

Autonomous relations between phonological and phonetic encoding in perception*

Michael Key

University of Massachusetts, Amherst

1. Introduction

Is knowledge of a phonological process applied in the same module as that in which auditory processing occurs? Most of the work in the literature on the influence of phonological knowledge in perception has focused on showing that it can cause listeners to identify sounds in a manner consistent with underlying values rather than surface values. In the case of knowledge of the English process of regressive place assimilation, Gaskell & Marslen-Wilson (1996) showed that *lea[m]* primes the word *lean* when pronounced in a viable assimilation context (e.g. *bacon*), but not when pronounced in a non-viable context (e.g. *gammon*); Gaskell & Marslen-Wilson (1998) showed that non-coronal stops pronounced in non-coronal contexts facilitates detection of coronal stops, in both word and non-word stimuli.

These results demonstrate that knowledge of a phonological process can influence the course of perception, but do not provide evidence for the *locus* of its application. As pointed out by Norris et al. (2000), Gaskell and Marslen-Wilson's (1998) detection effect in non-words supports the conclusion that knowledge of the place assimilation process is applied *pre-lexically* (as Gaskell and Marslen-Wilson argue), or after lexical processing, given that the context effect was stronger in words than in non-words.

In modeling terms, these results are consistent both with models that would locate the application of knowledge of place assimilation in the same module as auditory processing (*interactive models*, such as TRACE (McClelland & Elman 1986)), and those that would insist that it occurs in a separate module from auditory processing (*autonomous models*, such as Merge (Norris et al. 2000)). To distinguish these types of models, I compare the results of two psychophysically-different discrimination tasks designed to test for

*I would like to thank John Kingston for helping me with virtually every aspect of this project. Thanks are also due to Joe Pater, Diana Apoussidou, Della Chambless, Shigeto Kawahara, Dan Mash, Sharon Peperkamp, Sarah Watsky, and audiences at the UMass-MIT Phonology/Phonetics Meeting, Laboratory Phonology 11, and NELS 39 for their helpful assistance and feedback. All errors and deficiencies are mine alone.

the effects of knowledge of two assimilation rules. The contrast in discrimination performance in the two tasks is used in service of an argument that interactive models are insufficient to explain how phonological knowledge is applied in comprehension, a debate traditionally contested with lexical bias studies. I argue that an adequate model that includes phonological knowledge must posit modularity in some form, thus providing evidence for autonomous models. In addition, a response latency analysis of a subset of the results suggests that modularity may not be expressed as serially-ordered phonetic and phonological modules (e.g. as in Merge (Norris et al. 2000)), but is accommodated by a parallel architecture (e.g. as in Race (Cutler & Norris 1979)). I conclude with arguments that the results reflect a language-specific inference mechanism, rather than a feature cue parsing device (Gow 2003).

2. Phonological knowledge in perception

Some previous work examining how knowledge of a phonological process is applied in perception has found that listeners may *compensate* for the effects of a process, thereby apparently ‘undoing’ its effects by responding in a way that is consistent with underlying values rather than surface or phonetic values. In lexical decision or word detection experiments, listeners easily recognize words in which a process has applied, even though we might expect phonological mismatch to cause inhibited access to a single stored form. Lexical access is stronger when the process has applied in a context that matches the observed context for the process’ application in production, than when it applies in an inappropriate context (Gaskell & Marslen-Wilson 1996, Darcy et al. in press) In phoneme detection experiments, listeners recognize surface values which are the output of the process as tokens of an underlying category. Again, this effect obtains with particular strength in the appropriate context for the process (Gaskell & Marslen-Wilson 1998).

