

Do Very Advanced Users of English Accurately Assess Their Own Lexical Knowledge?

Paul Garside

Abstract: The ultimate goal of second or foreign language learning, albeit one which remains elusive for many learners, is to be able to use the language in an authentic context. In terms of reading, perhaps the largest impediment to understanding authentic texts, certainly once the rudiments of grammar have been acquired, is an adequate knowledge of vocabulary. In order to compensate for this, learners need to be able to look up unknown words, or infer (or guess) meaning from context. However, learners have to be able to accurately assess their own lexical knowledge in order to know which words to focus on in this way. Building on an earlier study by Laufer and Yano (2001), I argue that, although better able to do so than slightly lower proficiency learners, even very advanced users of English tend to overestimate their own lexical knowledge when viewing items in context. This implies that guessing or inferring from context alone should not be relied on as a strategy for dealing with unfamiliar vocabulary. In pedagogical terms, this suggests that a program of extensive reading should be integrated with and conducted alongside intensive reading and direct vocabulary study.

Keywords: *vocabulary, inferring from context, extensive reading*

Introduction

The ultimate goal of second or foreign language learning, albeit one which remains elusive for many learners, is to be able to use the language in an authentic context. In terms of reading, this means being able to comprehend original texts that are not made with language learners in mind. There are many factors that stand in the way of comprehension, including knowledge of grammar and syntax, discourse norms, and cultural references. Yet perhaps the largest impediment to understanding, certainly once the rudiments of grammar have been acquired, is an adequate knowledge of vocabulary.

It is clearly unrealistic for second language learners to know all of the words they encounter in authentic texts. This, therefore, necessitates the use of compensatory strategies, such as looking up unknown words in a dictionary, and guessing or inferring from context. In fact, it has been argued that a combination of both of these strategies is optimal (Nation, 2013). However, in order for this to be successful, learners must first be able to notice unknown words in a text, and also to accurately gauge their ability to infer meaning (Laufer and Yano, 2001). If these preconditions are fulfilled, dictionary use (or

some other means of checking meaning) can be employed for words that remain unclear. On the other hand, if learners misjudge their ability to infer meaning then this could lead to serious misunderstandings within the text. In other words, recognizing what one *doesn't* know, as well as what one does, is a critical aspect of second language learning.

Literature review

It is notoriously difficult to estimate the amount of words known by native speakers of a language. This is complicated by the fact that individuals vary considerably in their knowledge of vocabulary. Nevertheless, even modest estimates place the amount of word families known by native speakers of English at around 20,000, or roughly 32,000 individual lexical items (Nation, 1990). This, of course, is complicated by the issue of what knowledge of a word actually entails. Knowing a word consists of several aspects, including form, collocations, grammatical function, and constraints of use, but at its most fundamental it consists of the relationship between form and meaning. This is the aspect to which this study pertains.

In order to attain a comfortable level of unsupported comprehension, Hu and Nation (2000) maintain that ideally 98% of words in a text need to be already known. Another way of stating this is that no more than about one word in 50 should be unknown to the reader. Such a level of text coverage, as well as facilitating overall comprehension, gives the reader enough context clues to have a reasonable chance at guessing or inferring the meaning of the remaining words.

For these unknown items, guessing or inferring from context has been suggested as a way of compensating for a lack of vocabulary knowledge, particularly in extensive reading (Day & Bamford, 1998). Such incidental learning is undoubtedly how native speakers learn the majority of their vocabulary, although whether this is also the best way for second language learners to acquire vocabulary is a matter of some debate. Even optimistic views suggest that multiple encounters with a word in meaning-focused context are required before the meaning is well understood, such is the cumulative process involved in the incidental acquisition of many vocabulary items (Nation, 2013). Other researchers have argued that vocabulary development is more successful when learners are required to actually focus on such words in specific activities, including productive tasks (Hulstijn & Laufer, 2001).

Moreover, guessing from context is by its very nature an imprecise exercise. Even if learners are motivated to look up unknown words in a dictionary, how can we be sure that they choose appropriate items to do so? In other words, how can we be certain that learners know what they think they know? One such attempt to ascertain the accuracy with which learners assess their own knowledge of vocabulary items in context was undertaken by Laufer and Yano (2001). Their study focused on advanced-level learners from three countries: Japan, China and Israel, with one of the aims being to compare the results of these three nationalities. They found that learners from all three countries,

regardless of gender, overestimated their knowledge of selected items, although the Japanese learners did so less than their Chinese and Israeli counterparts. This was the case both before and after they were required to explain or translate the items.

