The Perception of Nasalized Vowels in American English: An Investigation of On-line Use of Vowel Nasalization in Lexical Access

Joyce McDonough[1], Heike Lenhert-LeHouiller[1], Neil Bardhan[2]

[1]*Linguistics and* [2]*Brain & Cognitive Sciences, University of Rochester, Rochester New York*

joyce.mcdonough@rochester.edu, hlehnert@ling.rochester.edu, nbardhan@bcs.rochester.edu,

Abstract

The goal of the presented study was to investigate the use of coarticulatory vowel nasalization in lexical access by native speakers of American English. In particular, we compare the use of coarticulatory place of articulation cues to that of coarticulatory vowel nasalization. Previous research on lexical access has shown that listeners use cues to the place of articulation of a postvocalic stop in the preceding vowel. However, vowel nasalization as cue to an upcoming nasal consonant has been argued to be a more complex phenomenon. In order to establish whether coarticulatory vowel nasalization aides in the process of lexical access in the same way as place of articulation cues do, we conducted two perception experiments: an off-line 2AFC discrimination task and an on-line eyetracking study using the visual world paradigm. The results of our study suggest that listeners are indeed able to use vowel nasalization in similar ways to place of articulation information, and that both types of cues aide in lexical access.

1. Introduction

The fact that listeners consider lexical candidates equally if they are identical in the make-up of their initial sound sequences up to the point of disambiguation has lead to models of spoken word recognition that posit that lexical access proceeds incrementally (cf. [13], [15] among others). Whether these increments are of phonemic or sub-phonemic size has been a focus of recent research (cf. [8], [14]). While some recent studies suggest that listeners use sub-phonemic allophonic information to the place of articulation of an upcoming consonant (cf. [7], [8]), previous studies on the perception of coarticulatory vowel nasalization ([10], [2], and [3]) and its use in lexical access (cf. [11]) seem to suggest that vowel nasalization is not used by listeners in the same way. Some studies have shown that, out of context, listeners can discriminate between an oral and a nasalized vowel (ṽ ~ v) [10]. However, when given a lexical item or a nasalized vowel followed by a nasal consonant, they are much more likely to classify a nasalized and an oral vowel as 'the same', indicating that listeners at least partially ignore the nasality in the vowel and rely on their phonemic knowledge; i.e. they compensate for coarticulation [10] [2][3]. Furthermore, Lahiri & Marslen-Wilson [11] report that listeners did not use coarticulatory vowel nasalization as cue to an upcoming nasal consonant in lexical access during a gating task.

These differences in the use of coarticulatory information to the place of articulation (PoA) and to the nasality (Nas) of an upcoming consonant reported in the literature and summarized above may have at least two sources: 1. Listeners exploit these two coarticulatory cues (PoA and Nas) indeed differently or 2. The gating task used by Lahiri & Marslen-Wilson [11] was not sensitive enough to show the use of coarticulatory vowel nasalization as cue to an upcoming nasal.

It is very well possible that listeners exploit PoA and Nas cues differently, since the two cues have different phonological as well as phonetic properties: A vowel carrying coarticulatory information about the PoA of an upcoming stop is considered the same phonemic vowel independent of its context (i.e. The [ɪ] in *kit* is the same sound as a [ɪ] in *kick*.). A nasalized vowel, however, can be phonemic in some languages, which, in turn, may lead to less pronounced coarticulatory nasalization in vowels preceding a nasal consonant [18]. Phonetically, vowel nasality is considerably more

complex in the acoustic domain than cues to PoA of consonants. Cues to the place of articulation of a stop consonant are acoustically encoded by well-defined differences in the formant transitions [12]. However, there is no simple measurable acoustic correlate of nasality in vowels. Vowel nasality is characteristically affected by vowel height, speaker [9][4][5], and, arguably, there exist timing differences between languages in the production of nasality [18][6] that may arise due to phonemic or allophonic status of vowel nasality. These properties make nasality in vowels a very different cue from PoA cues.

