From variation to the emergence of linguistic regularities
Rémi Lamarque

To cite this version:
Rémi Lamarque. From variation to the emergence of linguistic regularities. Current Trends in Linguistics, 2017, Hamburg, Germany. hal-01469827

HAL Id: hal-01469827
https://hal.archives-ouvertes.fr/hal-01469827
Submitted on 16 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
**Introduction:**
- In languages, two rules may be in competition. Regularization occurs when one of them emerges as an optimal solution and completely overcomes the other one.
- Regularization is often seen as a change in individuals’ grammar.
- Evidences from cultural evolution (Derex, 2015, 2016) or complex systems (Becker & al., 2009) studies suggests that changes at a macro level doesn’t necessarily reflect changes at a micro level.
- In linguistic terms, this suggests that regularization at a community-level does not reflect a change in individuals’ grammar.
- In this study, I explore how community-level factors may explain the regularization of certain derivational processes (Compound Abbreviated Loanwords) in Japanese.

**What is a Compound Abbreviated Loanword (CAL) ?**
CALs are abbreviations:
- Based on two constituents of foreign origin
- Frequent and employed in various situations (not only in informal speech)
- Phonologically constrained derivation process: Conservation of the initial morae of each constituent to form a four morae abbreviation (see pokémon below).
- Sometimes irregular: Majority of exceptions include three morae, maintaining only the first mora of the second constituent (see potetā).
- Individuals show different patterns of irregularity, reflecting competition between rules.

**Examples :**

<table>
<thead>
<tr>
<th>English</th>
<th>Japanese</th>
<th>Katakana</th>
</tr>
</thead>
<tbody>
<tr>
<td>pocket monster</td>
<td>ポケットモン</td>
<td>ポケモン</td>
</tr>
<tr>
<td>potato chips</td>
<td>ぷケトチ</td>
<td>ポテチ</td>
</tr>
</tbody>
</table>

**Community- vs. individual-level regularization:**
- Corpus surveys of attested forms (Labrune, 2007; Lamarque, 2015) show that three morae CALs are often created when the second mora of the second constituent is either the lengthening of a vowel (/R/) or the first part of a geminate consonant (/Q/), as in potetā. Moreover, the deletion of these morae has become more systematic with time:

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>% deletion of /R/</td>
<td>45%</td>
</tr>
<tr>
<td>% deletion of /Q/</td>
<td>45%</td>
</tr>
</tbody>
</table>

- This suggests that the deletion of /R/ and /Q/ when they appear in final position of a CAL has regularized.
- However, an experimental study (Lamarque, 2016) showed that individuals still have a highly variable treatment of /R/ and /Q/ in the creation of new CALs.
- /R/ and /Q/ were deleted in less than 30% of cases (see fig.1).
- The regularity in attested forms, representing the regularization at a community-level, does not reflect the distribution of preferences in individuals’ grammar.

**New experimental approach:**
**Questions arising from previous results:**
- How can regularity emerge at community-level when individuals have different preferences?
- What community-level factors play a role in the regularization process?

**Multi-player online game (see Fig.2):**
- Players create new CALs together.
- Players assign names to imaginary creatures.
- Players are motivated to anticipate group-level outcomes.
  - E.g., they get more points if everyone in a group chooses the same name.

- Focus on how players converge on naming solutions when variation in the grammar presents multiple options (e.g., when /R/ or /Q/ should appear at the end of a CAL).

**Community-level parameters:**
- The size of groups of players (Derex & al., 2013)
- Connections within and between groups (i.e., network structure, Derex & Boyd, 2016)
- Mode of interaction (e.g., negotiation via chatbox).
- Social factors (e.g., variant prestige, Tamariz et al., 2011)
- Distribution of biases in the community (Pierrehumbert et al., 2014)
- ...

**Main objectives :**
- Explain discrepancy between individual-level and community-level preferences.
- Identify the role of different community-level factors in the process of regularization.

**Next steps:**
- Complete the creation of the game (select items, design the creatures, run tests, ...).
- Create multiple versions with different parameter manipulations.
- Collect the data.
- Analyze the data: what factors significantly influenced the players’ attitudes and scores? In what way?
- Conduct agent-based modeling simulations (Netlogo) and compare to empirical results. (Wilensky, 1999)

**References:**
- Wilensky, Uri, Netlogo. http://ccl.northwestern.edu/netlogo/, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999

**Acknowledgements:**
This study is partially supported by a grant from A*MIDEX (n° ANR-11-IDEX-0001-02) funded by the Investissements d'Avenir French Government program, managed by the French National Research Agency (ANR).