The relation between phonetic and phonological encoding in perception: Interactive or autonomous?

Michael Key

University of Massachusetts, Amherst

Much previous work on the influence of phonological knowledge in perception has found evidence that listeners 'compensate' for the effects of a native phonological rule in perception (e.g. Gaskell & Marslen-Wilson 1996, 1998, Darcy *et al.* in press). This suggests that listeners encode the signal in the currency of phonological categories, which serve as the input to some kind of perceptual phonology component.

However, this literature leaves unclear whether lower-level encoding in terms of the signal's acoustic properties is *autonomous* from higher-level phonological encoding. This is likely due to the fact that compensation effects have usually been demonstrated *recognition* tasks (e.g. word detection), which access listeners' phonological encoding and are therefore unable to detect any possible autonomous phonetic encoding. If phonetic encoding is autonomous from phonological encoding, there should be some perceptual task that can access it and such a task should fail to find compensation effects, which are the result of phonological encoding. Thus, the question is whether the relation between phonetic and phonological encoding in perception is *interactive* or *autonomous*, as sketched in (1).

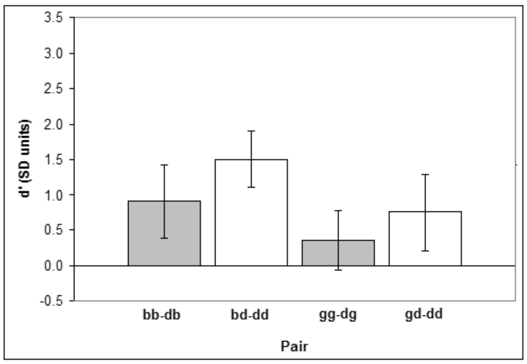
The studies reported here aim to clarify the interface between phonetic and phonological encoding in perception by using *discrimination* tasks aimed at tapping the hypothesized autonomous phonetic encoding. Discrimination tasks have previously been shown to be successful in accessing phonetic encoding by failing to produce categorical perception (Schouten *et al*. 2003); the present studies tested whether discrimination would similarly fail to produce compensation effects. This was done using non-word stimuli that match the well-known coronal place assimilation rule of English (EPA) (e.g. *good girl → goo***[g]** *girl)*. Non-words were used to control for possible influences from the lexicon. Previous recognition studies on place assimilation show that Eng. listeners can recognize assimilated pronunciations of words, particularly in viable assimilation contexts (Gaskell & Marslen-Wilson 1996), or identify the phoneme that matches the input representation in viable contexts (Gaskell & Marslen-Wilson 1998). If discrimination tasks access an autonomous phonetic encoding, then we predict that stimulus pairs that mimic the input and output of a phonological rule should not be more confusable than pairs that mimic two phonologically-unrelated legal forms of the language.

In Exp. 1, the Eng. voiced oral stops were were self-combined in all possible non-word stimuli of shape VC.CV in an AX discrimination task. The figures in (2) show place discrimination (*d′* units) by Eng. listeners (2a) and French listeners (2b) (used as a control) on pairs in which C_1 varies between a coronal and a non-coronal, thus mimicking EPA. The left pair of bars in each graph represents the pairs in which C_1 varies between [b] and [d], in a viable [b]-context (gray bar) or unviable [d]-context (white bar), while the right pair of bars shows the analogous contrast in context for [g] and [d]. Disc. of non-coronals from coronals is worse in viable contexts than in unviable contexts in Eng. listeners, but no worse in Fre. listeners, counter to the prediction. In Exp. 2, the same stimuli were presented in a 4IAX task; in this format, four stimuli, of which one is always different, are presented on each trial and the listener decides whether the different stimulus occurs in the first or second pair of stimuli. The figure in (3) shows no difference in disc. of non-coronals from coronals as a function of context in Eng. listeners, as predicted. Finally, the same tasks were tested using stimuli mimicking the Fre. rule of voicing assimilation (FVA) (e.g. *lac gelé → la***[g]** *gelé*), which follows up a recognition study comparing EPA and FVA (Darcy *et al.* in press). The figures in (4) show voicing disc. in Fre. listeners (4a) and Eng. listeners (4b) in an AX task. As in the case of place, disc. of voicing is much poorer in viable contexts (obstruents, left bar) than in unviable contexts (sonorants, right bar) for native listeners (Fre.), although it is also marginally poorer in non-native listeners (Eng.). The 4IAX task for FVA in Fre. listeners is underway.