In the current study, we seek to determine whether listeners differ in how they exploit coarticulatory cues to an upcoming nasal consonant as opposed to coarticulatory cues to the place of articulation of an upcoming stop in lexical access. For this purpose, we conducted two perception experiments with native speakers of American English. A 2AFC discrimination task to test listeners sensitivity to the two types of coarticulatory information off-line, and a 4AFC eyetracking experiment using the visual world paradigm [1] to test the online processing of coarticulatory cues to PoA and nasal consonants.

2. Off-line discrimination of vowel nasality and place of articulation cues

We conducted a discrimination study to examine how listeners use the coarticulatory information present in the vowel of a CVC word to predict the upcoming final consonant. In order to do so, a stimulus list was constructed containing two experimental conditions: the PoA condition and the Oral/Nasal (O/N) condition. In both, the stimuli consisted of word pairs that differed in the final consonant of the word. The PoA stimuli consisted of pairs differing in the place of articulation of the final consonant (e.g. *tack* vs. *tap*) and the majority of these stimuli consisted of the same word pairs used in the study by Dahan et al. [7]. The O/N stimuli consisted of word pairs differing only in the nasality of the final consonant (e.g. *bong* vs. *bog*). All words were picturable, and matched for frequency using the American National Corpus. The final consonant in all stimulus words was excised using Praat [17].

Overall, there were 18 word pairs (36 items total). 11 word pairs contrasted in the nasality of the final consonant and 7 word pairs contrasted in the place of articulation of the final consonant.

Listeners were presented with each of the 36 items 6 times, amounting to 216 trials per subject (66 nasal trials, 66 oral trials, 30 velar PoA trials, 36 alveolar PoA trials, 18 bilabial PoA trials). The trials were presented in randomized order. The differences in the numbers between PoA and O/N contrast items was unavoidable, due to the need to present picturable words.

During the experiment, participants (5 male and 6 female native speakers of American English with no known hearing impairment) were seated in front of a computer screen and listened to the stimuli over headphones while two pictures representing the contrast pair (i.e. a picture of a *bong* and a picture of a *bog* while hearing 'bo') occurred on the screen. The participants were asked to click on the picture representing the word from which they thought the sounds had come.

The response on each trial was recorded, and the error rate for each condition was calculated. Listeners were largely able to correctly identify the correct words based on the coarticulatory information present in the vowel, without the final consonant present. The overall error rate across all trials was 9.5%. However, listeners' accuracy varied depending on the experimental condition: 7.4 % out of the 9.5% were errors that occurred on PoA trials, while the remaining 2.1% of erroneous trials occurred on O/N trials ($p < 0.0001$).
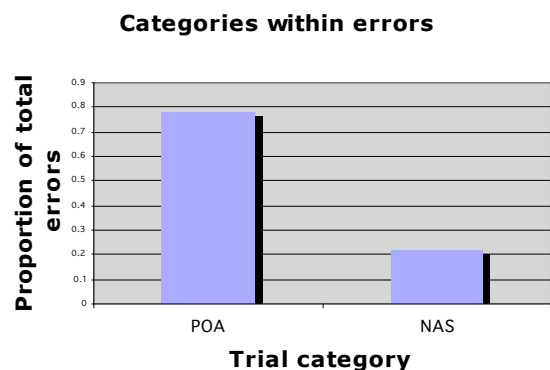
Categories within errors

Figure 1: *Percentage of errors made in discriminating the place of articulation of the final consonant (POA) and the nasality of the final consonant (NAS) relative to all errors made.*

As shown in Figure 1, listeners made better use of the nasality cue in the vowel than of the place of articulation cues, since the error rate for the nasal/oral contrast pairs was significantly lower than that for the PoA pairs, replicating the findings

by Kawasaki [10]. There was, however, no bias toward either nasal or oral vowels: oral vowels were mistaken for nasalized vowels in 0.09% of the trials, and nasalized vowels were mistaken for oral vowels in 0.12% of trials.

3. On-line discrimination of vowel nasality and place of articulation cues

Based on the results from the off-line discrimination task, we would expect that listeners will use vowel nasality as cue for an upcoming nasal consonant at least as much - if not more so - during lexical access. However, previous research using a gating task seems to suggest that nasalization of vowels is not exploited in such a way [11]. In order to test whether coarticulatory nasalization is used similarly or different from place of articulation cues, we conducted a 4AFC eyetracking experiment.

