

Assimilation of English Vocabulary into the Japanese Language

Keith Barrs

Abstract

The Japanese language contains numerous words which have been borrowed from English. As many of these words display a high degree of formal and semantic similarity with their English language equivalents, there is the potential for the benefits of these similarities to be exploited in the English language learning classroom. However, still little is known of how many and to what extent English loanwords are integrated into Japanese. This article discusses a small-scale study into both the number of cognates between Japanese and English and their frequencies in the respective languages.

Introduction

As part of the wider concept of language awareness (LA), the raising of students' crosslinguistic awareness (CLA) is seen as both a teachable and useful instructional activity in the language classroom (White & Horst, 2012). Defined as the "ability to reflect upon similarities and differences across languages" (White & Horst, 2012, p. 182), students' CLA can be nurtured through the use of explicit language instruction that draws attention to how a student's first language (L1) can influence the learning of a subsequent language (L2). A specific example of CLA raising in action is the explicit vocabulary instruction activity of focusing on the number and nature of cognates that exist between languages. Previous research has shown that such a pedagogical focus on cognates has been beneficial in

various English language learning contexts, such as for Francophone (Granger, 1993) , Polish (Otwinowska-Kasztelanic, 2009) and Spanish learners of English (Nagy, Garcia, Durgunoglu, & Hancin-Bhatt, 1990).

The Japanese language contains tens of thousands of English words and phrases (Daulton, 2008; Shibatani, 1990; Stanlaw, 2004). This leads Daulton (1999, 2004, 2008) and Ringbom (2007a, 2007b) to suggest that Japanese learners of English may benefit from explicit awareness of the cognate vocabulary between their L1 and L2. However, to date, still very little is known about the nature of these cognates. There is a lack of data concerning not only the number of Japanese-English cognates, but also the extent of integration of the English words into the Japanese language, the form they exist in, and the meanings they contain.

This article presents findings from a small-scale investigation into the quantity and extent of integration of English loanwords in Japanese. In this paper, the quantity of loanwords is investigated by way of a dictionary study and the extent of integration is judged by reference to a corpus analysis of the Japanese language. The overall aim of the study is to further understanding of the assimilation processes of English vocabulary into the Japanese language, and how this has implications for English language study in Japan.

Background

The Japanese language, having been shaped by millennia of contact with nations all across the globe, is an amalgamation of native and foreign linguistic elements. The majority of the Japanese lexicon has been imported from China and nativised

as 漢語 /*kango*/'words of Chinese origin'. One analysis of 76,536 words of modern Japanese puts the ratio of *kango* to 'native words' (和語 /*wago*) at 49.4% (37, 834) to 33.2% (25,365) (Okimori, Kimura, Tanaka, Chin, & Maeda, 2011, p. 40). Japanese also contains an extensive collection of loanwords that have entered the language from about the sixteenth century onward, called 外来語 /*gairaigo*/'words coming in from outside'. This lexical strata includes items from a wide array of the world's languages; such as French (e.g. ピーマン /*pīman*/'green pepper'), Portuguese (e.g. ブランコト /*buranko*/'swing'), German (e.g. アルバイト /*arubaito*/'part-time work') and Dutch (コーヒー /*kōhī*/'coffee'). The lexical analysis referenced above puts the percentage of the *gairaigo* strata at 9% (6,886 words). If compound words made up of *gairaigo* mixed with *kango* and/or *wago* are included then the figure rises to 10.4%. It is further estimated that over 80% of *gairaigo* are words and phrases imported from English, called 英語外来語 /*eigo gairaigo*/'English loanwords' (Daulton, 2008; Shibatani, 1990; Stanlaw, 2004). Examples include サラダ /*sarada*/'salad' and アクション /*akushon*/'action'. This means that around 8% of the modern Japanese vocabulary is made up of English words and phrases.

