>
<td>0.07</td>
<td>0.10</td>
</tr>
</tbody>
</table>
Scott Crossley and Thomas Lee Salsbury

Table 13

<table>
<thead>
<tr>
<th>Bigrams</th>
<th>Number of learners</th>
<th>L2 frequency 1st quarter</th>
<th>L2 frequency 4th quarter</th>
<th>L1 frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>And I</td>
<td>6 (6)</td>
<td>0.23</td>
<td>0.26</td>
<td>0.38</td>
</tr>
<tr>
<td>But I</td>
<td>5 (5)</td>
<td>0.33</td>
<td>0.41</td>
<td>0.13</td>
</tr>
<tr>
<td>And they</td>
<td>5 (3)</td>
<td>0.07</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td>And we</td>
<td>2 (2)</td>
<td>0.12</td>
<td>0.21</td>
<td>0.06</td>
</tr>
<tr>
<td>So I</td>
<td>2 (1)</td>
<td>0.21</td>
<td>0.16</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 14. Expression of self

<table>
<thead>
<tr>
<th>Bigrams</th>
<th>Number of learners</th>
<th>L2 frequency 1st quarter</th>
<th>L2 frequency 4th quarter</th>
<th>L1 frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am</td>
<td>6 (6)</td>
<td>0.73</td>
<td>0.79</td>
<td>0.40</td>
</tr>
<tr>
<td>I have</td>
<td>6 (6)</td>
<td>0.25</td>
<td>0.33</td>
<td>0.21</td>
</tr>
<tr>
<td>I did</td>
<td>4 (1)</td>
<td>0.21</td>
<td>0.27</td>
<td>0.10</td>
</tr>
<tr>
<td>I want</td>
<td>4 (4)</td>
<td>0.27</td>
<td>0.39</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 15. Discussion of possibility bigrams

<table>
<thead>
<tr>
<th>Bigrams</th>
<th>Number of learners</th>
<th>L2 frequency 1st quarter</th>
<th>L2 frequency 4th quarter</th>
<th>L1 frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maybe I</td>
<td>4 (5)</td>
<td>0.17</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>You can</td>
<td>2 (2)</td>
<td>0.18</td>
<td>0.28</td>
<td>0.18</td>
</tr>
</tbody>
</table>

The L2 learners. Additionally, the use of these bigrams increased over time in a trend opposite of native speaker frequencies (details provided in Table 14).

L2 learners also used a bigram category that we label "discuss possibilities." The two common bigrams for this function (maybe I and you can) demonstrated that L2 learners use both modals of possibility and adverbs of possibility. However, the frequencies of use for these bigrams were more frequent than those of native speakers. The trend in the more common bigram, maybe I, decreased toward native speaker levels, while the bigram you can did not (details provided in Table 15).

4.1 Analysis 2 discussion

The qualitative analysis found in analysis two demonstrates that the examined L2 learners produced spoken bigrams that had pragmatic and syntactic func-
The development of lexical bundle accuracy and production

19

tions that were similar to those that distinguish native speakers written and spoken text. In addition, the learners in this study produced other bigrams that served pragmatic and syntactic functions. In general the frequency of these bigrams was on par with the native speakers, but many instances of over-production were noted. The over-production is likely the result of L2 learners having less lexical diversity and lexical coverage, producing an outcome similar to over-generalization. If the bigrams were not on par with native speakers, they generally demonstrated a trend toward native speaker levels with the exception of bigrams for expression of self.

From a pragmatic stance, the bigrams produced by L2 learners shared similar pragmatic and syntactic properties with bigrams common to native speakers. This included the use of the filler phrase I mean, which was used by 4 of the 6 learners. However, native speakers, unlike L2 learners, use a much richer diversity of filler bigrams. L2 learners also produced bigrams for expressing opinions. These bigrams, centering on the verbs think and like, occurred at a greater frequency than native speaker production, and, like most examples, were not as lexically diverse. Additionally, L2 learners produced bigrams for asking questions. These bigrams were produced at a frequency similar to native speaker production, but the bigrams were all formed around the pivot word what. Thus, as in the other examples, L2 learners did not demonstrate high bigram diversity. Lastly, L2 learners produced multiple examples of bigrams pivoting around coordinating conjunctions. Many of these bigrams were produced at a similar rate as native speakers, but many were seemingly over-produced. L2 learners, however, did exhibit a high degree of bigram diversity with regard to coordinating conjunctions. These bigrams likely prove important in connecting segments of speech and holding turns.

In addition to the lexical bundles that fit the criteria of Biber et al. (2004) and Crossley and Louwerse (2007), the L2 learners in this study depended on other pragmatic bigrams to produce meaning in their discourse. These included the production of bigrams for expressing self and bigrams for discussing possibility. These bigrams, while not generally paralleled with common bigrams found in native speaker spoken discourse, likely assist in helping L2 learners discuss personal states and needs as well as assisting L2 learners in using politeness strategies and hedges. Moreover, the bigrams produced in both these categories were generated at a greater frequency than found in native speaker discourse and demonstrated trends away from native speaker frequency levels.

