Vocabulary Learning and Instruction
Volume 1, Issue 1, August 2012
http://vli-journal.org

The Frequency Model of Vocabulary Learning and Japanese Learners

Dale Brown
Nanzan University
doi: http://dx.doi.org/10.7820/vli.v01.1.brown

Abstract

The frequency model of vocabulary learning, the idea that words are learnt broadly in order of their frequency, is routinely applied in language teaching, testing and research. There has, however, been little research actually confirming it. This paper reports on a small-scale study which investigated the extent to which the vocabulary knowledge of a group of Japanese university students follows the model. Forty-nine low-intermediate proficiency participants took a 100-item Yes/No test containing 20 words from five frequency bands. It was found that, as expected, for the group as a whole, knowledge of the words was related to frequency, with the proportion of words known falling as frequency declined. Following Milton's study, however, an analysis of the results of individual participants revealed that around 20% showed different patterns of knowledge and did not follow the frequency model. One question, however, is whether the frequency information on which the Yes/No test is based is the best approximation of these learners' experience of English. A re-examination of the data in terms of JACET8000 levels found that once more the group as a whole followed the frequency model, and in addition that more of the individual participants conformed to the model. The study thus demonstrates the importance of using frequency data that is relevant to the learners in question and provides confirmation that frequency is a key determiner in the learning of vocabulary. It seems that teachers can reasonably make use of frequency information to support their students' learning.

Keywords: vocabulary learning; word frequency; frequency model; JACET8000.

1 Background

The frequency model of vocabulary learning is the idea that words are learnt broadly in order of their frequency. This model is central to ideas about principled vocabulary teaching, is a key consideration in the design and writing of textbooks and is used on a daily basis by teachers as they consider which words are likely to prove challenging in their classes, even if the only frequency resource used is their own intuition. In vocabulary testing, the frequency model is involved in practically every well-known test: indeed it is hard to think of a vocabulary test that does not take account of the model in some way, whether in the selection of items, controlling the difficulty of contexts or analysing responses. In vocabulary research, and indeed language learning research more generally, frequency is routinely controlled for in accordance with the model. More generally, frequency is recognised as a key driver of all aspects of language learning, and indeed of human learning in general (Ellis, 2002).

Please cite this article as: Brown, D. (2012). The frequency model of vocabulary learning and Japanese learners. *Vocabulary Learning and Instruction, 1*(1), 20–28. doi: 10.7820/vli.v01.1.brown

Perhaps, because the idea of the frequency model is so ubiquitous and seems so obvious, there is surprisingly little research actually confirming it. Milton (2007) traces the idea back to Palmer (1917) and notes how other authors have accepted it without question. Nation (2006b) cites Read (1988) and Laufer, Elder, Hill, and Congdon (2004) as having found that learners' scores drop on tests as you move from higher to lower frequency levels; while Schmitt (2010) also cites Read along with Schmitt, Schmitt, and Clapham (2001) similarly. None of these studies, however, set out to explicitly examine the frequency model.

Two studies that did examine the frequency model are Aizawa (2006) and Milton (2007). Aizawa tested 350 Japanese university students on their knowledge of items in the eight 1000-item levels of the JACET8000 list (JACET, 2003). Aizawa's results (Figure 1) show the frequency model working well for the most frequent bands, as the number of items known declines step-by-step over the first four levels. Beyond this, however, other factors seem to come into play.

One explanation for Aizawa's results is the nature of word frequency. Figure 2 shows a frequency curve, based on data from the Corpus of Contemporary American English (Davies, 2008). As the curve makes clear, the items comprising the first 1000 lemmas are much more frequent than those in the second thousand, but as we move further to the right, the differences between frequency bands become much smaller.

