

Life-long English learning: How lexis-based instruction and use of corpora can help

Stephen Jennings

Introduction

Gaining access to native speaker speech is becoming easier with the advent of Internet resources such as the US adult literacy program, which uses short news reports from CNN, the Voice of America's simple English broadcasts, the BBC's learning English pages; and let us not forget youtube.com. Youtube, in particular, contains a myriad of countless clips of English speaking people in natural situations. However, the problem is that the speech found in all of the above Internet resources is not brokered in an accessible way for the average learner in the EFL context of Japan.

As I am very interested in the use of authentic English speech and writing in the English as a Foreign Language (EFL) context, these research notes aim to give a brief outline for an approach encouraging learners to analyse speech and writing of native speakers and, very importantly, their own speech and writing. Students will, subsequently, become aware of the important notion of Discourse Analysis, through use of an on-line corpus⁽¹⁾.

Many scholars believe that it is very important to have undiluted native speaker dialogue for language learners, for example, followers of *the Natural Approach* (Krashen & Terrell, 1983) and *the Lexical Approach* (Lewis, 1993). It is becoming more apparent, though, that not only is vocabulary acquisition for students of a second or foreign language very important but also, is the study of how to create students who are able to gain vocabulary and a better general command of English, in a more autonomous way. This *autonomy* may be best gained by accessing English words and their collocates in situ from naturally occurring language (for more on using corpora in language teaching see, Biber, Conrad & Rippen, 1998). This is a very complex undertaking; however, the limits of space for these research notes require that I may only briefly outline the approach with which I aim to take towards building the chance for our students to pursue a life-long approach towards English language learning.

Benefits of a Lexis-based Language Awareness approach

This, learning gained by analysing natural speech with the help of corpus-based analysis, is a way for students of English in the EFL environment to have the access to

language that learners living in an English-speaking environment (English as a Second Language – ESL) take for granted.

To that end, and as a way to enhance a life-long view of English learning, learners must, I believe, be shown that every language is lexis-based i.e. it builds *formulaic expressions*. By definition, these expressions appear in language, more than just by chance, and, if attended to in the classroom by using language awareness activities, language improvement can be furthered in autonomous settings outside the formal learning environment (while watching movies, reading books or talking with English speaking friends).

Problems with a Lexis-based Language Awareness approach

A lexis-based language awareness approach may be considered a good approach for EFL learners in the context of Japan; however, there are problems, for both teacher and learner, with becoming aware of language. The basis of a Lexis-based Language Awareness Approach lays in noticing the difference between the learners' own and the target language. A problem may also be found within the lesson materials and tasks, for example, the Input Hypothesis (Krashen, 1982) states that for learners to get valuable learning chances from the lesson, the input received must be structured so learners are able to attend to both its form and content and they must have sufficient control over the language to do this (for more on control in the learning process see Bialystok, 1994). Moreover, according to Vygotsky's (see for example, Lantolf and Appel, 1994) theory of the Zone of Proximal development, by working with language of the correct level for the student their language will improve with interaction with the text, peers and the teacher. Thus, we can see there is a struggle to make authentic English accessible to lower level learners.

There follows, in Section 1.; notes on the general format of introducing; and making authentic-English accessible to lower level learners. Thereafter, Section 2 discusses the basing learning of vocabulary; language in general, and the noticing of patterning in groups of words.

Section 1

Language Awareness

The reason to focus on a holistic or contextualised view of language comes from the Language Awareness Approach, or movement, as it has been called, which provides chances for learners to gain more knowledge about learning strategies.

Content-Based Instruction

Content-Based Instruction (CBI) leads learners to employ certain strategies with the goal of moving the learning responsibility away from the teacher and toward the learner (Rosen & Sasser, 1997). When describing a program in the EFL context, Brinton, Snow and Wesche (2003) describe principles on which successful CBI program is based.

- 1) Theme and content (specified in terms of communicative goals) are of primary importance
- 2) The language exercises are derived from the text rather than imposed on it
- 3) The texts are authentic – they have been produced for and by native speakers of English; in other words , they have not been simplified for pedagogical purposes (p. 34)

The Use of CBI in the context of Japanese Universities

CBI is one of the main underlying curricular components behind the use of authentic theme-based content in a structured, integrated manner. This is due, in the main, to CBI's direct relationship to the use of English as a tool to form the basis of gaining knowledge of the world.

