

## **Ubiquitous variability in the phonological form of loanwords: Tracing early borrowings into Japanese over five centuries of contact**

Within the field of Loanword Phonology, the issue of diachronic change and social variation has only recently begun to be addressed. With respect to diachronic change, Kang (2010) argues that the driving forces in loanword adaptation can evolve over time, with early adaptation being primarily perceptually-driven and later adaptation involving more categorical mappings at the abstract/systemic level. With respect to social variation, Crawford (2009) argues convincingly that the loanword borrowing should be broken into two independent stages: [1] the initial *adaptation* of the word by the first generation speakers (who may be in direct contact with the source-language speakers) and [2] the subsequent process of *transmission* of the adapted word-forms throughout the borrowing community (even after contact has ceased). Crawford (2009) also models how multiple variant forms of the same loanword can coexist and compete with each other until a particular form ultimately emerges as 'dominant' within the community. This account makes the strong prediction that, in the early stages of language contact, there should be pervasive variability in the phonological form of loanwords. However, Paradis & LaCharité (2008) claim that borrowings during early contact are just as categorical as in the modern language, with only a slight difference in variability between the two. However, it is possible that this finding is merely an artefact of the fact that only a single data-source was used to represent the early-contact period. In order to get a richer (and more realistic) picture of the full scope of variation at each time point, multiple data sources must be triangulated.

Using loanwords into Japanese as a case study, the present study reports on a database compiled with precisely this goal. First, a representative sample of the historically earliest Western loanwords in Japanese was identified. To do so, all Western loanwords in a 130-year old Japanese dictionary (Yamada 1892-1893) were identified, giving a comprehensive picture of the loanword forms that enjoyed common currency in the 1890s. Then, to get a sense of historical depth, each loanword form was assigned an approximate date of first being borrowed into Japanese, based on that word's earliest written attestation. The source for this data was *Nihon Kokugo Daijisho* (Nihon Daijiten Kankōkai 2000-2001), which, much like the *Oxford English Dictionary*, provides dated citations from primary-source historical documents (as early as 1569). Next, the status of these loanword forms within the modern Japanese language was ascertained by cross-referencing these loanwords in modern synchronic dictionaries. To avoid the possibility of random sampling bias associated with one particular dictionary, two different dictionaries were used in the process – *Daijisen* (Matsumura 1998) and *Daijirin* (Matsumura 2006). Finally, in every source that was consulted, the patterns of dictionary-article redirects were analyzed. Under the assumption that the 'dominant' loanword form at a given time point has the full-text definition (with all other variant forms redirecting to it), this can be taken as a rough indication of which loanword forms have shifted in popularity over the past century. The resulting database contains a total of 748 variant forms of 255 loanwords, with rich information on the history behind the 'pool of variation' for each word-family.

Results indicate that the extent of variation in loanword phonology has been grossly underestimated. As many as 70% of the loanwords in the database have more than one variant form; indeed, many words have well over a dozen. Surprisingly, even though the relevant loanwords arose out of five wildly different contact situations, the overall proportion of the extents of variation is remarkably stable across source languages. These findings have a number of important implications for the field of Loanword Phonology. First, the data typically used in studies of loanword phonology is far too sparse. Rather than taking "form X in language A is borrowed as form Y in language B" as the raw data, it is imperative that the phonologist draws on a richer dataset that encompasses the full envelope of variation. Moreover, clearly, no single principle (e.g. "all early borrowing is the product of perception" or "all later borrowing is systemic/categorical in nature") can possibly generate the full range of variation attested in the present study. As Crawford (2009:17) aptly stated, approaches like these suffer from the same fundamental flaw: they seek to explain "the behavior of a single idealized speaker at a particular moment in time". Instead, it is concluded that the field ought to seek correlations between aspects of the contact situation and the ultimate 'dominance' of certain forms within the pool of variation.

## References

- Crawford, C. (2009). *Adaptation and transmission in Japanese loanword phonology*. Unpublished Ph.D. dissertation, Cornell University.
- Kang, Y. (2010) The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology* 27, 225-253.
- Matsumura, A. (1998). *Daijisen [Great Fountainhead of Words]*, (2nd ed.). Tokyo: Shōgakukan.
- Matsumura, A. (2006). *Daijirin [Great Forest of Words]*, (3rd ed.). Tokyo: Sanseidō.
- Nihon Daijiten Kankōkai (2000-2001) *Nihon Kokugo Daijiten [Great Dictionary of the Japanese National Language]*, (2nd ed.). Tokyo: Shōgakukan.
- Paradis, C. & LaCharité, D. (2008). Apparent phonetic approximation: English loanwords in Old Quebec French. *Journal of Linguistics* 44, 87-128.
- Yamada, B. (1892-1893). *Nippon Daijisho [Great Dictionary of Japan]*. Tokyo: Nippon Daijisho Hakkōjo.