This study

The study is based on that of Laufer and Yano mentioned earlier, and follows the same methodology. There are some important differences, however, meaning that it is less an attempt to replicate the earlier one than to contribute some additional information to the question of how well learners evaluate their own lexical knowledge.

First, and most significantly, this study looked at only Japanese speakers of English; therefore, no cross-cultural comparison can be conducted. Second, whereas the original study investigated advanced learners of English, the participants in this study can be described as very advanced in their English ability. In fact, all ten of the participants use English on a daily basis, either in their personal or professional (or in some cases both) lives. Therefore, it was decided to adapt the original text slightly, in order to increase the level of difficulty. This was done by substituting some of the selected items for more challenging vocabulary, and also by removing some of the context clues (see Appendix A for the adapted text used in this study). This was done in order to set what was felt to be a suitable level of challenge for the participants in terms of inferring the meaning of unknown, or partially known, vocabulary items.

Further, the participants in this study are older than the university students investigated by Laufer and Yano. Although precise ages were not requested, they are thought to range from late 20s through 40s. One of the conclusions tentatively suggested in the original study was that the Westernization of younger Japanese people had caused them to become less humble than older generations, which may have been a factor in the overestimation of vocabulary knowledge. It is beyond the scope of this study to investigate such potentially deep cultural shifts, although the fact that the subjects are more mature might provide some additional insight into this issue.

Seven females and three males participated in this study, which is insufficient to draw any conclusions based on gender. Because of such differences between Laufer and Yano's original study and this one, the research questions are necessarily different. Nevertheless, it is felt that, by following the same methodology, some useful comparisons can be made between the two sets of data. These comparisons are set out below.

Research questions

The specific research questions in this study were as follows:

1. Do very advanced users of English assess their lexical knowledge more accurately than advanced learners?
2. Do more mature Japanese users of English assess their lexical knowledge more accurately than Japanese university students?

Procedure

The same methodology was used in this study as in Laufer and Yano's original study, therefore only a brief outline will be given here.

Twenty lexical items were selected from a modified version of the original text, and highlighted for the purposes of investigation. The target words were listed on a separate sheet (see Appendix B) and the participants were asked to indicate whether they did not understand (0), understood approximately (1), or fully understood (2) each item, having seen them in the text provided. They were then given a separate sheet and asked to explain each item in English, or translate it into Japanese (see Appendix C). Access to the text was also permitted at this stage. Following this, the participants were asked to rate their understanding of the words once more, using an identical self-assessment sheet as in the first stage. The purpose of this stage, according to Laufer and Yano, is to find out if self-perception is altered by a task that necessitates the production and demonstration of knowledge. The vocabulary translations and explanations were later rated by the researcher, working together with a native speaker of Japanese, according to the following scale: left blank, mistranslation or incorrect explanation = 0 points; approximately, but not fully, correct = 1 point; correct translation or explanation = 2 points.

Each participant then received a score for each of the three comprehension tests (two self-assessments and one translation or explanation). The maximum possible score for each test was 40 (20 words x 2 points).

Results and discussion

Table 1 (below) presents both the *real* (i.e., externally-rated) and the *self-given* scores of the participants, as well as the difference between the two following each self-assessment test. The *real score* is the score awarded for the translation/explanation test. Although this score was given only once, it is shown twice for reasons of clarity. The mismatch value shows the difference between the real and self-given scores.

Table 1
Self-Assessment Versus Real Knowledge (Current Study)

	1 st self-assessment (maximum = 40)			2 nd self-assessment (maximum = 40)		
	Mean	Min	Max	Mean	Min	Max
Real score	22.7	6	37	22.7	6	37
Self-given score	24.7	8	36	25.7	3	35
Mismatch	2.0 9% higher	-6	9	3.0 13% higher	-8	9

Table 2 (below) presents the real scores, the self-given scores, and the differences (mismatch) between the two, as recorded in the original Laufer and Yano study. The results are shown for each of the three nationalities represented in that study.

