

Pronouncing the Academic Word List: Features of L2 Student Oral Presentations

Rebecca Hincks

Department of Speech, Music and Hearing, CTT

KTH, Stockholm, Sweden

E-mail: hincks@speech.kth.se

ABSTRACT

This paper is an analysis of lexical choices, pronunciation errors, and discourse features found in a corpus of student presentation speech. The speakers were Swedish natives studying Technical English. Particular emphasis is given to the pronunciation of the words most often used in academic texts. 93% of words used in the corpus came from the most frequent 2570 lexemes of academic written English, 99% of all words were acceptably pronounced, disfluencies occurred at relatively stable inter-student rates, and 30% of all new sentences began with the conjunction ‘and’.

1. INTRODUCTION

Swedish university students are fluent speakers of English who generally produce English phonemes without any significant difficulties. Their facility in English can be attributed to a number of factors, such as early instruction following communicative practices [1], exposure to spoken English in the media, the genetic relationship between Swedish and English, and the necessity of becoming fluent in an international language. Pronunciation problems can however arise in the production of multi-syllabic words that are familiar to students mostly in written form. Uncertainty as to the placement of primary stress within the word can lead students to use the avoidance strategies that are easily available to them as fluent speakers. It is important for the development of an academically appropriate vocabulary that students become confident as to the prosodic characteristics of morphologically complex words.

Swedish higher education is coming under increasing pressure to offer degree programs conducted in English, in order to attract students from foreign countries. University administrators and politicians assume that students’ English skills are sufficient to follow such programs, but language experts maintain that it is not clear how well students are able to transfer their excellent receptive skills in English into articulate and precise productive skills. One of the goals of this ongoing study is to examine the

sophistication of the student vocabulary used in oral presentations.

Very little research has been done on the nature of vocabulary in spoken academic English. In a multidimensional discourse analysis of a 2.7 million-word corpus of university English, Biber et al. [2] point out the real need for detailed vocabulary studies of different registers. Research on the lexis used in L2 student writing has been done by Laufer and Nation [3] both of whom would like to see similar studies done on L2 spoken English.

Analysis of oral presentations presents a good opportunity to study spoken L2 English. Presentations are to a large extent monologues, and can therefore be more easily transcribed than dialogues. Presentations can furthermore be an assignment comparable to a written report, requiring research and planning, and finally receiving a mark. This can provide students with the motivation necessary to stretch their productive use beyond the range of their everyday English.

Assessing the quality of oral presentations made in a language-learning context involves a complex evaluation of content, delivery and language. Those assessing the presentation, whether teachers or peers, can easily become blinded by the content and delivery and be left unsure of what to say about the quality of the language. Vocabulary in particular is difficult to assess at any more than an impressionistic level.

This study is part of a long-term project within the framework of applications for speech technology in language learning. Current applications for speech recognition are designed for beginning students of foreign languages [4]. Such applications provide training in and feedback on production of L2 phonemes, and have little to offer most intermediate and advanced language learners. One research question underlying the analysis presented in this paper is in what ways sophisticated speech recognition could help advanced L2 users. Just as writers of today are supported by spellcheckers, perhaps speakers of tomorrow will rehearse oral presentations with the aid of a ‘speech-checker’ able to identify mispronounced words. This sort of application, admittedly futuristic given today’s ASR technology, would

place the computer in the role of a tool rather than a tutor for the language learner [5].

In order to make the recognition work, we need to understand the nature of the phonetic and phonological deviations the speakers are likely to make, the words they are likely to use, the nature of their disfluencies, and the way they use prosody and vocabulary to structure the presentation of information. This study looks briefly at these various aspects with the intention of pinpointing areas worthy of further exploration.

2. BACKGROUND

Within the field of English for academic purposes, work has been done to identify the words most difficult and necessary for learners to know. Working from a corpus of 3.5 million words of cross-disciplinary academic texts, Coxhead [6] produced an academic word list of the 570 word families that appear with a 10% frequency above a 2000-word baseline. Schmitt and Zimmerman [7] studied the acquisition of the derivative forms of these lexical items in a written format. Laufer and Nation [3] used an earlier version of the same word list to develop a method for assessing the lexical richness of student writing. None of these studies looked at the acquisition of the words' pronunciations or their use in oral production.

3. CORPUS COLLECTION

For this study, a corpus of 100 minutes of semi-spontaneous (or rehearsed) student English monologue was collected using a mini-disc recorder and a tie-pin microphone. The recordings were made in the classroom as students made ten-minute oral presentations on technical topics of their choice. There were eleven different speakers, all native Swedish speakers between the ages of 22 and 30. The speech was transcribed orthographically, omitting question-and answer portions of the presentation. The total number of words in the corpus was 13,471. Data was gathered regarding the frequency of the lexical items used, the number and nature of mispronunciations, lexical and grammatical errors, disfluencies, and the vocabulary used to signal change in topic. The transcripts were prepared for lexical analysis by removing filled pauses, hyphenated phrases, unfinished words, and proper nouns.