These results suggest that listeners tolerate phonological mismatch (by recognizing output values as instances of underlying categories) and that their willingness to do so depends on whether the mismatch occurs in a context in which a relevant process is known to apply. The dependency between toleration of mismatch and appropriateness of context is an important finding in understanding how speech is recognized because it does not support the proposal that words have a single lexical representation that is underspecified for place (cf. Lahiri & Marslen-Wilson 1991). Instead it recommends the proposal that listeners make use of a perceptual *inference mechanism*, which includes knowledge of the structural change and the context of a phonological process (Gaskell & Marslen-Wilson 1996, 1998, Darcy et al. in press).¹

Other work has found results that suggest that listeners do not compensate for the

¹Context dependence falsifies the claim that listeners rely on underspecified lexical entries because this theory fails to explain why e.g. *lea[m]* primes *lean* before a labial consonant (an appropriate context), but not before a non-labial consonant (an inappropriate context). There is also evidence that listeners know the *full structural description* of the process, rather than just the context. Additional data (not reported here) from the place discrimination experiments (see below) shows independent effects of the undergoer of assimilation and the match between the structural change and the context; these data will appear in Key & Kingston (in prep.).

effects of phonological processes in certain situations. Mitterer et al. (2003) tested to see whether listeners would compensate for the effects of a Hungarian liquid assimilation process ($/l/ \rightarrow [r]/_ [r]$). Mitterer et al. compared the results of two experiments run on both Hungarian and Dutch listeners, who were used as a control group. In a “l”-“r” identification task presented in a $[r]$ context, Hungarian but not Dutch listeners compensated for the effects of the process – their decision criterion was shifted in favor of “l” responses. This result would predict that Hungarian (but not Dutch) listeners would perform poorly at discriminating $/l/$ from $/r/$ in the context of $[r]$. However, a 4IAX discrimination task found no differences in performance between the two groups; both groups performed rather poorly in the $[r]$ context. The equally poor performance by both groups suggests an explanation based on an inherent phonetic subtlety in the contrast between $[l]$ and $[r]$ (in a $[r]$ context). For instance, discrimination may be poor in both groups due to a tendency to misparse spectral information which properly belongs to the ($[r]$) context as belonging to the $[l]$ or $[r]$ target.

Kabak & Idsardi (2003) examined various Korean phonotactic restrictions in an AX discrimination study on Korean and English listeners. They wanted to see if native knowledge of a given phonotactic restriction and its repair would cause an illicit sequence and its repaired alternative to be poorly discriminated. The results showed that the subset of restrictions explicitly involving illicit coda or onset segments (e.g. $/c\text{-}t^h/ \rightarrow [t^h]$) induced the predicted inability to discriminate illicit and repaired sequences (e.g. $*c.t^h$ vs. $t.t^h$) by Korean listeners (relative to English listeners). In contrast, restrictions on particular clusters in which each segment *is* syllabically licensed (e.g. $/l\text{-}n/ \rightarrow [l.n]$) did not produce poor discrimination of the analogous pairs (e.g. $*l.n$ vs. $l.l$). In Korean, knowledge of phonological processes that repair illicit syllables seem to be applied in a different perceptual module than those that repair illicit sequences.²

The two cases just discussed show that discrimination tasks are useful for investigating how phonological knowledge is applied in perception. First, we can observe that knowledge of one phonological process may be applied differently from that of another, even within the same task (Kabak & Idsardi 2003). Second, discrimination results may not display a translation of compensation effects demonstrated in other tasks that make explicit reference to categories (Mitterer et al. 2003). The latter result in particular suggests how discrimination tasks can be used to distinguish competing models of perception. In particular, Kabak and Idsardi’s data are compatible with the predictions of autonomous models, but are difficult to reconcile with the predictions of interactive models.

3. Experiments: Assimilation processes and discrimination

Does native knowledge of a place or voicing assimilation process influence the discriminability of place or voicing contrasts? Compensation effects for place assimilation have been demonstrated by Lahiri & Marslen-Wilson (1991), Gaskell & Marslen-Wilson (1996,

²A more vacuous alternative conclusion is that knowledge of Korean sequence-repairing processes is not applied at all in perception. This is doubtful in light of the fact that, although Koreans’ performance was better on these pairs than those designed to test the syllable-repairing processes, it nonetheless remained below English listeners’ performance.