Table 2
Self-Assessment Versus Real Knowledge (Original Study)

	1 st self-assessment (maximum = 40)			2 nd self-assessment (maximum = 40)		
	Japanese	Israelis	Chinese	Japanese	Israelis	Chinese
Real score	14.73	15.5	18.15	14.73	15.5	18.15
Self-given score	19.76	27.15	25.73	19.05	26.78	25.38
Mismatch	5.03 34% higher	11.65 75% higher	7.58 42% higher	4.32 29% higher	11.28 73% higher	7.23 40% higher

The first research question was: *Do very advanced users of English assess their lexical knowledge more accurately than advanced learners?* In order to answer this question, it is necessary to compare the results of the current study with those of the original. Although the very advanced users of English in the current study did, on average, overestimate their lexical knowledge slightly, they did so substantially less than the advanced learners in the original study. The average overestimation of the former group was 9% in the first self-assessment, and 13% in the second self-assessment (conducted after the translation/explanation task). The overestimation in the latter group ranged from 34 to 75% in self-assessment one, and from 29 to 73% in self-assessment two, depending on the nationality of the learners.

Laufer and Yano claimed that very advanced learners (as well as beginning learners) tend to be more aware of what vocabulary they don't know and are hence less likely to overestimate their lexical knowledge when reading. This claim is clearly borne out by the results of the current study.

The second research question was: *Do more mature Japanese users of English assess their lexical knowledge more accurately than Japanese university students?* In order to answer this question, the results of the current study need to be compared only with those of the Japanese learners from the original study. On average, those university students overestimated their knowledge by 34% in the first self-assessment and by 29% in the second. Although this is less than the overestimation of the Israeli and Chinese students, it is clearly more than the very advanced users of the current study, who overestimated by only 9% and 13% respectively, as mentioned above. Therefore, the answer to this question can also be answered affirmatively.

Laufer and Yano expressed surprise that the Japanese learners in their original study had overestimated their lexical knowledge, as they felt it contravened “the cultural tradition of a more humble self-evaluation” (p. 560). It was suggested that this might have been due to a process of Westernization among young Japanese people (Yano, 1997). Applying cultural interpretations in linguistics is a notoriously thorny issue, however, and as such this claim warrants further attention. Rather than a lack of modesty among learners, it may be that the overestimation among all groups can be accounted for by task-related factors. Dealing with an unfamiliar word in context does not require a deep understanding of the item; rather it requires an approximate understanding, so that overall comprehension of the text can continue unimpeded. Therefore, if the reader feels that she has understood the sentence, perhaps she also

feels that she has generally understood the meaning of the unfamiliar word. Decontextualizing the word and explaining or translating it, however, calls for a different, more precise kind of understanding. Consequently, one interpretation of these results is that it is simply more difficult to accurately explain or translate a word than to gain a general understanding of it from context, to the extent that overall comprehension of the text is unimpeded. This difficulty can be partly explained by the lack of a precise one-to-one correspondence for many words, either when searching for a translation in the L1 or a synonym in the L2. This difficulty also extends to giving a clear definition for many terms, either in the L1 or L2 (Nation, 2013). Moreover, it introduces an element of subjectivity to the marking of such a task. These difficulties offer at least an alternative explanation for the apparent overestimation of lexical knowledge in general.

Conclusion

The results of this study suggest that very advanced non-native users of English do indeed assess their lexical knowledge more accurately than those of a slightly lower proficiency. They also suggest that more mature Japanese learners do so with more accuracy than Japanese university students, although this may be due, at least in part, to the difference in proficiency level. Despite my scepticism that such overestimation can be clearly attributed to factors other than task difficulty, the fact remains that differences *were* found between the nationalities in the original study, and between the participants in the original and current studies. It is, therefore, possible that cultural or attitudinal factors play some role in this discrepancy. Further studies are undoubtedly required before more concrete conclusions can be drawn. It would also be useful to ascertain whether overall text comprehension was successfully maintained, despite the incomplete knowledge of certain lexical items revealed by the translation/explanation task.

The key point made by Laufer and Yano in their original study was that guessing or inferring from context alone should not be relied on as a strategy for dealing with unfamiliar vocabulary, because it requires an accurate assessment by learners of their own ability to do so. This implies that direct vocabulary study is also necessary, in order to counter any such misunderstandings. In terms of course design, this justifies supplementing a program of extensive reading, such as that recently implemented at Rikkyo University, with more intensive reading and vocabulary-based activities. Regardless of cultural or attitudinal factors, this integrated approach can also help learners to raise their vocabulary knowledge to a level whereby such misunderstandings are likely to be reduced, as seen in the results of this study.