4. RESULTS

4.1 Vocabulary

Lexical analysis of the transcripts was performed using the program *Range* [8] to quantify and identify words according to their frequencies in the corpus of academic English as determined by [6]. Table 1 shows these results. Words belonging to the one thousand most frequent word families in academic written English account for 85% of the words used by the students. The academic word list, representing the words non-native speakers should know for successful studies at an English-speaking university,

accounted for the smallest share of words, 2.4%. The high

Table 1. Results from *Range*

	Base list number in <i>Range</i>	Frequency in corpus of academic texts	Percent of tokens
	List One	First 1000 word families	85.4%
	List Two	Second 1000	5.3%
	Academic Word List	2000-2570	2.4%
	Not on a list		6.8%

portion of non-list words (6.8%) could be due to the technical nature of most of the presentations, where terms describing e.g. engines or computers were often used.

4.2 Pronunciation

For the purposes of this study, pronunciation errors were defined as phoneme substitutions that could impede intelligibility, and errors in placement of lexical stress. Norell [9] provides useful guidelines as to the features of Swedish-accented English that are perceived most negatively by native speakers. One of the characteristics of Swedish-accented English is devoicing of /z/; however, the /s/-/z/ distinction carries a low functional load and is not among the most negative characteristics according to [9]. In our corpus, the distinction was potentially problematic only in the minimal pair *phase-face*. Devoicing of /z/ was therefore not counted as an error.

Mispronounced words were classified as to whether they were due to a phoneme substitution or presumed ignorance of the word's pronunciation, caused by unfamiliarity or difficult spelling. Examples of the latter would be stress problems: *opPonent* pronounced as *oppoNENT*, or problems stemming from opaque spelling: *weight* pronounced as *white*. Table 2 shows a classification of the types of error.

The most frequent phonemic error was substitution of /d/ for /ð/ in function words. This is a feature of many English dialects and while stigmatised, is not as serious an error as the next-frequent phoneme substitution, which was the substitution of /j/ for /dʒ/. Difficulties with /θ/, /tʃ/ and /w/ accounted for the remainder of phonemic errors, for a total of 50 tokens and 21 types. Errors at the lexical level accounted for 37 tokens and 28 types.

The list of 87 mispronounced words was compared with the list of 'academic' words produced by *Range*. Only 13 tokens coincided, representing 7 types of words. Nine of these mispronunciations were due to phonemic problems and four due to the incorrect placement of lexical stress. These four words were: *access*, *capacity*, *component*, and *contribute*, realized as acCESS, CAPacity, COMponent, and CONtribute.

Table 2. Classification of pronunciation error

Phoneme sub.	/ð/ → [d]	/dʒ/ → [j]	/θ/ → [t], [f], [s]	/tʃ/ → [ʃ]	/v/ ↔ /w/	Total
tokens	20	15	7	4	4	50
types	3	7	5	3	3	21
lexical error	Misplaced stress			Other (e.g. opaque spelling)		
tokens	15			22		37
types	14			14		28
Total tokens					87	

4.3 Morphological and lexical errors

Words used incorrectly, including morphological errors, e.g. incorrect verb forms, and words used in the wrong sense, accounted for 1.5% of the corpus. Thus these students' speech can be characterized as grammatically fairly correct.

4.4 Disfluencies

Of the 448 filled pauses transcribed in the corpus, 90% were transcribed as 'uh' or 'eh', thus as purely vocalic sounds. Only 8% were transcribed as 'um' or 'ehm'. Restarts and repetitions appeared at an average of 2.05% in relation to the number of words, with a low inter-student standard deviation.

4.5 Sentence boundaries

Analysis was made of the words used to begin what was perceived as the start of a new sentence. A total of 666 sentences were transcribed in the corpus, and of these, 200, or 30%, began with the conjunction 'And' or 'and' with a clitic filled pause: 'Eh and' or 'And uh'. The second most common sentence opener was 'So' with a frequency of 11.5%.

5. DISCUSSION

In this study we are attempting to identify the characteristics of non-native student presentations. What are the most salient features? We have seen that while the speech could be characterized as accented, serious pronunciation errors were few. Confusion within the pairs /dʒ/ → /j/

and /w/ ↔ /v/ was present to some extent in the corpus, and these errors are judged negatively by native speakers [9]. The lower number of different types of word containing phonemic errors indicates their recurring appearance in high-frequency words such as *the*, *just* and *very*, indicating potential problems of fossilization for some speakers.

Even more important for the advanced learner is the ability to confidently articulate words that may be more familiar in writing than in speech. We had expected to find more examples of misplaced lexical stress in words from the academic word list. Asking the students to read words from the academic word list would have run the risk that the pronunciation was worse than it otherwise might be, because of discomfort with the reading task. It may be that for these students, problems with stress placement appear for words with a lower frequency than those of the academic word list.

Another primary goal of the study was to look at the extent to which students productively used the words they are exposed to in the texts they read for their coursework. This would test the hypothesis that increased range of input leads to increased range of output. Laufer [10] has found that there is a threshold in the development of active vocabulary, even though passive vocabulary continues to grow. The format of the formal, graded oral presentation, where technical content must be explained, seemed to be appropriate to the goal of requiring students to use their most sophisticated vocabulary.