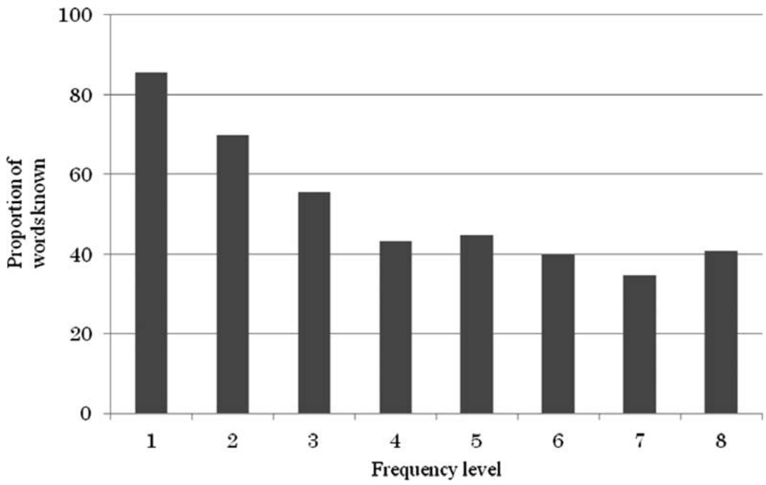

Figure 1. Aizawa's (2006) findings for the proportion of words known at eight JACET8000 frequency levels.

Milton (2007) similarly looked at the vocabulary profiles of learners, using the *X_Lex* test (Meara & Milton, 2003) with 227 Greek learners of English. He found that in terms of the overall results the frequency model worked well, but looking at individual results revealed a more complex picture. Around 60% of individuals produced profiles in line with the frequency model, while the remainder did not. Two alternative profile patterns were identified: a level 2 deficit, in which $1 > 2 < 3$, which accounted for around 25% of participants; and a structural deficit, in which $1 < 2 > 3$, which accounted for around 10%.

The above studies provide us as teachers and researchers with useful information on the frequency model. Aizawa (2006) shows that there are limits to the frequency model, but it works well over the first few thousand words; Milton

(2007) reveals that the model applies to the majority of learners, but not to all learners. This finding is important from a pedagogical perspective as it could mean that assumptions regarding learners' knowledge are faulty for a sizable number of students. The current study, which is a small-scale partial replication of Milton's work, thus seeks to look into how well the frequency model applies to a group of Japanese university students.

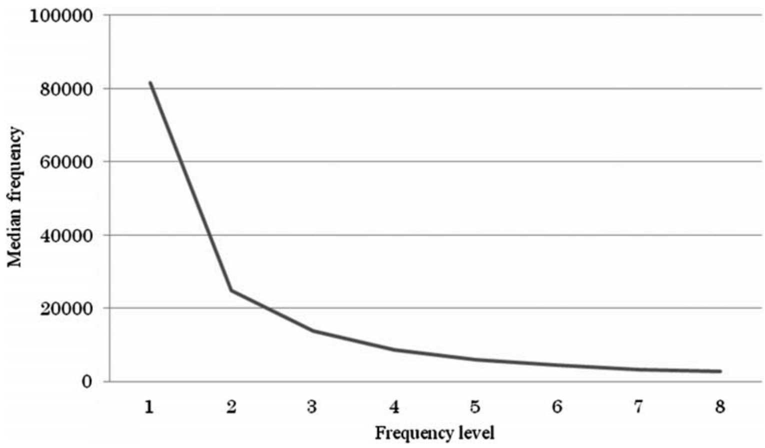

Figure 2. Median frequency of samples of items in COCA. The sample includes every seventh word from the most frequent 8000 lemmas in COCA. Data downloaded from www.wordfrequency.info.

2 Method

The data in this study were collected as part of a larger project looking at the lexical problems faced by learners when reading. Data were collected from a highly homogenous group of 49 learners, all non-English majors studying in two different, but related, departments at a Japanese university. All have progressed through the Japanese school system, reaching similar levels of general academic achievement and similar levels of English proficiency.

The 49 participants took a Yes/No test created for the wider research project. This test included 20 words from each of the first five 1000-word family bands in Nation's (2006a) BNC-based wordlists. The 20 words from each list were selected by first placing the headwords from each 1000-item list in alphabetical order, numbering them, and selecting every 50th word with the count beginning at a different number between 1 and 50 for each of the five lists. In addition to the 100 real words, 20 false words were taken from Milton's (2009) first sample *X_Lex* test. These were included in order to encourage the learners to think carefully about whether they knew each word. The 120 items were presented in random order. The test had good reliability (Cronbach's $\alpha = 0.82$) and relatively few false words were selected, an average of 1.49 per participant, for a false alarm rate of 7.45%.