1998)), and compensation for voicing assimilation has been shown by Snoeren et al. (2006). In addition, a cross-linguistic word detection study of both processes was carried out on English and French listeners by Darcy et al. (in press). Darcy et al. found that listeners compensated more for the effects of the assimilation process of which they had native knowledge than for the process that was non-native. English and French listeners were ideal for this design because English has a process of place assimilation, but French does not; conversely, French has a process of voicing assimilation, but English does not.³

Place and voicing assimilation processes are a good empirical domain because compensation effects for both types of processes have previously been demonstrated, allowing for a comparison between identification and discrimination results. Does the compensation effect, which is clearly based on phonological encoding, translate into poor discrimination performance in the relevant context(s) for the processes' application? If discrimination performance is also based on phonological encoding, the known compensation effects should be reflected in just this way. However, we know that discrimination sometimes reflects attention to sub-categorical stimulus differences, in which case we do not expect compensation effects to translate. In interpreting Mitterer et al.'s findings, Darcy et al. (in press) suggested that the compensation effect found in the Hungarian identification results might only hold at a (later) phonologically-influenced stage, which is why the phonetically-driven 4IAX task produced equally poor performance by both groups of listeners. If discrimination is instead based on a continuous phonetic encoding, we predict no difference in discriminability as a function of context. These two possible scenarios and their implications for both autonomous and interactive models are laid out in (1).

- (1) *Implications of basis of discrimination for models of perception*
- | | |
|---|---|
| a. Disc. based on phonological encoding | Consistent with interactive
Consistent with autonomous |
| b. Disc. based on phonetic encoding | Inconsistent with interactive
Consistent with autonomous |

The comparison in (1) implies a fundamental asymmetry between autonomous and interactive models; interactive models are disconfirmed by phonetically-driven performance, but autonomous models are not disconfirmed by phonologically-driven performance. That the former is true is evinced by interaction's central claim that feedback is always beneficial for speech recognition, so it is antithetical to suppose that available linguistic knowledge would not affect performance. The latter is true because autonomy does not deny top-down influence from linguistic knowledge under any conditions. Rather, it specifies the conditions under which the application of linguistic knowledge is available. The present study investigates these conditions with respect to knowledge of two phonological processes by examining the role of the format of discrimination tasks.

³English marginally has a coda devoicing tendency (e.g. *bi[k] place*, however it lacks a voicing process (e.g. **chea[b] glove*).

3.1 Experiment 1: English place assimilation

English place assimilation (EPA) causes a coda stop (or nasal) to be pronounced as a labial (2a) or a velar (2b)) when a following onset labial or velar stop is present. Therefore, labial and velar stops are viable EPA contexts, while coronal stops are unviable contexts (2c).

- (2) *EPA exemplified*
- a. *goo*[g] *kisser*
 - b. *freigh*[p] *bearer*
 - c. **war*[n] *dinner*

Methods and participants English voiced oral stops /b,d,g/ were combined in all possible non-word, naturalistic stimuli of the form VC.CV, where V = /a/. The ISI was 500 ms. The participants were 20 native speakers of American English from the UMass community, and 16 native Continental or Québécois French speakers from the UMass or McGill University community. The French listeners provided a control group for the EPA studies, as in the study of Darcy et al. (in press). All subjects reported normal hearing. The stimuli were presented over Sennheiser HD pro headphones in a quiet room, and participants responded “same” or “different” by pressing the corresponding button on a Cedrus RB-834 response pad. The software used for the presentation of the trials was SuperLab (PC - v. 2.0.4).

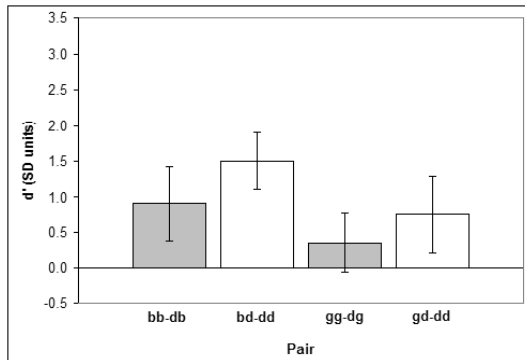
Results Figure 1 shows place discrimination performance measured using the d' statistic of Detection Theory (Macmillan & Creelman 2005) by English listeners and French listeners on pairs in which C1 varied between a coronal and a non-coronal stop, thus mimicking EPA. The left pair of bars in each graph represents the pairs in which C1 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].