Besides the pragmatic implications of most of these bigrams, it is also interesting to note that the majority of common bigrams produced by L2 learners pivot on pronouns, and generally on first person pronouns. This finding is different from past studies of written L2 production which demonstrated that verb base forms combined with objects have the highest multi-word potential (Git-saki 1996; Shei and Pain 2000). While it is likely true that L2 learners produce
many bigrams that pivot on verb forms in written discourse, the bigrams in this study that seemed to have the most pragmatic and syntactic salience pivoted around first person pronouns. The production of first person pronoun bigrams might signal a dependence on speaker-centered discourse.

Results from the qualitative analysis highlight the importance of exploring interlanguage development from the learner perspective and not solely as an approximation to target-like norms. Bigrams pattern in predictable ways across the learners and show many similarities to native speaker oral discourse (Biber et al. 2004; Crossley and Louwerse 2007). At the same time, L2 bigrams are somewhat different in both form and function from native speaker bigrams. We observe some forms that are more frequent than forms produced by native speakers. This finding suggests that bigrams such as the expression of self (I am, I have, I did, and I want) or bigrams for possibility (maybe I and you can) carry a greater pragmatic load in the L2 oral discourse. In the absence of a larger vocabulary, learners use what they know to carry a greater number of meanings. Thus, the qualitative analysis highlights specific phenomena related to interlanguage multi-word lexical units that would be lost with a focus only on target-like forms and functions.

This second analysis, combined with the first analysis, demonstrates that L2 learners begin to produce bigrams at a more accurate frequency as a function of time studying English and English proficiency. It also shows that the bigrams produced serve linguistic functions. Taken together, these two studies provide evidence for lexical growth in both form and in function. Importantly, this growth, while explicit on the surface level, provides evidence for the growth of seldom explored levels of lexical proficiency, namely depth of knowledge lexical features. The findings also strengthen notions about what it means to know and acquire a word. It is not enough to merely know a word and its meaning. To fully acquire a word, one must know, among other elements, how words combine and what combinations are acceptable (Bogaards 2001). The findings of this study provide additional knowledge about L2 lexical growth and suggest what a full understanding of lexical proficiency might entail.

5. Conclusion

This paper has demonstrated that L2 learners increase their accuracy of bigram use over the course of a year. This finding has important implications for the development of L2 lexical networks and lexical acquisition. The study also demonstrated that L2 learners begin to produce bigrams that serve pragmatic and syntactic functions that are similar to those used by native speakers. That these bigrams were not produced at equivalent frequency as native speakers
The development of lexical bundle accuracy and production

is likely attributable to a lack of lexical diversity and coverage on the part of the L2 speaker. Overall, the trends in our data set support the notion that as a learner gains control of a new language, the learner produces word sequence patterns that parallel those of expert users. The acquisition of these acceptable lexical bundles are likely crucial for gaining communicative competence (Hyland 2008).

While the findings from this study help strengthen our knowledge of lexical proficiency and the acquisition of lexical bundles, the study does have limitations. First, the acquisition of multi-word lexical units is the result of many diverse factors including input frequency, complexity, native language influences, accuracy, and syntactic and pragmatic properties (Gitsaki 1996). While this study could not control for all these variables, future studies should consider the roles they perform. Additionally, this study could be faulted for the size of the population sampled. However, the strength of the statistical and qualitative analysis found here does not rely on the number of subjects but the number of data points available for analysis. With 99 data points available, we do not feel that the size of the subject population is a serious limitation. Lastly, it is not inherently obvious that normalizing the frequency of bigrams between L1 speakers and L2 learners allows for a comparison of corpora that are of unequal size or that data from a single individual can be compared to a group. However, we believe that the normalization techniques used in this study provide the best methodology for comparing corpora of unequal sizes and populations.

While we recognize these possible limitations, we nevertheless feel that this study illuminates important aspects related to the development of L2 lexical proficiency. Far from L2 learners relying on single words to convey meaning, this study demonstrates that L2 learners have a variety of lexical bundles that they are able to functionally employ. This study also demonstrates that L2 learners that exhibit increased language proficiency significantly increase their accuracy of bigram frequency over the course of a year. This finding shows that we should not ignore the role that lexical bundles, frequency effects, connectionist models, and implicit learning may play in the development of L2 language proficiency.

Mississippi State University
(a544@msstate.edu)
Washington State University
(tsalsbury@wsu.edu)
22 Scott Crossley and Thomas Lee Salsbury

References


Crossley, Scott, Tom Salsbury, Danielle McNamara & Scott Jarvis. (in press-b). What is lexical proficiency? Some answers from computational models of speech data. TESOL Quarterly.
The development of lexical bundle accuracy and production


24 Scott Crossley and Thomas Lee Salsbury


Hyland, Ken. 2008. As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes* 27: 4–21


The development of lexical bundle accuracy and production 25