3 Results

Table 1 provides descriptive statistics for the Yes/No test. With regard to the results in terms of the frequency bands, the group results reveal the expected pattern whereby the first 1000 items are best known and each subsequent band of words is less well known (Figure 3, Table 2). There is a reasonably consistent fall in the number of words known between each level, with only level 4 appearing to deviate from the general pattern. Statistical tests confirm these impressions. As the results for the fifth frequency level were not normally distributed, the non-parametric Friedman's test was used and found the differences between the levels to be significant ($x^2 = 171.42$, asympt sig $= 0.000$). Wilcoxon Signed-Ranks tests were used as post hoc tests to compare the differences between the frequency levels. The alpha was set at 0.05, adjusted with a Bonferroni correction to 0.005. All the bands were found to be significantly different from each other with the exception of levels 3 and 4.

Table 1. Participant Results ($k = 100$, $n = 49$)

Mean	SD	Range	Low	High	Reliability
76.10	7.853	43	49	92	0.822

Note: Reliability = Cronbach's.

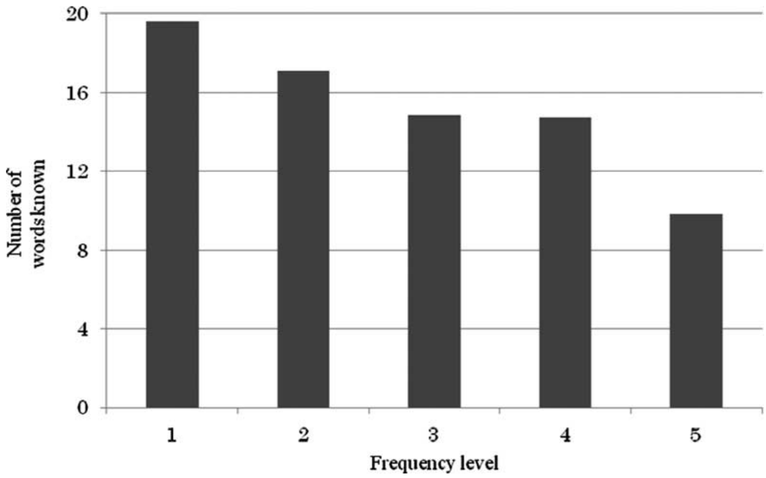

Figure 3. The number of words known at each frequency level.

Table 2. Results for Each Frequency Band ($k = 20$ per band, $n = 49$)

Band	Mean	SD	Range	Low	High
1	19.61	0.640	2	18	20
2	17.10	1.699	9	11	20
3	14.82	2.098	11	8	19
4	14.76	2.488	12	7	19
5	9.82	3.046	13	4	17

Turning to the individual results, we find similarities with Milton, with a large majority of the participants, 37 (76%), following the normal pattern. The level 2 deficit pattern is seen in the results of five participants, 10% of the total, while a structural deficit is seen in only one case. Compared with Milton, there are a relatively large number of participants displaying other patterns: 6 (12%). However, in all cases this is because equal numbers of words from different levels were indicated as known. In two of these cases, in fact, the participants indicated all 20 level 1 words and all 20 level 2 words to be known. Counting these as examples of the normal pattern, it accounts for 80% of the participants.

4 Discussion

In this group of Japanese learners the frequency model of vocabulary learning would appear to operate quite strongly. This is reassuring, considering the importance of the model in the field, but also somewhat surprising. As Milton (2007) notes, "frequency information drawn from a wide variety of native speaker sources may not be relevant to foreign language learners who are not exposed to this sort of language but have only textbooks to draw on" (p. 49), while Nation (2004) comments that "the British National Corpus (BNC) is predominantly a corpus of British, adult, formal, informative language, and most English learners in primary and secondary school systems are not British, are children, and need both formal and informal language for both social and informative purposes" (p. 4). Nation attempted to overcome this problem by basing his BNC wordlists on the spoken component of the BNC alone, but the description of the language above remains largely accurate even of this spoken component.