Most English loanwords can readily be associated with English language equivalents. Examples include チーズ /*chīzu*/cheese, プリンター /*purintā*/printer, and スプーン /*supūn*/spoon. Most of these loanwords have been taken/borrowed/loaned more or less directly from the English language and absorbed into Japanese to become part of the Japanese lexicon. The extent of the assimilation and adaption into Japanese varies depending on the individual word. These loanwords are frequently seen on posters, T-shirts, product packaging and

TV adverts; found in government documents and academic speeches; and make up a fairly substantial part of people's everyday Japanese vocabulary. Whilst many have only a fleeting existence, quickly forgotten and replaced with new imports, or are highly restricted to specialised fields, a certain number are regularly seen and heard throughout Japanese society (Daulton, 2008; Irwin, 2011).

It is also important to recognise a specialised sub-category of English loanwords called 和製英語/*waseieigo*/'made-in-Japan' English. This is a kind of 'pseudo-loanword' category. These words and phrases are clearly modelled on English, but have then been re-modelled into a form predominantly only found in Japan. In this way, they contribute to the identification of a Japanese variety of English, often called Japanese English, Japlish or Janglish. The category of 'English made-in-Japan' includes ベビーカー/*bebīkā* /baby car (pram/stroller/buggy) and ナイター/*naitā* /nighter (night baseball game). These are words which are not readily apparent in other varieties of English (Irwin, 2011). For the purposes of this article, the term 'English loanwords' (*eigo gairaigo*) excludes this specialised sub-category of loanwords.

With regard to English language education in Japan, English loanwords may be able to assist learners in their study of English vocabulary. The extent of assistance offered by each word depends on its similarity to its English language equivalent. Daulton asserts that in relation to the high frequency words of English, there are extensive enough similarities with loanwords in Japanese for this potential lexical resource to be exploited (Daulton, 2008). Nation states that native language (L1) knowledge helps to reduce the learning burden associated with the target

language (L2) vocabulary (Nation, 1990, 2008). This suggests that English loanwords in the native Japanese lexicon can have a facilitating cross-linguistic effect on the Japanese students' learning of English. Indeed, previous research has shown that L1 knowledge of English loanwords can assist the learner in their comprehension, spelling, listening and writing ability (see Daulton, 2008 for an overview).

However, with the sub-category of 'made-in-Japan' English loanwords, problems can arise for the learners and teachers. The words resemble English, especially when written out in the katakana script which highlights them as foreign, but it is difficult to associate the waseieigo with direct English language formal and semantic equivalents. For example, whilst it is not unknown for a few waseieigo creations to make their way back into English, such as サラリーマン / *sararīman*/salary man andウォークマン / *wōkuman*/Walkman, the vast majority have no such direct formal and semantic equivalency with English language vocabulary. The waseieigo items resemble English vocabulary, but are not equivalent, and this issue of cross-linguistic difference can cause confusion for the learner. This is different from cross-linguistic 'zero relations' where the learner is unable to make any associations between a linguistic element in their L1 and L2; for example an English student of Japanese trying to learn the 漢字 / *kanji*/'Chinese characters'. With cross-linguistic difference the fact that the items are similar yet different (a 'contrast relation') can cause linguistic confusion. For example, a Google.com search for the baseball-related 'nighter' returns only the term 'all-nighter', in reference to staying up all night. Similarly a Google image search for ベビーカー / *bebīkā* /'baby car' in katakana script brings

up pictures of prams and buggies, but a search for 'baby car' in the Latin alphabet brings up pictures of plastic toy cars. Because of the proliferation of waseieigo throughout Japanese society (Barrs, 2011, 2013), in places such as shop signs, T-shirts, product packaging, TV subtitles, text messages and daily conversation, the classroom-associated problems for learners and teachers are real and prominent. Unfortunately, it is this omnipresence of waseieigo in Japan that leads many to the conclusion that English loanwords in general are problematic for learners (Martin, 2004; Nuttall, 2000; Shepherd, 1995; Simon-Maeda, 1995).

The aim of this research is to investigate the quantity of loanwords which exist in the Japanese language. The study aims to shed light on what English vocabulary is already present in the L1 lexicon of Japanese students of English, with future directions of the research examining the important and under-researched issue of the formal and semantic adaptations that the loanwords undergo when brought into Japanese.