References

- Day, R. & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, UK: Cambridge University Press.
- Hu, M. & Nation, I.S.P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13(1), 403-430.

- Hulstijn, J. & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning* 51(3), 539-558.
- Laufer, B. & Yano, Y. (2001). Understanding unfamiliar words in a text: Do L2 learners understand how much they don't understand? *Reading in a Foreign Language* 13(2), 549-566.
- Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Cambridge, UK: Cambridge University Press.
- Nation, I.S.P. (2013). *Learning vocabulary in a foreign language* (3rd ed.). New York, NY: Heinle and Heinle.
- Yano, Y. (1997). Contact linguistics to contact cultural studies: Some western influence on modern Japanese and its use. In W. Wolck & A. De Houwer (Eds.), *Recent studies in Contact Linguistics* (pp. 474-80). Bonn: Dummler.

Appendix A: Text

Most children take maleness or femaleness as their first identification of themselves. But once this identification is made, the growing child then begins to compare itself not only in physique, but even more importantly in IMPULSE, with those ABOUT it.

Are all of its interests those of its own sex? "I'm a boy, but I love colour, and colour is something that interests only women." "I'm a girl, but I'm FLEET of foot and love leaping." "I'm a boy, but I love to RUN materials THROUGH my fingers." Such a child is forced to reject such parts of its particular biological inheritance as conflict ACUTELY with the sex stereotype of its culture.

Moreover, a sex stereotype that DECREES the interests and occupations of each sex is usually not completely lacking in VERACITY. The stereotyped idea of the male or female in a given society may ADHERE very closely to the temperament of a particular type of male or female. Those who do not belong to these preferred types are DOOMED throughout life to feel less a man, or less a woman, simply because the cultural ideal is based on a different PARADIGM.

We can understand what is meant by different cultural MORES if we look at the roles of males and females in the Tchambuli tribe in New Guinea. In Western society, women are supposed to be ORDINARILY passive and men active, powerful, BELLICOSE and achieving. In the Tchambuli, it is the women who have the real position of power in society. Men never fish unless a sudden school of fish appears in the lake, upon which they may leap into canoes in a FROLICHSOME spirit and spear a few fish. But the real business of fishing is controlled entirely by the women. Moreover, the women control the money, even though they may permit the men to do the shopping, both for food at the market and in trading in mosquito bags.

The men make a GALA occasion of these LATTER shopping trips. When a man has the final negotiations for one of his wives' mosquito-bags IN HAND, he goes off RESPLENDENT in feathers and shells to spend a delightful few days over the transaction. He will take his time and enjoy SCRUTINIZING what he might buy in the same way as a modern woman with a well-filled purse enjoys shopping in a big city. But the Tchambuli male can only spend the money that he brings back with the ASSENT of his wife.

Appendix B: Self-assessment

Name _____

Look at the following words. They are written in capital letters in the text. For each word, say whether you understand its meaning as used in the text, or not. Rate your understanding on a scale: 0 – don't understand, 1 – understand approximately, 2 – fully understand.

Circle the appropriate number for each word.

Word	Word understanding		
1. impulse	0	1	2
2. about	0	1	2
3. fleet	0	1	2
4. run through	0	1	2
5. acutely	0	1	2
6. decrees	0	1	2
7. veracity	0	1	2
8. adhere	0	1	2
9. doomed	0	1	2
10. paradigm	0	1	2
11. mores	0	1	2
12. ordinarily	0	1	2
13. bellicose	0	1	2
14. frolicsome	0	1	2
15. gala	0	1	2
16. latter	0	1	2
17. in hand	0	1	2
18. resplendent	0	1	2
19. scrutinizing	0	1	2
20. assent	0	1	2

Appendix C: Translation/Explanation

Name _____

Explain or translate the following words in English. Look at the text to see what the words mean in the text and explain or translate (or both) them accordingly.

Word

Word understanding

1. impulse
2. about
3. fleet
4. run through
5. acutely
6. decrees
7. veracity
8. adhere
9. doomed
10. paradigm
11. mores
12. ordinarily
13. bellicose
14. frolicsome
15. gala
16. latter
17. in hand
18. resplendent
19. scrutinizing
20. assent