Why then does the frequency model appear to operate so strongly? One factor may be that, at least at the relatively high frequency levels under consideration here, most of the frequent words are the same whatever variety of English one examines. That is, whether one looks at British English or otherwise, formal or informal language, adult or child language, informative or social language, the most frequent words may be, to a considerable extent, the same. Another possibility is that the participants' exposure to English, which has largely taken place through their formal school education, has predominantly focused on adult, formal, informative language, even if probably not focused on British English.

Nevertheless, it is likely that the BNC is not the best representation of the participants' exposure to English. Besides the foreignness of the BNC to the participants as a group, the BNC, or any general corpus, cannot be taken as a representation of any individual's experience of the language. In addition, Nation's wordlists are in any case not purely based on frequency; range and dispersion also being taken into account when drawing up the lists, further muddying the waters. In trying to examine the validity of the frequency model, I am using the BNC wordlists as an approximation of the frequency of words in the participants' experience of English. It is then possible that the frequency model in fact operates more strongly than is shown in the above figures. Indeed, as we can never measure the frequency of each individual's exposure to the words of a language, we can never judge the full extent of the frequency model. Nevertheless, it may be possible to find better approximations of these participants' experience of English. As we are looking at Japanese learners, an obvious potentially better approximation is the JACET8000 list (JACET, 2003). This list, while also in part based on the BNC, tries to give a sense of how English is experienced in Japan, and so takes account of the language of school textbooks, popular tests of English, children's books and TV and film scripts.

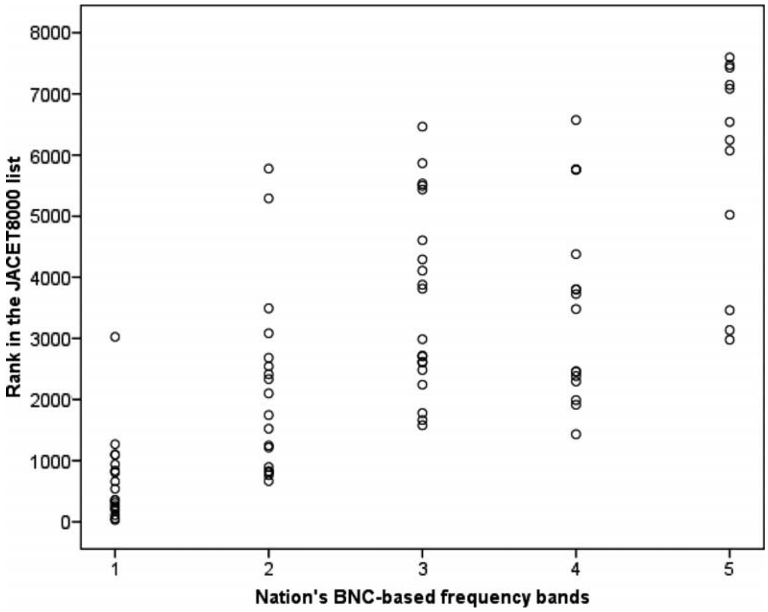

Figure 4. The JACET8000 rank of the items in each frequency band of the Yes/No test.