The relationship between CBI and Learner Autonomy

CBI is a means of transfer of knowledge. It follows, then, that there will often be a need for teacher intervention and support to a greater or lesser extent. That is to say, a greater emphasis will have to be placed on the concept of noticing in the language learning process i.e. the teaching of strategies to aid students to pay attention to and reflect on how language is used.

Implications for CBI and the theme-based approach

The learning of English is a complex task. There are a great many variables with which to contend, however, within the limits of this paper, it has been contended that content-based instruction, if properly implemented, can aid improvement in gaining information through English, it tries to ensure a more active role from learners, promoting autonomy and motivation from well-chosen themes. This culmination is widely believed in the SLA community to improve learner language skills.

Section 2

Lexis-based not Grammar-based teaching

Widdowson outlines what he believes an English teaching program in the EFL context should be. It should be one where 'The language exercises are derived from the

text rather than imposed on it [because] it is generally assumed the main concern is the meanings of words as lexical items' (1996: p. 53).

What I believe the learners in the Japanese context must be helped to see is *the way* words are ordered, i.e. *the patterning* of words in a sentence. This is a difficult process because there can be; meanings encoded in words (metaphor), words with more than one meaning, phrasal verbs and other phrases, such as proverbs. For example, lexical items can come as large clusters of words e.g. "(he's) gone and done it again", these are formulaic phrases and are complete units of meaning. The difficulty lies in that, as Widdowson notes, these types of phrases are *compounds* of words; not *combinations* of words, i.e. they cannot be grammaticalised; as a result the whole pattern, or chunk of language, must be noticed if it is to be learned. As mentioned earlier, the poverty of input of authentic language in the EFL environment does not lend itself well to the noticing of such phrases.

Another main difficulty for the learners of second languages is that they do not have a firm grip on one of the most common types of phrase, collocations⁽²⁾. An example could be the word "problem" a native or proficient speaker of English would be able to collocate this word with "serious", where speakers of other languages may have difficulty in finding the most suitable word due to interference from their own mother tongue. For example a learner with an as yet imperfect grasp of English may utter something like "deep problem" or "strict problem". This inability for a learner to find "natural phrasing" makes communicating with non-native speakers harder work because the native speaker cannot rely on what the learner will say next.

It is the relationship between the compounds or phrases, their malleability, and the meaning given to the phrase and how it is most usually arranged, that makes the language to be learned a complex web of connections. Learners need sense to be made from this complex arrangement. From a chaotic jumble, learners need to be shown a method of noticing these habitual co-occurrences (patterns).

How Corpus Linguistics can help

These research notes outline how compounds of words, or phrases, may be critically analysed by students in an inductive way. In other words, set phrases or patterns of words are noticed and reflected upon with an attempt to show students how to analyse language, build vocabulary and, through a process of reflection and productive use, improve their communicative competence. Once this cycle has become second nature, students will be able to notice contextualised lexis (chunks of meaningful language) and how the chunk is used, aiding language acquisition and accuracy. If the

premise is taken that teachers should show learners how to notice language chunking by using inductive classroom tasks⁽⁴⁾, we must turn to the ability for students in the EFL context to retain chunks, phrases and individual words in long term memory.

Bearing in mind it is widely acknowledged to be human nature to forget language that is not continually used, especially in the EFL context, there is a need for post-lesson tasks to reinforce language which has gone through the input stage but has not been accepted as part of the learners own repertoire of English. Consequently, there is a need for students to be able to use a searchable database of the language they have had access to in lesson materials. If the language used in a textbook were to have a searchable corpus, it would provide an extra dimension to language learning, to the students in the EFL context.

English Language Corpora

A computer is an extremely useful resource in holding vast sums of data and the ubiquitous way to gain access to such data is through the Internet. Unfortunately⁽³⁾, the Internet is analogous to a large city, and to navigate safely through all of the information on offer; with its one way streets, dead ends (and all its other diversions) takes a skilled information broker for teacher. It stands to reason, like anyone, a learner will need guidance about where to find appropriate information at the correct level to be able to learn from it.

Large corpora of English written and spoken English exist, where a skilled learner can make use of lists of contextualised vocabulary. The way in which learners are shown how to use a corpus and the actual analysis of text; to promote autonomous learning through a guided analysis of lexical items; and through being helped to notice certain language “patterns”, will be outlined below.

