Tobias Scheer, Fabien Mathy, Fanny Meunier (Université Côte d'Azur)

How to determine whether an alleged phonological process is real

Problem. For the past 50 years or so (starting with Kiparsky 1968-73), a central question for linguistic theory is whether the pieces (morphemes) that may be identified by the analyst are really the ones that are stored in long term memory and manipulated by the cognitive system of present day natives. Given for example *electric* - *electricity*, are we facing suppletion (two independent lexical items /electric/ and /electricity/), allomorphy (three lexial items /electric/, /electris/, /-ity/, the s-allomorph being chosen in the presence of /-ity/) or morpho-phonology (two lexical items /electric/, /-ity/, their concatenation provoking k→s)? Upon production, suppletion requires a lexical access but no concatenative or phonological computation. Allomorphy on the other hand engages a lexical access and concatenation, but no phonological activity. Finally, morpho-phonology mobilizes all three actions.

For decades, especially phonologists have tried to establish a set of criteria (called evaluation measure or evaluation metrics) that are able to decide for any given alternation which category it belongs to. Despite the effort, all attempts remained inconclusive (Bermúdez-Otero & McMahon 2006: 383ff). Therefore the popperian competition among theories is significantly biased: a theory that accounts for velar softening *in the phonology* cannot be compared to a theory considering that this process has nothing to do with phonology. The set of things to be explained is not the same, and wildly diverges already at the scale of a language. Before theories can compete, the question what a true phonological phenomenon is thus needs to be assessed.

Experimental setup. Based on electro-physiological evidence from a speech production task (data acquired through deep electrodes implanted in the brain of three patients for medical evaluation), Sahin *et al.* (2009) showed that there are at least three linguistically distinct processes (ERPs, Event Related Potentials) that can be separated in time and space in the brain: first lexical access occurs (at ~200*ms*), then concatenation is performed (at ~320*ms*), then phonology is operated (at ~450*ms*). They have thus established an experimental setup that is able to discriminate the three production categories that are at stake in the linguistic debate mentioned.

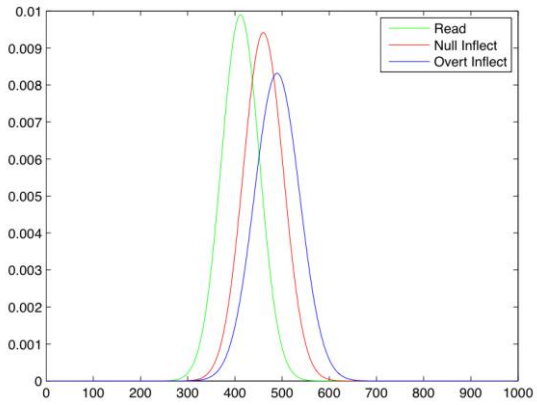
Sahin et al.'s protocol is based on three experimental conditions: *Read* merely requires the subject to repeat a stimulus word (e.g. *rock*). Only lexical access is required for this task, since the word is not inserted into any grammatical context: the word only needs to be retrieved from long term memory in order to be pronounced. In the *Null Inflect* condition, a carrier sentence such as *this is the* __ followed by the (visually presented) stimulus *rock* enforces the participant to grammatically process the stimulus word (concatenation). This operation, however, leaves no phonetic trace since the sg. remains unmarked in English. Finally, in the *Overt Inflect* condition things are as before except that the body of the stimulus *rock* is now modified: the carrier sentence *these are the*__ leads the participant to produce *rock-s*. In addition to lexical access and concatenation, a phonological operation needs to be performed: under lying /-z/ (cf. *tree-[z]*) is devoiced by the stem-final k. Significantly, Read produces a peak in electric activity only at ~200*ms* (lexical access), Null Inflect at ~200*ms* and at ~320*ms* (concatenation), while Overt Inflect shows an ERP at ~200*ms*, ~320*ms* and at ~200*ms* and ~450*ms* (phonological computation).

Experiment. We report on experimental evidence that replicates Sahin et al. in behavioural modality (reaction time, RT) and using stimuli sets from four French alternations: consonant-zero in masc. vs. fem. forms of adjectives (*grand* [gʁã] - *grande* [gʁã**d**] "big") and in conjugational paradigms (*il part* [paχ] "he leaves" vs. *vous part-ez* [paχ**t**-e] "you pl. leave"), schwa-ɛ alternation in conjugation (*elle ach[ɛ]te* "she buys" vs. *vous ach[ə]tez* "you pl. buy") and the French equivalent of velar softening (*électrique* [elɛktʁi**k**] "electric" - *électric-ité*

[elɛktʁis-ite]). The time elapsed between stimulus presentation (in spelling on a screen) and voice onset (measured by a vocal key, i.e. a microphone) was recorded for every trial of 40 French natives responding each to a 154 words in 3 experimental conditions (462 trials total). We expected that more cognitive activity takes more time. Hence, that all other things being equal, the reaction time hierarchy is *Read* < *Null Inflect* < *Overt Inflect* (as in Sahin et al.'s experiment).

Results. Results are of three types, in decreasing granularity. Across all four alternations, reaction time indeed follows the prediction *Read* (mean RT 412ms) < *Null Inflect* (mean RT 460 ms) < *Overt Inflect* (mean RT 489ms) (see figure).

At the scale of individual alternations, consonant-zero alternations in adjectives and schwa-ɛ did not show any effect across the three conditions (about the same RT for Read, Null Inflect and Overt Inflect). Consonant-zero alternations in conjugation did produce a significant effect, but only when comparing Read and Overt Inflect (the difference between Null Inflect and Overt Inflect is not significant). Finally, French velar softening induces a significant effect between both Read and Null Inflect on the one hand and Overt Inflect on the other hand.



Finally, going down to individual words within a given alternation, bayesian statistics indicate for any given word the best a posteriori hypothesis regarding the way it was produced (Read, Null or Overt Inflect) given its own RT (p(H/RT)) and based on the a priori distribution of RT of the three categories (p(RT/H)). In French velar softening for example, in the Overt Inflect condition, the greater plausibility for *historicité* (*historique*) is the hypothesis Overt Inflect, while the best hypothesis for *électricité* (*électrique*) is Null Inflect. Therefore, the production of *électricité* (unlike the production of *historicité*) did not seem to involve any phonological activity.

Finally, note that significant RT effects within a given alternation cannot be due to just token frequency since the same lexical items were used.

Discussion. While on a macro scale the prediction is borne out (extra cognitive activity needs more time in production), all alternations do not follow the same pattern, and within a given alternation different words may follow different production logics. While schwa-ɛ (*achète* - *achetez*) and consonant-zero alternations in adjectives (*grand* - *grande*) seem to be lexicalized (e.g. against Tranel 1981: 251ff who concludes that the latter is the result of phonological computation), French velar softening appears to involve phonological computation – but not for all words. While *historicité*, *unicité*, *scientificité* and *technicité* follow this pattern, *électricité* falls into the category where concatenative activity but no phonological computation is observed. That is, *électricité* has climbed up one step on the lexicalization ladder. The reason for that is the well documented effect produced by token frequency: *électricity* (12,97 token per million) is much more frequent than the other words (token frequency between 0,00 and 0,41 per million). That is, unsurprisingly morphologically complex words stand a better chance to become lexicalized over time when they are frequent. This notwithstanding, the phonological process of velar softening is part of the phonology of present day natives who use it for less frequent (or unknown) words.