Looking at the words in the Yes/No test and where they appear in the JACET8000 list is revealing, as Figure 4 shows. One fundamental difference between the two is immediately clear: Nation's lists are of word families while the JACET8000 list seems to be of lemmas, hence the much higher numbers in the latter. There still appears to be, nevertheless, an overall sequence of steps across the bands, with band 4, once again, the only band seeming to not clearly follow the pattern. This is confirmed by the median JACET8000 rank of the items in each band, shown in Table 3, and also by a correlation, Kendall's $\tau = 0.60$, $n = 85$, $p < 0.001$, between the BNC-based bands and the JACET8000 rank. Still, the figure makes it clear that while the band 1 words are quite tightly bunched together in terms of their JACET8000 ranking, the words in the other bands are much more spread out and there is a great deal of overlap between the bands. This raises the question of whether a Yes/No test based on the JACET8000 list would show stronger evidence of the frequency model. Indeed, the Yes/No results correlate more strongly with the JACET8000 ranks (Kendall's $\tau = -0.594$, $n = 85$, $p < 0.001$) than with the BNC-based frequency bands ($\tau = -0.510$, $n = 100$, $p < 0.001$). Unfortunately, it was not possible to construct and administer a new test as part of this project. What can be done, however, is to look at the percentage of the items known from different JACET8000 frequency levels. Of the 100 test items, 23 are in JACET level 1 (counting the three items on the JACET Plus250 list as level 1 items), 13 in level 2, 17 in level 3 and 11 in level 4. There are only four words from JACET level 5, so this level and those above it are not considered in this analysis. Figure 5 (see also Table 4) presents the percentage of these words indicated as known. As before, there are the familiar steps down across the frequency levels, though levels 1 and 2 are close to full knowledge. Turning once more to the individual results, the ceiling effects mean that a number of participants indicated full knowledge of all 36 level 1 and level 2 words. Counting these participants as also displaying the normal frequency profile, it accounts for 42 (86%) of the 49 participants. The level 2 deficit profile is seen in only a single participant, while the structural deficit appears in four participants. Among these five participants, in four cases one more word checked would have made their profile normal. The remaining two participants indicated full knowledge of all 53 words in the first three JACET levels.

Not all of the 100 items are included in this figure: three items appear in the separate JACET Plus250 list (a list of basic items such as numbers, days of the week and months, frequent country names, frequent abbreviations and frequent irregular past tense and past participle forms); 12 do not appear in the JACET8000 list at all. These 15 items do not have a JACET8000 rank and so do not appear in the figure.

Table 3. The Median Rank in the JACET8000 List of the 20 Items in Each Frequency Band of the Yes/No Test.

Frequency band	Median rank
1	355.5
2	1634.5
3	3400.5
4	3802
5	7451.5

Note. Items appearing in the JACET Plus250 list were treated as being ranked 1 and items not in the JACET8000 list as ranked 8001.

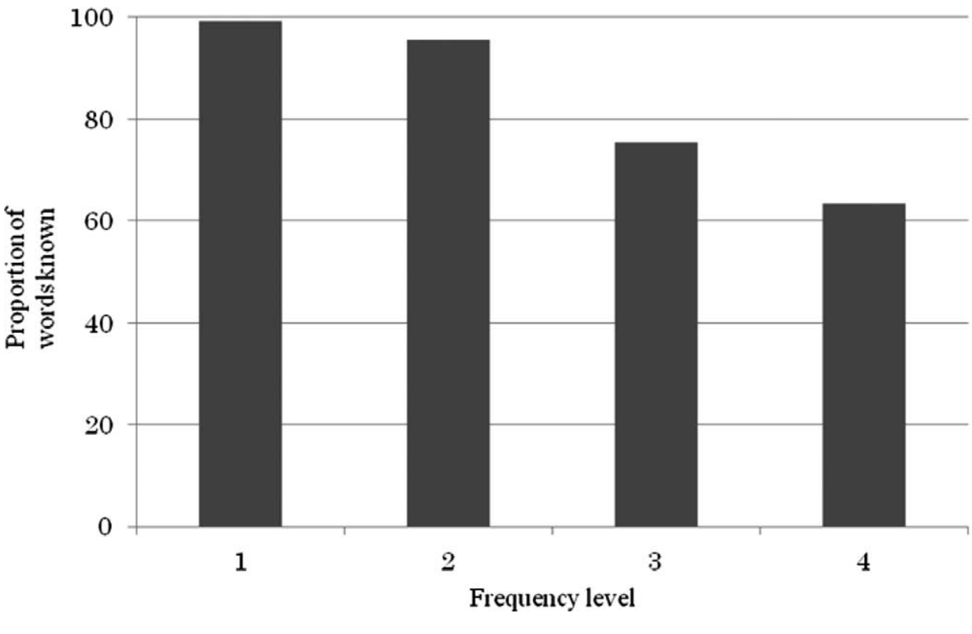

Figure 5. The percentage of words known from the first four JACET8000 levels.