Discrimination of non-coronals from coronals is worse in viable contexts than in unviable contexts in English listeners, but no worse in French listeners, which supports the phonological encoding basis for discrimination (1a). As stated in (1a), the homomorphism between previous compensation effects and these results is compatible with both interactive and autonomous models. Interactive models predict this outcome because they deny modularity, and therefore knowledge of EPA is expected to be applied both in identification and discrimination tasks.

However, autonomous models cannot be disconfirmed because of independent evidence that the AX task encourages discrimination on the basis of phonological categories. To give autonomous models a fair test, it is worth presenting the same EPA stimuli in a 4IAX discrimination paradigm, which was shown by Gerrits & Schouten (2004) to contrast with the AX task by inducing discrimination on the basis of sub-categorical phonetic differences. The 4IAX task failed to find non-discrimination of neutralizing liquids in Hungarian listeners relative to Dutch listeners (Mitterer et al. 2003). This manipulation provides a reasonable chance for finding evidence for a phonetic basis of discrimination scenario (1b). In the 4IAX format, four stimuli, of which one is always different from the

Michael Key

English listeners



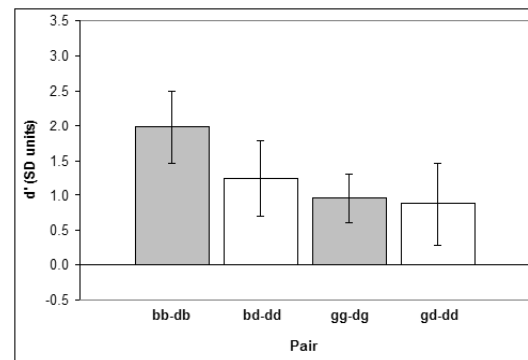
[b] vs. [d]/-[b] < [b] vs. [d]/-[d]:

One-tailed $t(19) = 2.85, p = .005$

[g] vs. [d]/-[g] < [g] vs. [d]/-[d]:

One-tailed $t(19) = 1.81, p = .043$

French listeners



[b] vs. [d]/-[b] $\not<$ [b] vs. [d]/-[d]:

One-tailed $t(15) = 3.70, p = .001$

[g] vs. [d]/-[g] $\not<$ [g] vs. [d]/-[d]:

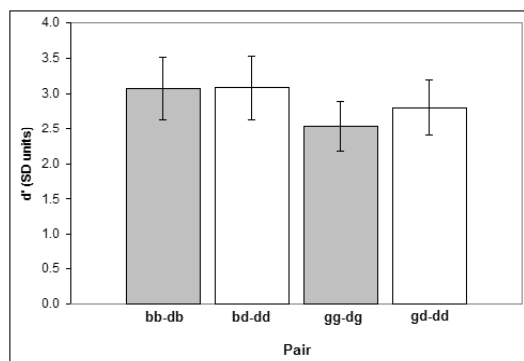
One-tailed $t(15) = 0.32, p > .10$

Figure 1: Place discrimination as a function of context (AX task)

other three, are presented on each trial and the listener decides whether the different stimulus occurs in the first or second pair of stimuli. The psychophysics of the 4IAX task are described as a decision based on a comparison of the size of physical differences within the first vs. second stimulus pair (e.g. Rousseau & Ennis 2001).

Figure 2 shows discrimination of non-coronals from coronals as a function of context by English listeners in the 4IAX task. Since no difference in discriminability as a function of context was found in French listeners in the AX task, there is no reason to expect a difference in the 4IAX task, hence this task was only presented to English listeners.

English listeners



[b] vs. [d]/-[b] $\not<$ [b] vs. [d]/-[d] : One-tailed $t(15) = 0.14, p > .10$

[g] vs. [d]/-[g] < [g] vs. [d]/-[d] : One-tailed $t(15) = 2.48, p = .013$

Figure 2: Place discrimination as a function of context (4IAX task)

Paired one-tailed t -tests support a context-based difference in discriminability between [g]

and [d], but not between [b] and [d]. Given the lack of a *consistent* effect of context in the 4IAX task (cf. the effect in the AX task), data from additional discrimination pairs not presented here, and an acoustic analysis of the stimuli, I argue that the context-based difference in discriminability of [g] and [d] in the 4IAX task reflects a *phonetic* rather than a *phonological* context effect. The full evidence and argument will appear in Key & Kingston (in prep.).