One of the main reasons for using corpora as a way to promote a bottom-up, lexis-based or word or phrase driven approach to learning in the EFL context is that

bilingual dictionaries do not record ... complex units of meaning in a methodical way ... that is why they do not help us to translate into a language we are not familiar with (Teubert, and Cermakova in Halliday et al. p.123)

As a remedy to this, after training, learners will be able to use a corpus to find out where a pattern of English usage is uttered or, if they are not sure how to use a particular phrase, the student can look up the word in a corpus. The student then may search through the corpus to get examples of how the word is most usually used in a similar sentence, changing the original after noticing the difference between their own work and that of a native speaker.

English Language Corpora and the textbook

If the premise is taken that it is a good idea for textbooks in the EFL context to have easy access to language corpora; textbooks in the EFL context should then have access on-line to the searchable database. Students will then have access to a source of English when they want to plan a certain speech act, or check back upon, when they are unable to smoothly communicate. Unfortunately, the complexity of speech utterances that may be undertaken can never be held inside a textbook. Therefore, students must be shown certain strategies for coping with occasions when impedance to communication occurs.

Future textbooks must surely have their own on-line corpus if they are to become more representative of the type of *access to knowledge* students in the EFL context of Japan are becoming more and more familiar i.e. in the on-line medium. Textbooks must take into account both learning strategies and pre-empt the types of sentence structures deemed useful to a particular group of students, bearing in mind the target language environment goals.

Students in the EFL learning environment, it is hoped, would have high motivation for the process of critical thinking because with help from a corpus of language to which they have already accessed, a communication difficulty would become a puzzle to solve rather than an insurmountable wall. A motivated learner, would go through the process of noticing the difficulty, thinking critically, and planning, or re-casting; the learner would then be on the way to becoming more aware of language and how to solve errors in an autonomous way.

Errors in Writing and Speaking

There is a difference in the way such errors need to be approached in the fields of Writing and of Speaking:

In the field of Writing, this would be most likely done with teacher intervention e.g. a piece of writing may be pointed to by use of an underline and a short annotated reason for the error (Bologna, Jennings & McCrostie, 2005). Students must then be allowed to re-cast their ambiguous sentence in order for the meaning to be negotiated again.

In the field of Speaking, a student would not have ready access to a computerised corpus of language. At instances of communication breakdown, learners would have to re-cast their sentence or negotiate the meaning of their utterance by using body language combined with gesture, inflection, pitch or even draw or spell out a word, or sentence⁽⁵⁾. In the case of speaking then, learners would best make use of an on-line

corpus by pre-empting a language breakdown by considering what to say before they say it, or remember what they said and search for a more appropriate instance of this on line. Of course, this is not always possible. Therefore, learners would need to use the strategies for communication learned in the language awareness component of the course.

If this process of gaining better control of learning strategies can be continued, it follows, that a learner will be able to build on his/her level and gain a better level of control of his/her English thereby a larger repertoire of accurate English. They will become more proficient writers and speakers. They will attain more accurate English causing less miscommunication.

Conclusions

These research notes are an attempt to show the sound theoretical underpinnings which lay beneath the reasons for use of authentic theme-based native speaker; speech, and writing in class materials, in the EFL context of Japan. To that end, the concepts of language awareness, learner autonomy and content-based instruction have been outlined. Moreover, it has been shown how focussing on lexis instead of form and accessing a corpus to note patterns of language; is a worthy manner of instruction. This author believes that accessing a corpus of language found in a textbook is a way to bring the notions outlined above in a useful way.

The context of the research note format is limiting, however, I believe the concept and application of noticing-based activities will show students very important language learning strategies that remain relevant and motivating during college life and into the wider world after graduation.

Notes

(1) A corpus is defined as ‘a collection of naturally language texts in electronic form, often compiled by according to specific design criteria and typically containing many millions of words’ (Halliday et al: p.168). An example is the Collins Concordance Sampler (available online at <http://www.collins.co.uk/Corpus/CorpusSearch.aspx>).