Table 4. Results for the First Four JACET Levels ($n = 49$)

JACET level	Mean	SD	Range	Low	High	Mean percentage	SD (%)
1 ($k = 23$)	22.82	0.441	2	21	23	99.20	1.919
2 ($k = 13$)	12.41	0.934	4	9	13	95.45	7.181
3 ($k = 17$)	12.82	2.224	14	3	17	75.39	13.080
4 ($k = 11$)	6.98	2.106	10	0	10	63.45	19.149

5 Conclusion

The above analyses suggest first and foremost that the frequency model of vocabulary learning is indeed very strong among this group of participants. The majority of those whose frequency profiles do not follow the normal pattern only narrowly deviate from that pattern. Furthermore, it may be the case that those learners who do not appear to follow the frequency model have somehow had differing experiences of English. That is to say, it is possible that their knowledge does follow the frequency model, but the frequencies of items in their language experience differ from that of others. As was noted earlier, we can only judge the frequency model by working with approximations, and, again as the analyses suggest, the quality of those approximations has a bearing on the results. The *X_Lex* test Milton used in his research draws on both general frequency information and word lists designed for EFL (Milton, 2007, p. 50), but there must be some possibility that for some of his Greek participants who did not follow the normal pattern, those lists were simply not a good representation of their experience of English.

Regarding language teaching and learning, this study confirms two points. First, we should try to use frequency data that makes sense for the learners in question. For teachers of Japanese learners, the JACET8000 list appears to offer advantages, as it was designed to do so. Second, when dealing with the highest frequency levels, we can be fairly confident in assuming that frequency correlates with the likelihood of a word being known. Certainly, there are other factors that affect vocabulary learning, and some items of low frequency are quickly picked up by learners while some high frequency items often remain unknown. For example, looking at the items in this study, in terms of the BNC-based frequency bands, items such as *client* (Band 1), *chuck* (Band 2) and *legislate* (Band 4) were marked as known by relatively few participants compared with other items in the same band. Similarly, among the items in the first four JACET8000 levels relatively few participants indicated that they knew items such as *inch* (level 2), *ministry* (level 3) and *ass* (level 4). Yet despite there being other factors affecting the learning of words, the overall results demonstrate that the effects of frequency are powerful. For us as teachers, frequency can provide extremely helpful guidance whether we are engaged in small-scale tasks such as checking the vocabulary load of a reading text or selecting words for explicit teaching or large-scale tasks such as planning a course. As Milton (2009) concludes: "the importance of frequency in vocabulary learning is as near to a fact as it is possible to get in L2 acquisition" (p. 242).

References

Aizawa, K. (2006). Rethinking frequency markers for English-Japanese dictionaries. In M. Murata, K. Minamide, Y. Tono & S. Ishikawa (Eds.), *English lexicography in Japan* (pp. 108–119). Tokyo: Taishukan Publishing Company.

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present*. Retrieved from http://www.americancorpus.org

Ellis, N.C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(02), 143–188.

JACET Basic Word Revision Committee (2003). *JACET list of 8000 basic words.* Tokyo: Japan Association of College English Teachers.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, *21*(2), 202–226. doi:10.1191/0265532204lt277oa

Meara, P., & Milton, J. (2003). *X_Lex, The Swansea Levels Test*. Newbury: Express.

Milton, J. (2007). Lexical profiles, learning styles and the construct validity of lexical size tests. In H. Daller, J. Milton & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 47–58). Cambridge: Cambridge University Press.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

Nation, I.S.P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 3–13). Amsterdam: John Benjamins Publishing Company.

Nation, I.S.P. (2006a). *BNC-based word lists.* Wellington: Victoria University of Wellington.

Nation, I.S.P. (2006b). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review/La revue canadienne des langues vivantes*, *63*(1), 59–81.

Palmer, H.E. (1917). *The scientific study and teaching of languages.* London: Harrap.

Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, *19*(2), 12–25. doi:10.1177/003368828801900202

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Basingstoke: Palgrave Macmillan.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55–88. doi:10.1177/026553220101800103