Putting aside the [g]-[d] data, the 4IAX task is apparently being performed on the basis of a phonetic encoding (Gerrits & Schouten 2004), due to the lack of a context effect in the discriminability of [b] vs. [d] in the 4IAX task compared with the context effect found with the same stimuli in the AX task. This phonetic basis for discrimination is only consistent with autonomous models, because interactive models cannot explain why knowledge of EPA is applied to the same stimuli in one task, but fails to apply in another.

3.2 Experiment 2: French voicing assimilation

French voicing assimilation (FVA) applies to change the voicing value of a coda stop to match the value of a following obstruent (3a,b), but not before a sonorant (3c). Therefore, voiced and voiceless obstruents are viable contexts, while sonorants are unviable contexts.

(3) FVA examples

- a. *ro[p]e sale* 'dirty dress'
- b. *la[g] gelé* 'frigid lake'
- c. **ca[b]e neuve* 'new cape'

Methods and participants Non-word VC.CV stimuli were used again. The oral stops /p,b,t,d,k,g/ were the codas, and voiceless fricatives /f,s,ʃ/, voiced fricatives /v,z,ʒ/, /l/, and /n/ were the onsets; /p,b/ were paired with /ʃ,ʒ/, /t,d/ were paired with /f,v/, and /k,g/ were paired with /s,z/. The stimuli were pronounced by a native French speaker. The participants were 11 native Continental or Québécois French speakers from the UMass and McGill University communities, and 20 native speakers of American English from the UMass community. The English listeners provided a control group for the FVA studies, as in the study of Darcy et al. (in press). All subjects reported normal hearing. The same procedures used in Experiment 1 were applied here.

Results Figure 3 shows voicing discrimination performance by French and English listeners in an AX task. As in the case of place in the AX task, discrimination of voicing is poorer in viable contexts (before a following obstruent, left bar) than in unviable contexts (before a following sonorant, right bar) for French listeners, although it is somewhat poorer in English listeners.

The repeated measures ANOVAs show a main effect of C_2 (the context) for both groups of listeners. The ANOVAs with Helmert-coded contrasts show two language-based differences: (i) discrimination of the voicing contrast is significantly poorer for both groups in obstruent contexts than in sonorant contexts, but the difference is much greater for French listeners, who showed a d' difference of 0.55 between obstruents and sonorants,

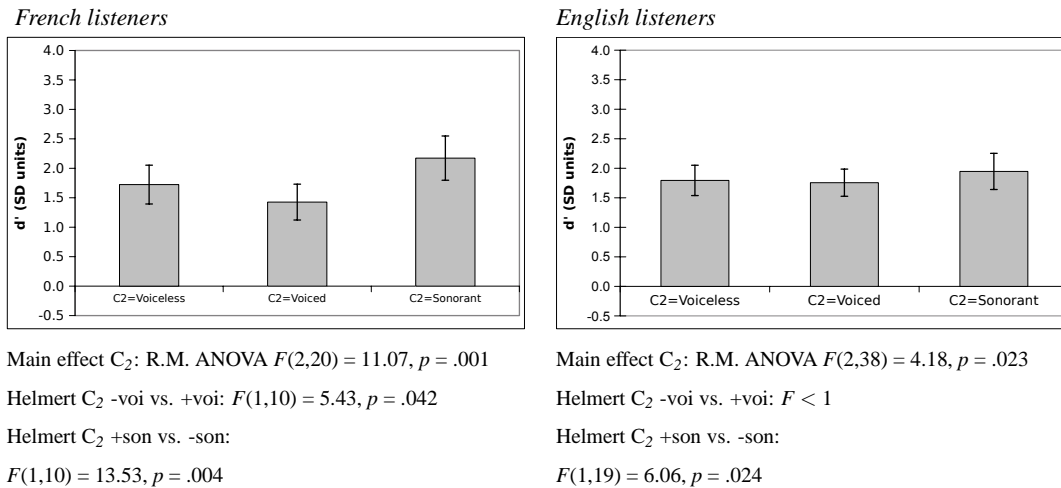


Figure 3: Voicing discrimination as a function of context (AX task)