However, learners need training to be able to use the corpus effectively. Another more accessible corpus is the google search engine

(see: <http://googleresearch.blogspot.com/2006/08/all-our-n-gram-are-belong-to-you.html> for more information)

(2) Collocation is defined by Halliday, Teubert, Yallop and Cermakova (2004, 168) as ‘the habitual meaningful co-occurrence of two or more words in close proximity to each other; as a lexical relationship’

- (3) As with learning English itself, learning a foreign language means decoding a complex web of connections
- (4) Research has shown that students in the EFL context of Japan, students find gaining language awareness skills fresh, practical and motivating (Falkus and Jennings 2004).
- (5) The miscommunication may be that of a more cultural misunderstanding, in such cases miscommunication is more difficult to solve, and may not be accessed readily in a corpus (see Scollon and Scollon)

References

- Biber, D., Conrad, S. & Rippen, R. (1998). *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge University Press. Cambridge.
- Bialystok, E. (1994) Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16, 157–168.
- Bologna, D, Jennings, S. & McCrostie, J. (2005). Student perceptions on the Benefit of Error Correction. (Vol. 7/1. pp. 47-62). *Asia Pacific Journal of Language in Education*.
- Brinton, D., Snow, M., and Wesche M. (2003). *Content-based Second Language Instruction*. The University of Michigan Press: Ann Arbor.
- Dornyei, Z. and Schmidt R. (2002). *Motivation and Second Language Acquisition*. University of Hawaii at Manoa: Hawaii.
- Falkus, J., & Jennings, S. (2005). Helping Learners Reduce Japanese English. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT 2004 Conference Proceedings*. Tokyo: JALT.
- Grabe, W. & Stoller, F. (1997). Content-based instruction: Research foundations. In M.A. Snow & Brinton, D. M. (Eds.) *The content-based classroom: Perspectives on integrating language and content*. (pp. 5-21). New York: Addison Wesley Longman.
- Halliday, M. Teubert, W. Yallop, C . & Cermakova, A. (2004). *Lexicology and Corpus Linguistics*. New York, London: Continuum.
- Jennings, S. (2004). English derived loanwords in Japanese: The changes they undergo and how they interfere with Japanese EFL learners' pronunciation. (Vol. 15 pp. 557-571). *The Journal of Kanda University of International Studies*.
- L.F. Kasper (Ed.), *Content-based college ESL instruction*. Mahwah, NJ: Lawrence Erlbaum.
- Kinsella, K. (1997). Syllabus design in content-based instruction. In M.A. Snow & Brinton, D. M. (Eds.), *The content-based classroom: Perspectives on integrating*

language and content (pp. 132-141) New York: Addison Wesley Longman.

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.

Krashen, S.D. & Terrell, T.D. (1983). *The natural approach: Language acquisition in the classroom*. London: Prentice Hall Europe.

Lantolf, J., Appel, G. (1994). *Vygotskian Approaches to Second Language Learning*. Westport, Connecticut: Ablex Publishing Corporation.

Lewis, M. (1993). *The Lexical Approach*. Hove, UK. Language Teaching Publications.

Rosen, N.G. & Sasser, L. (1997). Sheltered English: Modifying content delivery for second language learners. In M.A. Snow & Brinton, D. M. (Eds.), *The content-based classroom: Perspectives on integrating language and content* (pp. 35-68). New York: Addison Wesley Longman.

Schmidt, R. (1995). *Attention and Awareness in Foreign Language Learning*. (pp. 1-65). Honolulu, Hawai'i: University of Hawaii Press.

Scollon, R. & Scollon, S. (2003). *Intercultural Communication: A Discourse Approach (second edition)*. Oxford: Blackwell.

Straight, S.H. (1997). Language-based content instruction. In Stryker, Stephen B., & Leaver, Betty Lou (Eds.) *Content-based Instruction in foreign Language education* (pp. 236-260). Washington, DC: Georgetown University Press.

Widdowson, H.G. (1996). *Linguistics*. Oxford University Press.

要旨

英語教育における生涯学習： 語彙中心の教授法とコーパスの使用の利点

ジェニングズ スティーブン

本研究ノートは、英語を母国語とするネイティブスピーカーの会話を教材とした語彙中心の教授法によって、どのように学生が言語を分析し得るようになるかを述べたものである。この分析とは、言語学分野におけるいわゆる「言語意識」と「コーパス言語学」に関わる知識を蓄積していくことであり、言語分析は、学習者を以下のように導く。

- 1 学習することへの動機づけを促す
- 2 学習において、より自律、独立する

この2点における可能性を高めていくことで、学生は、卒業後も自身で英語学習を続けられる生涯学習の糧となるのである。