These results speak in favor of an autonomous relationship between phonetic and phonological encoding in perception (1b). While the AX results (2a, 4a) show poorer disc. in viable contexts in native listeners, seemingly in support of the interactive architecture (1a), these results can be straightforwardly accounted for by assuming that listeners make use of a phonological encoding in this task, perhaps due to their uncertainty from trial to trial. The 4IAX results for place (3) show no difference in disc. as a function of context in native listeners, and support a model in which phonetic encoding of the signal can be accessed autonomously from the phonological encoding.

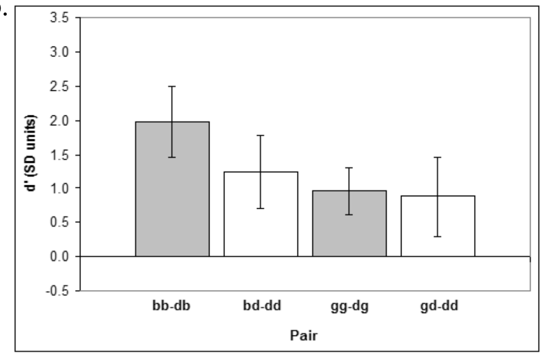
(1) a. *Interactive* → |acoustic form| → [phonological form] → /underlying form/
b. *Autonomous* → |acoustic form| **(→)** [phonological form] → /underlying form/

(2) a.



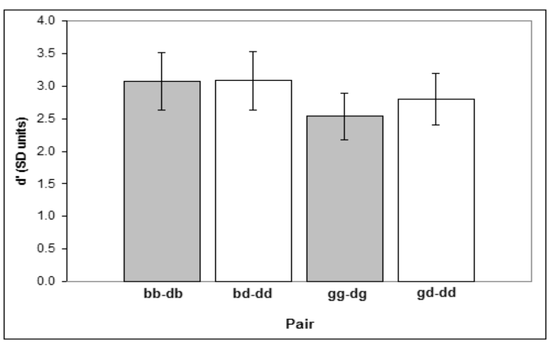
[b] vs. [d]: $t(19) = 2.85$, $p = .005$
[g] vs [d]: $t(19) = 1.81$, $p = .043$

b.



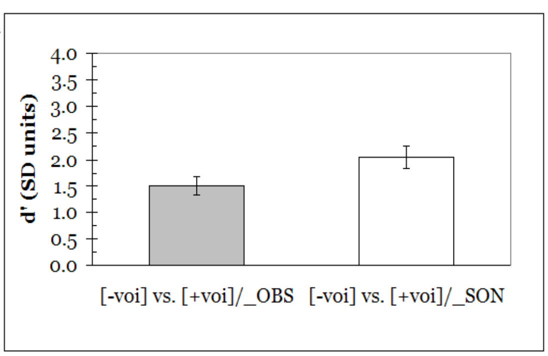
[b] vs. [d]: $t(15) = 3.70$, $p =.001$
[g] vs. [d]: $t(15) = 0.32$, $p > .10$

(3)



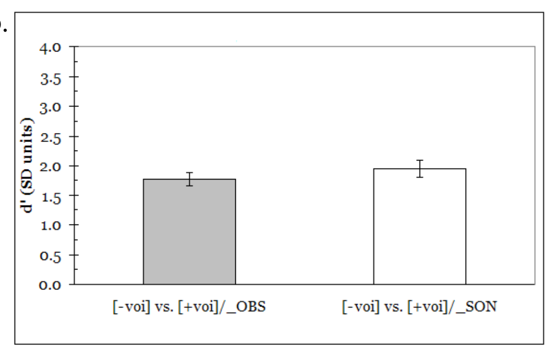
[b] vs. [d]: $t(15) = 0.14$, $p > .10$
[g] vs. [d]: $t(15) = 2.48$, $p = .013$

(4) a.



C2 Manner: $t(11) = -3.68$, $p = .002$

b.



C2 Manner: $t(20) = -2.46$, $p = .012$

References

Darcy, I., Ramus, F., Christophe, A., Kinzler, K., and E. Dupoux. (in press). Phonological knowledge in compensation for native and non-native assimilation In: F. Kügler, C. Féry & R. van de Vijver (eds.) Variation and Gradience in Phonetics and Phonology. Mouton De Gruyter.

Gaskell, M.G. and W.D. Marslen-Wilson (1996). Phonological variation and inference in lexical access. Journal of Experimental Psychology: Human Perception and Performance 22. 144-158.
Gaskell, M.G. and W.D. Marslen-Wilson (1998). Mechanisms of phonological inference in speech perception. Journal of Experimental Psychology: Human Perception and Performance 24. 380-396.
Schouten, B., Gerrits, E., and A. van Hessen (2003). The end of categorical perception as we know it. Speech Communication 41. 71–80.