compared to English listeners, who showed a d' difference of only 0.17; (ii) discriminability is poorer in voiced obstruent contexts than in voiceless ones for French listeners, but there is no difference between these contexts for English listeners. This suggests the voicing assimilation process in French is more of a voicing process than a devoicing process, which accords with the phonetic description of FVA in Snoeren et al. (2006), and with the perceptual asymmetry reported in Darcy & Kügler (2007). To complete the cross-linguistic study, I am in the process of collecting data for the 4IAX task for FVA on French listeners. These data will be presented in Key & Kingston (in prep.).

4. Discussion

4.1 Interactive vs. autonomous models

As argued above, the task format effects found in the EPA studies are incompatible with interactive models, and consistent with autonomous models. In this section, I lay out two basic types of autonomous models, and present some additional data from the EPA AX experiment that motivates a choice between the two.

The various extant autonomous models all share the feature that they posit modularity in some form. They can be partitioned with respect to how modularity is expressed. The first type expresses modularity by proposing that modules are *serially-ordered*, as in the Merge model (Norris et al. 2000). While the operation of a later module may be affected by the operation of any module that is ordered before it, any module that has no modules ordered before it operates autonomously from the influence of any other module in the system. The most obvious application of this feature is to explain the failure of linguistic knowledge to influence the outcome of lower-level processes. In the case of the EPA discrimination results, we can analyze the context effect found in the AX task as being the product of a module in which knowledge of EPA is applied, whereas the lack

of a context effect found in the 4IAX task would reflect discrimination performance in the phonetic module. If we assume that the phonetic module is ordered before the module in which knowledge of EPA is applied, we can explain why that knowledge has no effect in the 4IAX task. Babel & Johnson (2008) argued in support of this view, based on their finding that a speeded AX task uncovered no difference in the discriminability of intervocalic /f/ vs. /θ/ or /h/ vs. /x/ in English and Dutch listeners, despite the fact that each language only contrasts one of the two pairs.

An alternative conception of modularity is lies in assuming that distinct modules may operate in parallel, as in the Race model (Cutler & Norris 1979). In this type of model, responses may be based on either the auditory or phonological module, leaving it to factors such as task and stimulus manipulations to determine which module will win the race. Applied to the EPA results, this type of model also attributes the context effect to the module in which knowledge of EPA is applied and the lack of an effect to the phonetic module, but does not claim that the latter module must operate prior to the former. This point of contrast between the two types of autonomous models sets up a testable question: does the context effect (i.e. lowered d' values) found in the AX task correlate with increasing response latencies? If so, this is good evidence in favor of serial rather than parallel modularity because we predict fast responses to be produced by the phonetic module, while longer responses should be generated by the module in which knowledge of EPA is applied. This prediction is subject to the additional assumption that the AX task can *in principle* be performed on the basis of a phonetic encoding, particularly when response latencies are short.

Figure 4 presents additional data from the EPA AX experiment to test the prediction that the context effect will strengthen with increasing response latency. A median split of latencies was done for each English listener, and mean d' values were calculated separately for each discrimination pair below and above that listener's median. We thus have a mean d' score for each pair at short latencies ('RT < median < 1500') and at long latencies ('1500 > RT > median').⁴ The serial view predicts that the pairs in which the context effect was found (bb-db and gg-dg) will have significantly smaller mean d' values in the '1500 > RT > median' condition than in the 'RT < median < 1500' condition. This difference is found for the gg-dg pair (paired, one-tailed $t(19) = 2.26, p = .018$), but not for the bb-db pair ($t(19) = 0.07, p > .10$).⁵

However, the serial view also predicts a greater difference between viable and unviable pairs at long rather than short latencies. The difference in mean d' scores between gg-dg and gd-dd is significant at *both* long latencies ($t(19) = -2.45, p = .012$) and short latencies ($t(19) = -2.02, p = .029$). The difference between bb-db and bd-dd is actually *positive* in sign at long latencies ($t(19) = 0.74, p = > .10$) and negative at short latencies ($t(19) = -2.30, p = .016$), which is opposite the prediction of the serial view.

