4pSC23

The Perceptual Time Course of Coarticulatory Nasalization

Patrice S. Beddor, Julie Boland, Andries Coetzee and Kevin B. McGowan

Department of Linguistics, University of Michigan, Ann Arbor

Abstract

Listeners' moment-by-moment processing of anticipatory vowel nasalization and a following nasal consonant was investigated. English-speaking participants' eye movements were monitored as they heard instructions to look at one of two pictured objects on a computer screen. Trials included pictured pairs for naturally produced words of the form CVNC-CVC (e.g., bend-bed), CVNC-CVNC (bend-bent), and CVC-CVC (bed-bet). Vowels in CVNC words were coarticulatorily nasalized.

Results: When participants heard CVNC (bend), they fixated the correct picture earlier when the competing picture was CVC (bed)—that is, when the vowel in the competitor would be expected to be non-nasal—than when the competitor was another CVNC (bent). The more heavily nasalized the vowel, the faster the correct fixations to CVNC, indicating that participants used coarticulatory nasalization to plan their eye movement in CVNC-CVC trials. A non-nasalized vowel was not similarly helpful for selecting CVC over CVNC, although there was an effect of voicing context due in part to vowel quality factors distinct from nasalization.

Conclusion: That Ṽ is more informative about CVNC than V is about CVC suggests that listeners are not simply responding on the basis of familiarity with ṼN and VC sequences, and is consistent with perceptual compensation (e.g., Mann & Repp 1980). Overall, results indicate that listeners actively track information in the coarticulated signal and that this information speeds the time course of lexical activation.

Background

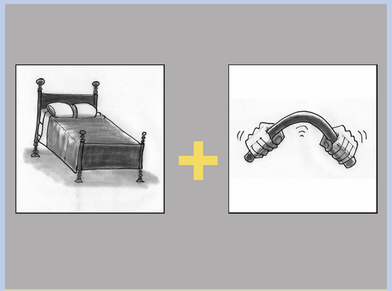
- Coarticulatory cues influence listeners' decisions:
  - *Identification*: Vowel nasalization alone, with no N, elicits N percepts in American English (Malécot 1960; Beddor *et al.* 2007).
  - *Gating*: When listeners hear CV fragments, they are more likely to report that the full word is CVN (rather than CVC) if the vowel is nasalized (Lahiri & Marslen-Wilson 1991; Ohala & Ohala 1995).
  - *Reaction time*: Absence of anticipatory vowel nasalization slows listeners' reaction times in identifying N vs. C (Fowler & Brown 2000).
- There is consensus that coarticulatory information affects listener judgments, but theorists disagree on the perceptual *usefulness* of the information:
  - Coarticulation is not ideal for the listener because it introduces variation (Ladefoged 2001, Tatham & Moreton 2006) or because it renders contrasts less distinct (Lindblom 1990).
  - Coarticulation can be useful, enhancing information if the quantal information is obliterated (Stevens & Keyser in press).
  - Coarticulation is useful information that aids listeners in tracking the speakers' articulatory gestures (Fowler 1996).

Theoretical Goals

We hypothesize that coarticulation is perceptually useful: listeners process the input acoustic signal using the rich, time-varying information in coarticulation to determine what speakers are saying.

- Is listeners' use of coarticulatory information **dynamic**?
  - As the acoustic signal unfolds over time, do listeners' perceptual assessments evolve?
  - Do English-speaking listeners use coarticulatory Ṽ to predict N?
  - Does absence of coarticulatory nasalization (i.e., an oral vowel) predict C?

To answer these questions, we measured the **time course** of listeners' perception of coarticulatory nasalization using eye tracking. Previous eye tracking studies have shown that inappropriate (cross-spliced) coarticulatory cues slow listeners' fixations (Dahan *et al.* 2001). Our study investigates listeners' use of appropriate coarticulation.



Method: Tracking eye movements to audio-visual stimuli

Stimuli

T-NT	T-D	D-ND	NT-ND
bet - bent	bet - bed	bed - bend	bent - bend
watt - want	watt - wad	wad - wand	want - wand
set - scent	set - said	said - send	scent - send
let - lent	let - lead	lead - lend	lent - lend
wet - went	wet - wed	wed - wend	went - wend

- *Visual:* Pencil & charcoal pictures of each word.
- *Audio:* Manipulated natural-speech recordings presented over headphones. Each CṼNT-CVT and CṼND-CVD trial had 3 types: light Ṽ, heavy Ṽ, and Deleted-N.

