Morphophonological convergence and rule emergence in recent morphological processes
Rémi Lamarque

In morphology, variability arises when speakers have some freedom to choose between different forms in the creation of novel forms (Fehringer, 2004; Haber, 1976; Säily, 2011; Thornton, 2012). Type frequencies of the different patterns in existing forms have a considerable influence on the novel forms speakers create (Albright & Hayes, 2003; Bybee, 1995, 2001; Pierrrehumbert, 2001a, 2001b). The present study addresses the emergence of rules in recent morphological processes, for which the community has not yet converged on a single convention, and the role of morphophonological convergence in compensating for the lack of lexical information needed to abstract a rule. We explore this issue through the case study of a recent morphological process in Japanese: Compound Abbreviated Loanwords (hereafter CALs) formation. We ran an online game-like experiment in which Japanese speakers created novel abbreviations both alone and paired with an (unknown to them) artificial partner programmed to regularize at a higher or lower rate compared to the participants’ behavior. We argue that the behavior of the artificial partner can influence participants’ strategy in resolving conflicts arising in the creation of novel forms, especially in situations where lexical information is scarce.

CALs are formed by preserving the two first morae of two words of foreign origin (see ex.1, Labrune, 2006). Apart from this regular pattern, three irregular patterns are used in existing forms, mainly to prevent violation of phonological constraints. Notably, a violable constraint on long vowels, and an inviolable constraint on the first part of a geminate appearing in final position of a word (see ex.2, 3 and 4). In a previous study, we analyzed a lexicon of 701 CALs, among which 541 were not concerned by these phonological constraints (hereafter “no conflict” CALs), 113 were concerned by the violable constraint (hereafter “soft constraint” CALs), and 47 by the “hard” constraint. Soft constraint CALs display the most variability, likely since any pattern is an equally valid option for them, but occurrences are too rare for speakers to be sensitive to differences in their type frequencies. In our experiment, participants played a three-phase online game during which they were asked to create novel abbreviations. For each trial, the participants were presented with a pair of English words and asked to pick the most natural abbreviation for those words among four patterns. In the first phase, they played alone. In the second phase, they were connected to another player and told that the goal was for both players to give the same answers (Von Ahn & Dabbish, 2004). They would see the other player’s answer at the end of each trial. In the third phase, the participants played alone again. The pretend other player in Phase 2 was actually a robot programmed to randomly follow one of three possible behaviors based on the answers of the participants in Phase 1: using the regular pattern more often than the participant did (“over-regularizer” condition); using it less than the participant (“under-regularizer” condition); or using all patterns with the same frequency as what the participant did (“no change” condition).

The results (fig.1) show that as the game progressed, participants used the regular pattern more often on the “no conflict” items, regardless of the behavior of the robot. For the soft constraint items, we observe a similar slight increase in the “no change” condition, but a significantly greater increase in the “over-regularizer” condition, while the “under-regularizer” condition shows a significant decrease. A generalized logistic mixed model was fit to the data response type (regular/irregular) as the binary dependent variable, phase and robot condition as the fixed factors, and item as a random factor. For “no conflict” items, model comparison revealed a significant effect of phase ($\chi^2=29.605, p<0.001$), and no significant effect of robot condition ($\chi^2=0.2791, p=0.6$), or interaction ($\chi^2=2.6252, p=0.10$). For “soft constraint” items, it revealed a significant effect of robot condition ($\chi^2=5.4611, p<0.05$), phase ($\chi^2=32.751, p<0.001$), and interaction ($\chi^2=20.645, p<0.001$). These results suggest that morphophonological convergence plays a significant role in situations where the type frequencies are insufficient for
speakers to abstract a rule, possibly tipping the scale in favor of a convention over the competing ones, before its influence diminishes as the type frequency of the corresponding pattern increases.

**Example 1:** the regular pattern (*pokemon*). 4 mora long

<table>
<thead>
<tr>
<th>English</th>
<th>Japanese</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>pocket + monster</td>
<td>po.ke.t.to + mo.n.su.ta.a</td>
<td>po.ke.mo.n</td>
</tr>
</tbody>
</table>

**Example 2:** the irregular “follow” pattern (*purakeesu*). 5 mora long

| plastic + case   | pu.ra.su.ti.k.ku + ke.e.su | pu.ra.e.e.su    |

**Example 3:** the irregular “delete” pattern (*burapi*). 3 mora long

| Brad + Pitt      | bu.ra.d.do + pi.t.to       | bu.ra.pi        |

**Example 4:** the irregular “replace” pattern (*amefuto*). 4 mora long (but not contiguous)

| American + football | a.me.ri.ka.n + fu.t.to.bo.o.ru | a.me.fu.to      |

**Figure 1:** proportion of use of the regular pattern on the different type of items in each phase and each artificial partner behavior condition

![Figure 1](image_url)

**References:**