The results of the latency analysis suggest that a parallel autonomous model may be on the right track. Another line of attack on the architecture question would be to compare

⁴1500 ms was the maximum time listeners were allotted to respond on each trial.

⁵Discrimination performance across the four pairs was inversely correlated with latencies: the mean d' score in the RT < median < 1500 condition was 0.98, while in the 1500 > RT > median condition it was 0.61.

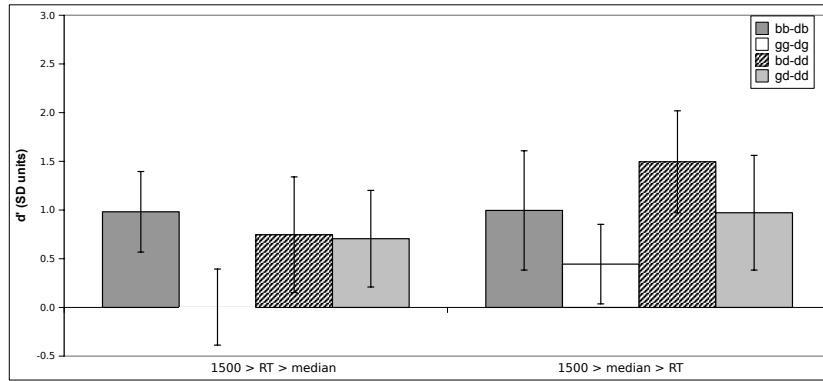


Figure 4: Individual median latency split of AX EPA performance

ERPs for viable pairs like bb-db with unviable pairs like bd-dd, thus following similar work by Mitterer & Blomert (2003), who found a difference in the size of the MMN associated with these pair types in Dutch listeners. They concluded that the smaller MMN in viable pairs reflects prelexical application of knowledge of Dutch place assimilation. While I share this interpretation, a ‘priming’ design that measures two ERPs – the PMN (~200-400 ms) and the P3 (~500 ms) – would be particularly useful for determining whether phonological knowledge is applied at the same time range as phonetic processing, or whether it is only applied after phonetic processing. We could motivate the former scenario if Mitterer and Blomert’s MMN results also obtain at both the PMN and P3 potentials. The latter scenario would be supported if their MMN results obtain at the P3, but not at the PMN. This possibility is pursued in Key & Kingston (in prep.).

4.2 Are these processes phonological?

The pattern of performance observed in both AX tasks presumably derives from the presence of a bias from native knowledge of the relevant process. This bias could be modelled in multiple ways; for the sake of the present discussion, I will assume it takes the form of an inference mechanism that applies the phonological production computation in reverse (i.e. mapping non-coronal–non-coronal sequences onto coronal–non-coronal sequences, or – α voice α voice sequences onto α voice – α voice sequences.).

Recent work by David Gow and others argues for an alternative to phonological inference mechanisms as the explanation for compensation results. Gow (2003) and Van Worthe et al. (2006) propose that an acoustic feature cue parsing mechanism is responsible for compensation effects for place or voicing changes that may or may not reflect the pattern of assimilation in any given language. The mechanism parses conflicting phonetic cues and attributes them to their segments of origin. Therefore, no linguistic experience with any particular pattern of assimilation is required for compensation effects on this view – a claim that stands in contrast with that made by a language-specific phonological inference account.

The lack of *symmetric* depression of discrimination performance in the AX tasks

seems to constitute a failed prediction of a feature-cue parsing mechanism. The feature cue account predicts that listeners would discriminate viable pairs like bb-db and unviable pairs like bd-dd equally well, counter to the findings of the AX place experiment. Similarly, this account fails to capture the effect of the context's voicing value on AX voicing discrimination by French listeners. Finally, the feature cue parsing account fails to explain the language-based differences found in both the place and voicing experiments, since the feature cue parser operates at a perceptual level, and is thus universal.