There is some evidence that students avoided or lacked words:

*They were known as com—as eh good seafarers
because that's the legal—eh, the smallest amount of time*

and some evidence that they tried to use more advanced words:

*and this process will turn—eh convert—eh the sugar
the word comes fro—uh originates from the Latin*

These sorts of revealing last-minute corrections are unavailable to analysts who look at student writing, unless the student is working at a computer logging key-strokes.

Words from the academic word list accounted for 2.4% of the total words in the corpus. Unfortunately, this figure cannot be compared with other studies of lexical frequency in spoken English, because such studies have not yet been conducted. However, [3] looked at academic word list frequencies from writing samples of English L2 university students, finding mean frequencies of 3.6% for lower intermediate students and 8% for advanced students. It is well known that writing is more lexically dense than speech. A possible methodological problem with this study is the suitability of using frequency lists generated from writing on speech, even in formal-register speech. However, until appropriate studies of spoken academic English are con-

ducted, we will need to use the tools currently available.

It is possible that the greatest improvements that can be made in student presentations are at the discourse level. It was surprising that as many as 30% of new sentences were opened by the simple conjunction *and*. It would be interesting to look at what sort of connecting phrases students could use when using their native language to make presentations. Using more filled pauses that end with a bilabial nasal could also help them to sound more native-like [11].

The limited size and consistent pronunciation of the vocabulary used in these presentations presents a potential opportunity for the use of speech recognition as a tool for advanced language learners. If a presentation on a certain technical topic can be expected to use words 93% drawn from the 2570 most frequent words in English, and the remaining 7% drawn from a known list of technical terms, then it ought to be possible to build speech recognition engines that could recognize most of a student's production. With really good recognition, unrecognized words would be indications that either the word has been mispronounced or that it is a low-frequency word, which in turn could be difficult for a non-native audience to understand. In this way, a speech recognizer could be used as a sort of filter for preparing talks for different sorts of audiences. It could also be used in a gaming sense to reward students for making certain lexical choices, providing the necessary incentives for advanced students to push continued growth of their active vocabulary.

In this paper we have been looking at what characterizes the corpus of presentation speech as a whole. However, this corpus was composed of speech from eleven individuals, and data has been calculated by speaker. Inter-speaker variation is another interesting path to pursue, but is beyond the scope of this paper. It can be mentioned, however, that a modest correlation (.49) was found between the percentage of academic word list words used and the global quality of the presentation as marked on a three-point scale.

4. CONCLUSIONS

In this study we find that 93% of non-native student oral production comes from the most frequent 2570 word families of academic English. Estimates of native speakers who have completed a high school education place their vocabulary knowledge in the range of 30,000 words [12]. It would be very interesting to be able to compare our non-native student presentations with native-speaker presentations in order to see to what extent the non-native speakers are using a relatively impoverished vocabulary.

ACKNOWLEDGEMENTS

Thanks to David House, the Centre for Speech Technology, and to students and colleagues at the Unit for Language and Communication at KTH.

REFERENCES

- [1] M. Oscarson, "A national evaluation programme in the Swedish compulsory school," *System*, Vol. 23, No. 3, pp.295-306, 1995.
- [2] D. Biber, S. Conrad, R. Reppen, P. Byrd, and M. Helt, "Speaking and Writing in the University: A Multidimensional Comparison," *TESOL Quarterly* Vol. 36, No. 1, pp.9 – 48, 2002.
- [3] B. Laufer and P. Nation, "Vocabulary Size and Use: Lexical Richness in L2 Written Production," *Applied Linguistics*, vol. 16, no. 3, pp. 307- 322, 1995.
- [4] R. Hincks, "Speech recognition for language teaching and evaluating: A study of existing products," *Proceedings of ICSLP 2002*, Vol. 2, pp. 733-737.
- [5] M. Levy, *Computer-Assisted Language Learning: Context and Conceptualization*, Clarendon Press, Oxford, 1997.
- [6] A. Coxhead, "A New Academic Word List," *TESOL Quarterly* Vol. 34, No. 2, pp.213 –238, 2000.
- [7] N. Schmitt and C. Zimmerman, "Derivative Word Forms: What Do Learners Know?" *TESOL Quarterly* Vol 36, No. 2, pp. 145-171, 2002.
- [8] A. Heatley and P. Nation, *Range* [computer software, version 12, available from www2.vuw.ac.nz/lals/]. Wellington, Victoria University of Wellington, 2002.
- [9] P. Norell, *Native-Speaker Reactions to Swedish Pronunciation Errors in English: Recognition, Intelligibility and Attitude*, Stockholm University: Stockholm Studies in English, 1991.
- [10] B. Laufer, "The Development of Passive and Active Vocabulary in a Second Language: Same or Different?" *Applied Linguistics* Vol. 19 no. 2, pp. 255-271, 1998.
- [11] H. Clark, "Speaking in time" *Speech Communication* Vol. 36, pp. 5-13, 2002.
- [12] D. Crystal, *The Cambridge Encyclopedia of the English Language*. Cambridge University Press, 1995, p. 123.