Manipulation 1: ṼN stimuli with two degrees of vowel nasalization:
- *Light nasalization*: 40% nasal vowel, 60% oral
- *Heavy nasalization*: 80% nasal vowel, 20% oral

Manipulation 2: Ṽ but no N ([n] deleted from bent, bend, want, wand, etc.)

Participants

16 native English-speaking undergraduate students at the University of Michigan.

Picture familiarization Participants learned the name of each picture prior to the main task.

Eye-tracking task

For each trial, participants heard:
- Look at the pictures (e.g., *bed* & *bend*)
- Fixate cross;
- Now look at [target] (e.g., *bend*. Cross disappears at onset of word.)

Predictions

Hypothesis 1: Ṽ signals an upcoming N.

Auditory CṼNC

1a. Initial correct fixations to CṼNC (bent) should occur *earlier for CṼNC-CVC (bent-bet) than for CṼNC-CṼNC (bent-bend)* because Ṽ is not a differentiating cue for CṼNC-CṼNC.

1b. In CṼNC-CVC (bent-bet) trials, initial correct fixations should occur *earlier for CṼNC with heavy nasalization than for CṼNC with light nasalization.*

Hypothesis 2: Presence of coarticulatory information is more informative than its absence. (Listeners are not simply responding on the basis of familiarity with ṼN and VC.)

Auditory CVC

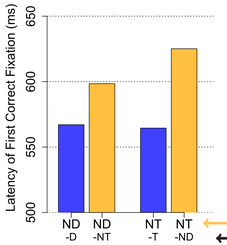
2. Initial correct fixations to CVC (bet) should occur *about the same time* in CVC-CVC (bet-bed) and CVC-CVNC (bet-bent) trials.

Hypothesis 3: English-speaking listeners use Ṽ in voiceless (bent) more than in voiced (bend) contexts because Ṽ tends to be heavily nasalized and N is often absent in VNCvoiceless (e.g., Malécot 1960, Cohn 1990).

Voicing

3. Fixations to Deleted-N trials should be less accurate, and show a greater decrease over time, for CṼ(N)D than for CṼ(N)T.

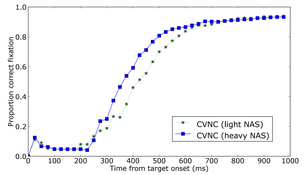
Results: Prediction 1a



1a. ✓ Prediction is upheld: Listeners correctly fixate CṼNC image faster when the competitor would have an oral vowel (CVC: blue bars) than when it would have a nasal vowel (CṼNC: orange bars). Results are shown for heavy nasalization.

Results: Prediction 1b

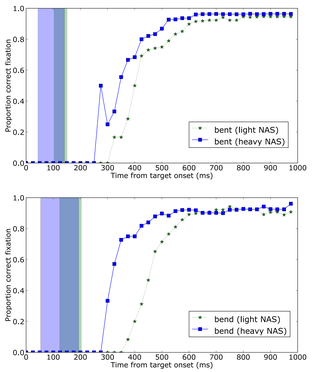
1b. ✓ Prediction is upheld: In CṼNC-CVC (bent-bet) trials, the more nasalized the vowel, the earlier listeners correctly fixate the CṼNC image: listeners use Ṽ as information about N before N is heard.



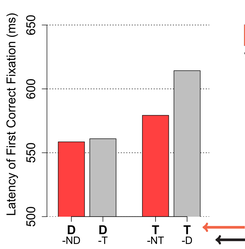
Results: 1b illustration

Light bar: heavy Ṽ
Dark bar: light Ṽ
Green bar: N onset

Note: a 200ms delay is generally assumed for programming an eye movement (e.g., Dahan *et al.* 2001).



