The limits of inductive learning: The case of Modern Irish mutation

Ruth Maddeaux (ruth.maddeaux@mail.utoronto.ca), Yoonjung Kang (kang@utsc.utoronto.ca)
University of Toronto

Introduction: Recent studies show that speakers acquire detailed knowledge of statistical patterns in their language and a great deal of what they know about the phonology of their language may be projected from the extant patterns available in the learning input (Ernestus & Baayen 2003, Zuraw 2010, Daland, et al. 2010, Hayes et al. 2011, etc.). Other studies point to the limits of inductive learning and claim that speakers are aware of universal preferences in phonological patterns, even if evidence for those preferences is absent from their own language (Wilson 2006, Moreton 2008, Berent, et al. 2008, Becker, et al. 2011, 2012, etc.). Such universal preferences reveal themselves in the form of learning biases leading to “underlearning” of patterns that are counter to or unsupported by purported universals. But, much remains to be explored about the limits of inductive learning and the precise nature of universal preferences that constrain the induction of phonological generalizations.

The present study contributes to this line of inquiry by examining Irish consonant mutation as a test case. In Irish, word-initial [d] and [g] in non-lenition contexts alternate with [ɣ] in lenition contexts, while [dʲ] and [gʲ] alternate with [j] (See (1)). In other words, for a lenited form beginning with [ɣ] or [j], there are two possible alternations: one with a coronal stop in non-lenited contexts and one with a dorsal stop.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[an ɣuːlán] ‘the shoulder’</td>
<td>[an yə] ‘the colour’</td>
<td>[an ɡalax] ‘the moon’</td>
<td>[an ɣʊlæv] ‘the statue’</td>
<td></td>
</tr>
</tbody>
</table>

In this paper, we report the findings from a nonce word experiment that probes Irish speakers’ preference between the two alternation options (coronal stop ~ [ɣ]/[j] vs. dorsal stop ~ [ɣ]/[j]). We examine to what extent the speakers’ preference reflects the statistical distribution of existing alternations and what role, if any, phonological universals play in shaping speakers’ preference in these alternations.

Dictionary data: To explore the distribution of alternation types in the extant language data, we extracted all nouns that begin with /d/, /ɡ/ or their palatalized counterparts (n=418) from an Irish dictionary (Seaton 2009). Note that the dictionary form shows the non-lenited form of the noun and indirectly indicates the type frequency of the two alternation types explained above. We then examined how the distribution of coronal and dorsal stops is conditioned by various phonological conditions, such as the palatalization of the stop itself and by the segmental features of the immediately following segment. Exploratory analyses and follow-up logistic regression analyses identified the distributional asymmetries summarized in Figs. 1-3. There was no significant effect of palatalization; the proportion of coronal stops was comparable between palatalized (41.2%) and non-palatalized (43.2%) word-initial stops (Fig. 1). The proportion of coronal stops (out of coronal and dorsal stops) is significantly higher when the following segment is a non-low vowel (57.9%) than a low vowel (36.2%) or a consonant (26.7%) with no statistically significant difference found between the low vowel and the consonant contexts (Fig.2). Among the following consonant conditions, coronal stops are categorically absent before /n/ (0%), very rare before /l/ (9.4%) but more frequent before /ɾ/ (39.2%) (Fig. 3).

Nonce word experiment – Methodology: In the nonce word experiment, participants heard a sentence with a nonce noun in a lenition context along with an orthographic representation of the same sentence on a computer screen, with a blank standing in for the target noun. The participants were then prompted to fill in the blank by writing out the noun that they heard in Irish orthography. The contrast between voiced coronal and dorsal stops is neutralized in the lenited condition ([ɣ] and [j]) as explained above but the Irish orthography forces the writers to choose between a coronal or a dorsal stop (<dh> or <gh>) thereby revealing their preference with respect to the non-lenited counterpart of the nonce noun they heard. The target nonce words all began with a voiced coronal or dorsal stop and were balanced for the phonological conditions of our interest, namely, palatalization (i.e., nonce lenited form began with [ɣ] (non-palatalized) or [j] (palatalized)) and the following segment quality. 54 target items were included along with 66 fillers, and 22 native speakers of Irish participated in the study. 1071 responses that chose one of the expected stops were included in analyses. Mixed-effects logistic regression was used for statistical analyses.
**Results and discussion:** The nonce word experiment results generally confirm the tenet of inductive learning that speakers possess implicit knowledge of the statistical patterns of their language, but we also found evidence for a substantial role of universal phonological preference. First of all, responses by the following consonant (Fig.6) fairly accurately matched the pattern of the dictionary (Fig. 3). *[^dn]/[^dn]* clusters were never chosen as a possibility and the proportion of coronal stop before [l] (6.5 %) remains far lower than the proportion of coronal stops before [r] (37.9 %). We conjecture that this accurate learning of the lexical pattern was possible because the lexical generalization is supported by reasonable universal phonological preference, namely, an OCP constraint against a sequence of [Coronal, -continuant] consonants. The responses to the following segment type (Fig.5) also followed the same trend as the dictionary data (Fig.2) but the two differ in a crucial way: while the dictionary pattern groups the following segments into non-low vowels vs. others (low vowels and consonants), the speakers derived a slightly different generalization, choosing coronal stops more frequently before a vowel than before a consonant (19.9 %), along a major natural class division. The difference between low (43.3 %) and non-low vowels (57.5 %), which is significant in the dictionary data is no longer significant in the experiment data. Note that there is no plausible phonological universal to support a stronger preference for coronal stops before non-low than low vowels and hence the speakers seem to have “underlearned” this pattern. Finally, the speakers had higher coronal stop responses when the lenited form of the target word began with [j] (indicating that the non-lenited form of the stop should be palatalized) (45.9 %) than when the word began with [y] (42.3 %) and this difference turns out to be statistically significant (Fig. 4). Note that the dictionary data had a non-significant trend in the opposite direction (Fig. 1). We attribute this emergent effect of palatalization to the universal preference for paradigm uniformity. For non-palatalized consonants, the dorsal stop paradigm incurs less featural changes (change in continuancy but no change in place) than the coronal stop (change in both continuancy and place) ([g]~[y] >> [d]~[y]). The alternation in the palatalized paradigm, on the other hand, involves comparable changes (deletion of major constriction) for the two stops ([g]~[j] ≈ [d]~[j]). So, coronal stops fair better in terms of paradigm uniformity in palatalized over non-palatalized stops in relative terms. In conclusion, our results indicate that Irish speakers have implicit knowledge of the statistical patterns in the distribution of these nouns but the generalizations they draw are filtered through the sieve of universal phonological preferences.