Methods

Because the focus of this cognate research is pedagogical, it was decided that an English vocabulary list based on frequency would be the starting point. This also allows the exclusion of waseieigo which, because it is a Japanese creation, would not appear on the list of items used to search for the loanwords. Nation regards high-frequency vocabulary (around 2000-5000 word families, depending on the educational focus) as the principal kind which should receive systematic classroom attention (Nation, 1990). Especially for beginner learners, it makes methodological sense to focus on the most frequent words of the target language.

Several popular frequency-based English wordlists exist, such as the General Service List and the British National Corpus 3000. However, the Japanese Association of College English Teachers (JACET) list of 8000 words of English (Aizawa, Ishikawa, & Murata, 2005) was chosen because of its appropriateness to the Japanese context (it uses the BNC lists combined with a Japan-specific corpora).

To investigate the number of direct English loanwords in Japanese, the top 200 words on the JACET 8000 list were analysed for a Japanese-language loanword equivalent by using the online version of the 大辞林辞書 /*daijirinjisho*/ 'great forest of words dictionary'. This dictionary was chosen because it has a constantly updated online version and also allows the user to search the dictionary in English. This latter function radically reduces the time necessary to search for each item. Future directions of the project will involve a more substantial number of items from the list.

The 113 loanwords were then fed into the Vocabulary Database for Reading Japanese (VDRJ), developed by Tatsuhiko Matsushita, which uses a 33 million word corpus (based on the larger Balanced Corpus of Contemporary Written Japanese developed by the National Institute for Japanese Language and Linguistics) to create a frequency list of 60,894 words of Japanese. 60,894 was considered the cut-off point below which the loanwords had an extremely low frequency.

Results

An analysis of the most frequent 200 English words reveals that there are 113 cognate correspondences with English loanwords in Japanese (Table 1 shows a sample of the correspondences, from list entry 100 to 120). A cognate ratio calculation of the 113 loanwords divided by 200 tokens results in the figure of 0.565, or 57%. Considering that the Japanese and English languages are not members of the same language family, the fact that 57% of the most frequent 200 words of English

Table 1: *JACET 8000 correspondences with the Daijirin dictionary.*

	JACET 8000 List Entry	Daijirin Dictionary Entry
100.	most	モースト
101.	much	モースト
102.	down	ダウン
103.	long	ロング
104.	why	モースト
105.	here	ヒア
106.	right	ライト
107.	even	イーブン
108.	never	×
109.	before	×
110.	too	×
111.	new	ニュー
112.	where	×
113.	tell	テル
114.	old	オールド
115.	ask	×
116.	its	×
117.	give	×
118.	call	コール
119.	any	×
120.	should	×

are cognate with English loanwords in Japanese sheds light on the potentially vast extent of lexical borrowing from English.

It can also be revealing to investigate the frequency of these loanwords within the Japanese language. Are these loanwords likely to be part of the general Japanese lexicon or do they have such a low frequency that they are unlikely to be generally known, as is the case with the vast majority of loanwords in Japan (Irwin, 2011). The results of the analysis (see Table 2) revealed that 79 of the loanwords appear in the list of 60,894 most frequent words of Japanese, with almost half (44.2%) of the loanwords appearing in the top 10,000 words of Japanese. Considering that a Japanese child of age 9 to 10 is estimated to have a vocabulary of about 10,000 words (Okimori et al., 2011, p. 34), it can be assumed that a quarter of the most frequent words of English exist in loanword form in the native lexicon of Japanese students of English. The most frequent loanword is パート /*pāto*/'part' at rank 1136 with the least frequent being プット /*putto*/'put' at position 58,249. This leaves 34 'off-list' loanwords of extremely low frequency, such as モースト /*mōsuto*/'most' and セイム /*seimu*/'same'. Whilst the corpus is primarily centred on written Japanese (although text from internet forums was included to 'balance' the corpus), it does give a good indication of the kind of Japanese language in use, along with the loanwords included in this language.

Table 2: *Frequency of the Loanwords in the VDRJ database*

Frequency Level	Number of Loanwords	Cumulative Number	Cumulative Frequency
0-1000	0	0	0%
-2000	7	7	6.2%
-3000	8	15	13.3%
-4000	8	23	20.4%
-5000	8	31	27.4%
-6000	5	36	31.6%
-7000	6	42	37.2%
-8000	2	44	38.9%
-9000	2	46	40.7%
-10000	4	50	44.2%
10001-60,894	29	79	69.9%
Off list	34	113	100%

Discussion & Conclusion

These findings reveal that even in this small collection of 200 words, there is a large quantity of cognates between the two languages. Many of these loanwords have high-frequency within the Japanese language, suggesting that classroom-based activities which focus on exploiting the benefits of cognates can have value in the Japanese context.