In any case, the feature cue parsing account is inapplicable to the present data because the VC portions were recorded separately from the CV portions, and therefore lack cues to coronal place. Note that this is not an experimental idealization of the assimilation processes. An EPG study by Ellis & Hardcastle (2002) showed that complete place assimilation is a fairly common pattern in natural English speech, while Darcy & Kügler (2007) showed that both voiced and voiceless Cs that result from FVA do not differ from underlying voiced or voiceless Cs in terms of the proportion of the closure in which voicing is present, nor in the duration of a preceding vowel. Therefore, the feature cue parsing only provides a possible explanation for results obtaining with partial assimilation. In contrast, the context and language-based effects found with total and partial assimilation are accommodated by a language-specific inference account.

References

- Babel, M., & K. Johnson. 2008. Accessing psycho-acoustic perception with speech sounds. Oral presentation at Laboratory Phonology 11. Victoria University of Wellington, New Zealand.
- Cutler, A., & D. Norris. 1979. Monitoring sentence comprehension. In *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, ed. W.E. Cooper & E.C.T. Walker, 113–134. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Darcy, I., & F. Kügler. 2007. Phonological context effects for voicing and devoicing in French. In *Proceedings of the XVI International Congress of Phonetic Sciences*, ed. J. Trouvain & W. Barry, 1025–1028.
- Darcy, I., F. Ramus, A. Christophe, K. Kinzler, & E. Dupoux. in press. Phonological knowledge in compensation for native and non-native assimilation. In *Variation and gradience in phonetics and phonology*, ed. F. Kügler, C. Féry, & R. van de Vijver. Berlin: Mouton de Gruyter.
- Ellis, L., & W.D. Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* 30:373–396.
- Gaskell, M.G., & W.D. Marslen-Wilson. 1996. Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 6:110–125.
- Gaskell, M.G., & W.D. Marslen-Wilson. 1998. Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 24:380–396.

Michael Key

- Gerrits, E., & M.E.H. Schouten. 2004. Categorical perception depends on the discrimination task. *Attention, Perception, & Psychophysics* 66:363–376.
- Gow, D. 2003. Feature parsing: Feature cue mapping in spoken word recognition. *Attention, Perception, & Psychophysics* 65:575–590.
- Kabak, B., & W. Idsardi. 2003. Syllabically conditioned perceptual epenthesis. In *Proceedings of BLS 29*, 233–245.
- Key, M., & J. Kingston. in prep. Discrimination of place and voicing contrasts in consonant clusters: A comparison of English and French listeners.
- Lahiri, A., & W.D. Marslen-Wilson. 1991. The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition* 38:245–294.
- Macmillan, N., & D. Creelman. 2005. *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McClelland, J.L., & J.L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18:1–86.
- Mitterer, H., & L. Blomert. 2003. Coping with phonological assimilation in speech perception: Evidence for early compensation. *Attention, Perception, & Psychophysics* 65:956–969.
- Mitterer, H., V. Csépe, & L. Blomert. 2003. Compensation for phonological assimilation in perception: Evidence from hungarian liquid assimilation. In *Proceedings of the XV International Congress of Phonetic Sciences*, ed. M.J. Solé, D. Recasens, & J. Romero, 2321–2324. Barcelona, Spain.
- Norris, D., J.M. McQueen, & A. Cutler. 2000. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23:299–370.
- Rousseau, B., & D. Ennis. 2001. A Thurstonian model for the dual pair (4IAX) discrimination method. *Attention, Perception, & Psychophysics* 63:1083–1090.
- Snoeren, N., P. Hallé, & J. Seguí. 2006. A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34:241–268.
- Van Worthe, D., J. Celi, & D. Gow. 2006. Parsing voicing cues: A context effect predicted by feature cue parsing theory. *Journal of the Acoustical Society of America* 119:3242–3242. Abstract.

Department of Linguistics
South College
University of Massachusetts
Amherst, MA 01003

mkey@linguist.umass.edu