

Evolution in the derivation of compound abbreviated loanwords in Japanese

Rémi Lamarque

Aix-Marseille Université, Aix-en-Provence, France

In Japanese, compound loanwords are often abbreviated. The kind of abbreviation we call Compound Abbreviated Loanwords (hereafter CALs¹) is very frequent and productive in Japanese. CALs are derived from a compound base typically consisting of two constituents of a foreign origin. The final derived CAL always includes elements from both constituents, and its formation is subject to prosodic constraints. Most of the time, CALs are abbreviated to a length of four morae. The Japanese word *pokemon*, for example, comes from the two English loanwords *pocketto* 'pocket' and *monsutaa* 'monster' (Shinohara, 1997). The most common way of deriving a CAL out of two loanwords is by maintaining only the first two morae of each constituent, as in the case of *po-ke-mo-n* (*po-ke-t-to* and *mo-n-su-ta-a*) (Labrune, 2007). However, not all CALs follow this general pattern, and there is a wide range of exceptions. The most frequent kind of exception is by far the derivation of a three morae long CAL consisting of the first two morae of the first constituent and only the first mora of the second constituent. For example, the name of the famous actor Brad Pitt is abbreviated *bu-ra-pi* in Japanese, where the first two morae come from the Japanese adaptation of the first name, *bu-ra-d-do*, and the third mora comes from his last name, *pi-t-to*.

When asked about the reason for such a tendency to abbreviate in three morae rather than the expected four morae pattern, some native speakers answered that it was somehow related to a newer way of derivating CALs. For those speakers, three morae CALs are perceived to be more frequent in newer CALs than in older ones. This could imply that the constraints governing the derivation of CALs may have changed slightly over time. In the present study, we seek to test this hypothesis inspired by native speaker intuition by comparing corpora of CALs drawn from different time periods. This not only allows us to test for a change in the proportion of three morae CALs, but also provides additional data for exploring possible *phonological* reasons behind the creation of three morae CALs. For older CALs, we used a large corpus created by Laurence Labrune (2006, 2007, 2008). This was created around 2001 and includes 711 CALs² from a variety of sources and prior time periods. As such, it serves as a reference for older CAL derivation processes. For newer CALs, a new corpus was created involving tokens drawn from the video gaming community. This media is still mostly used by the youngest generations and includes a large number of English loanwords, making it an ideal source of newer CALs. Overall, 300 CALs were collected from websites concerning video games.

As seen in Fig. 1, a comparison between the two corpora shows that the proportion of three morae CALs was indeed greater among newer CALs, confirming our hypothesis. However, three morae CALs do not appear at random and seem to be conditioned by certain phonological contexts. In a large proportion of three morae CALs across both corpora, for example, it appears that the fourth mora, which is deleted, corresponds either to the lengthening of a vowel or to the first part of a geminate consonant (as in *bu-ra-pi*, where the following mora should be the first part of the geminated consonant *t* of *pi-t-to*). The increase in three morae CALs is therefore closely linked to a change in the treatment of those two special morae when appearing at the end of the CAL.

These changes can be modeled using the Optimality Theoretic framework, which expresses phonological tendencies in terms of ranked, violable constraints (Prince & Smolensky (1993/2004)). The most important change concerns the relative importance of a constraint that enforces a bimoraic second constituent. Among younger speakers, this constraint appears to have been demoted so that it is sometimes violated in favor of two other constraints, one which forbids long vowels at the end of a constituent, and one which requires geminating slots to be followed by voiceless obstruents. In short, the new constraint hierarchy favors deletion of these morae, resulting in many more instances of second constituents that are monomoraic.

¹ The term CAL is borrowed from Labrune (2006)

² The corpus actually includes 799 CALs, but 88 of them were excluded from it because they were found in a dictionary named « Young people's words dictionary », making it clear that they were newer CALs.

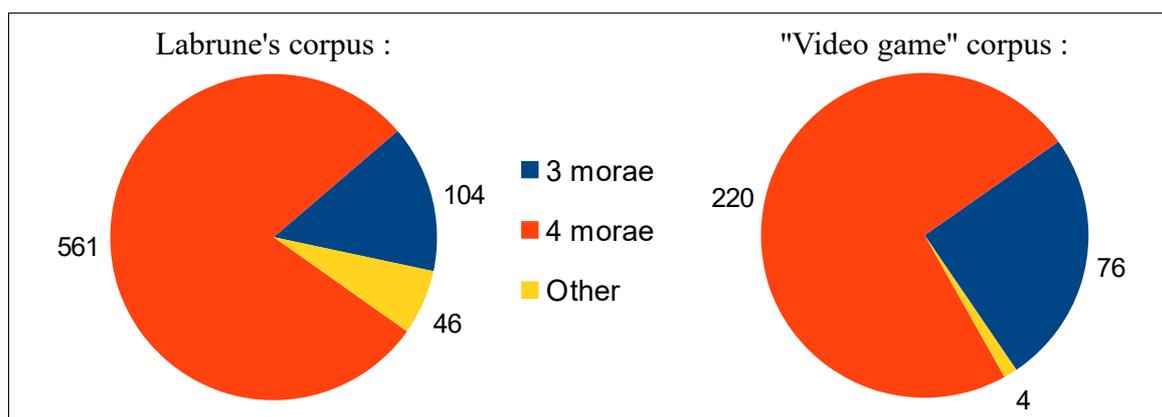


Figure 1 : Proportion of CALs by number of morae in both corpora.

Références bibliographiques :

- LABRUNE, Laurence, « Phonemic preferences in Japanese non-headed binary compounds : what waa-puro, mecha-kucha and are-kore have in common », in *Gengo Kenkyû, Journal of the Linguistic Society of Japan*, 129, 2006, pp.3-41.
- LABRUNE, Laurence, « Morphophonologie des emprunts abrégés en japonais : aspects prosodiques et mélodiques », in *Des sons et des sens : données et modèles en phonologie et en morphologie*, DELAIS-ROUSSARIE, E., & LABRUNE, L., (éds), 2007, Paris : Hermès, 28p.
- LABRUNE, Laurence, « Principes d'organisation phonémique des emprunts occidentaux composés abrégés », in *Revue d'Etudes Japonaises*, Paris : Université Paris 7, 2008, 14p.
- PRINCE, Alan, SMOLENSKY, Paul, *Optimality theory constraint interaction in generative grammar*. Malden : Blackwell, (1993/2002)2004, 289p.
- SHINOHARA Shigeko, « Analyse phonologique de l'adaptation japonaise de mots étrangers ». Thèse de doctorat sous la direction de RIALLAND Annie, Paris : Université de la Sorbonne Nouvelle - Paris III, 1997, 219 p.