Once the entire JACET 8000 list has been analysed for cognate correspondences, investigations can be made of the formal and semantic similarity between the word pairs. This will help to further understanding of the cognates' similarities and differences and can be used to help raise students' cross-linguistic awareness of the pedagogical value of their L1 vocabulary in the learning of the L2.

References

- Aizawa, K., Ishikawa, S., & Murata, M. (Eds.). (2005). JACET8000英単語「大学英語教育学会基本語リスト」に基づく[JACET 8000 English words: *Based on the JACET basic word list*]. Tokyo, Japan: Kirihara Publishing.
- Barrs, K. (2011). Unlocking the encoded English vocabulary in the Japanese language. *English Today*, 27(03), 15–20. Retrieved from http://www.journals.cambridge.org/abstract_S0266078411000320
- Barrs, K. (2013). L2 vocabulary in the L1 environment. *Working Papers in Language Education and Research*, 1(1), 3–10.
- Daulton, F. E. (1999). English loanwords in Japanese – The built-in lexicon. *The Internet TESL Journal*, 5(1), 1–5.
- Daulton, F. E. (2004). *Gairaigo: The Built-in lexicon?* Victoria University, Wellington.
- Daulton, F. E. (2008). *Japan's Built-in Lexicon of English-based Loanwords*. Clevedon: Multilingual Matters Ltd.
- Granger, S. (1993). Cognates: An air or a barrier to successful L2 development. *ITL Review of Applied Linguistics*, 99-100, 43–56.
- Irwin, M. (2011). *Loanwords in Japanese*. Philadelphia: John Benjamins Publishing Company. Retrieved from <http://books.google.com/books?id=iGFGIc06LE4C&pgis=1>
- Martin, A. (2004). The katakana effect and teaching English in Japan. *English Today*, 20(1), 50–55. doi:10.1017/S0266078404001087
- Nagy, W., Garcia, G. E., Durgunoglu, A., & Hancin-Bhatt, B. (1990). Spanish-English bilingual students' use of cognates in English reading. *Journal of Literacy Research*, 25(3), 241–259. doi:10.1080/10862969009547816
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston, M.A.: Heinle & Heinle Publishers.

- Nation, I. S. P. (2008). *Teaching Vocabulary. Strategies and Techniques*. Boston, M.A.: Heinle.
- Nuttall, M. (2000). English in Japan. *Japan Railway and Transport Review*, 25, 56–57.
- Okimori, T., Kimura, Y., Tanaka, M., Chin, R., & Maeda, N. (2011). *日本の語彙 [Japan's vocabulary]*. Tokyo, Japan: Sanseido.
- Otwinowska-Kasztelanic, A. (2009). Raising awareness of cognate vocabulary as a strategy in teaching English to Polish adults. *Innovation in Language Learning and Teaching*, 3(2), 131–147. doi:10.1080/17501220802283186
- Ringbom, H. (2007a). *Cross-linguistic similarity in foreign language learning*. Clevedon: Multilingual Matters Ltd.
- Ringbom, H. (2007b). The importance of cross-linguistic similarities. *The Language Teacher*, 31(9), 3–5.
- Shepherd, J. W. (1995). Loanwords: A pitfall for all students. *The Internet TESL Journal*, 2(2), 1–6.
- Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.
- Simon-Maeda, A. (1995). Language awareness: Use/misuse of loan-words in the English language in Japan. *The Internet TESL Journal*, 1(2), 1–5.
- Stanlaw, J. (2004). *Japanese English: Language and culture contact*. Hong Kong: Hong Kong University Press.
- White, J. L., & Horst, M. (2012). Cognate awareness-raising in late childhood: teachable and useful. *Language Awareness*, 21(1-2), 181–196. doi:10.1080/09658416.2011.639885