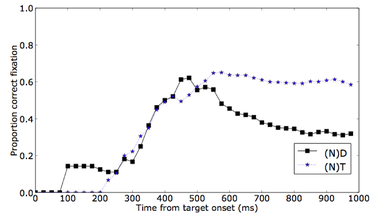
Results: Prediction 2



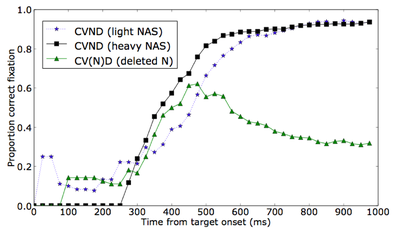
2. ✓ Prediction is upheld for voiced, but not voiceless, trials:
- *Voiced:* correct fixations to CVD are equally fast in CVD-CVND (where oral V is potentially informative; red bar) and CVD-CVT (where V orality is not informative; grey bar).
- Voiceless: contrary to predictions, fixations to CVT are faster in CVT-CVNT than in CVT-CVD (see discussion).

Results: Prediction 3

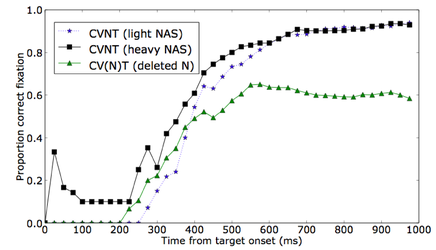
3. ✓ Prediction is upheld: Fixations to Deleted-N trials are overall less accurate and show a greater decrease over time to CṼ(N)D than to CṼ(N)T.



Results: Summary (voiced)



Results: Summary (voiceless)



Discussion

- **Listeners use Ṽ to plan their eye movement.**

CṼNC

CṼNC was fixated earlier when the vowel in the competitor was expected to be non-nasal than when it would be nasal. The more nasalized the V, the earlier CṼNC was fixated.

- **Listeners are less likely to use V.**

CVC

CVD was not fixated earlier when the vowel in the competitor was expected to be nasal than when it would be non-nasal. For CVT, oral V did lead to earlier fixations in CVT-CVNT than in CVT-CVD comparison. The voicing effect is due in part to vowel quality differences (in watt-want but not wad-wand) distinct from nasality.

- **Listeners use context-specific patterns of nasalization.**

Voicing

In production, English has heavier vowel nasalization and shorter N in VNT than VND. In perception, Deleted-N (with heavy Ṽ) elicits more fixations to CVNT than to CVND. That V is useful in CVT but not CVD may also be partly due to heavy nasalization in VNT: non-nasal V is useful only when the competitor would be expected to have especially heavy nasalization.

Overall the results indicate that listeners actively track information in the coarticulated signal and that this information speeds the time course of lexical activation.

References

Beddor, Patrice Speeter, Anthony Brasher, & Chandan Narayan. 2007. Applying perceptual methods to the study of phonetic variation and sound change. In *Experimental Approaches to Phonology*, ed. by Maria-Josep Sole, Patrice Beddor, & Manjari Ohala, 127–143. Oxford University Press. In Honor of John Ohala.

Cohn, Abigail C. 1990. Phonetic and phonological rules of nasalization. In *UCLA Working Papers in Linguistics* , volume 76. University of California at Los Angeles.

Dahan, D., J. S. Magnuson, M. K. Tanenhaus, & E. M. Hogan. 2001. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language & Cognitive Processes* 16.507–534.

Fowler, C. A., & J. M. Brown. 2000. Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Percept and Psychophysics* 62.21–32.

Fowler, Carol A. 1996. Listeners do hear sounds, not tongues. *JASA* 99.1730–1741.

Ladefoged, Peter. 2001. *A Course in Phonetics*. New York: Harcourt Brace Jovanovich, Inc., fourth edition.

Lahiri, A., & W. Marslen-Wilson. 1991. The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38.245–294.

Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In *Speech Production and Speech Modelling*, ed. by W. J. Hardcastle & A. Marchal, 403–439. Kluwer Academic.

Malécot, A. 1960. Vowel nasality as a distinctive feature in american english. *Language* 36.222–229.

Mann, Virginia A., & Bruno H. Repp. 1980. Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception and Psychophysics* 28.213–228.

Ohala, J.J., & M. Ohala. 1995. Speech perception and lexical representation: the role of vowel nasalization in hindi and english. In *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, ed. by B. Connell & A. Arvaniti, 41–60. CUP.

Stevens, K. N., & S. J. Keyser. in press. Quantal theory, enhancement and overlap. *Journal of Phonetics*.

Tatham, M., & K. Moreton. 2006. *Speech Production and Perception*. Palgrave.

Department of Linguistics, University of Michigan, Ann Arbor, MI

e-mail: beddor@umich.edu or clunis@umich.edu

http://ling.lsa.umich.edu/