The 11 participants in this study were again native speakers of American English with no known hearing impairment. Participants in the eyetracking study had not participated in the discrimination experiment.

The stimuli for this experiment consisted of the same lexical material as those used in the discrimination experiment. However, the recordings of the stimuli for this experiment consisted of the complete words (including final consonants). There were again two lists; one containing pairs that contrasted in the place of articulation of the final stop, and one containing word pairs contrasting in the nasality of the final sound.

Wearing a head-mounted eye-tracker, subjects saw four pictures on a computer screen while hearing the target word over headphones. The display on the screen always contained the target word (e.g. *bong*), the phonological competitor with different final consonant (e.g. *bog*) and two unrelated items that differed in the initial consonant-vowel sequence from the target word (e.g. *gut* and *neck*). As in the discrimination study, each word was repeated 6 times in randomized order. Participants were instructed to click on the named object. The eye movements of the participants were monitored as they listened to the word and clicked on the picture representing the word.

Our results suggest that listeners may indeed use nasalization in the vowel preceding a nasal consonant. Figure 2 shows the proportion of looks

to the target relative to the offset of the vowel in the CVC word. Since it takes on average 200 ms to initiate an eye movement to react to a stimulus, 200 ms after the offset of the vowel (0 ms on the time axis) marks the point relative to which the proportion of looks for each condition has to be evaluated. If the proportion of looks starts to increase before the 200 ms mark, listeners have most likely identified the target based on coarticulatory information in the vowel. If listeners' looks increases after the 200 ms mark, they have most likely not identified the word before hearing the final consonant.
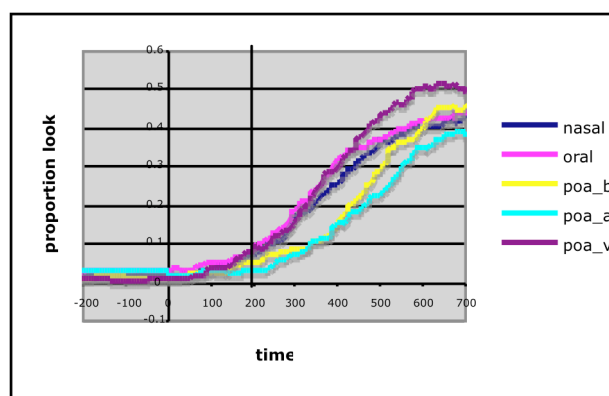

Figure 2: *Proportion of looks to the target word when the final consonant was a nasal, an oral, a bilabial, an alveolar, and a velar consonant.*

As can be seen in Figure 2, the proportion of looks to the target starts to increase already before the 200 ms mark for all but the words that have a final consonant at the alveolar place of articulation. Considering the 200 ms window before the off-set of the vowels, the proportion of looks to the nasal target in the last 100 ms before the off-set of the vowel differed significantly ($p < 0.001$) from the proportions of looks to the target in the first 100 ms.

This seems to suggest that the nasalization of a vowel aides as much in lexical access as the cues in a vowel that provide information about the place of articulation of an upcoming consonant. The only condition in which listeners did not seem to exploit coarticulatory information for lexical access was in the case of the alveolar stops. This may, however, very well be an artifact of our stimuli since most of the phonemically alveolar final stops were phonetically realized as glottal stops in our stimuli, hence leaving little coarticulatory information on the preceding vowel.

Overall, the results from this eyetracking study strongly suggest that listeners use acoustic cues indeed as they become available in the speech signal (cf. [14].

4. Discussion and conclusion

The current study investigated the use of coarticulatory vowel nasalization in lexical access. Both off-line and on-line experiments suggest that listeners are sensitive to nasalization in vowels preceding an upcoming nasal consonant and that they use this information during lexical access. While the claim that coarticulatory vowel nasalization facilitates lexical access cannot be deduced from our findings, the fact that listeners were affected in similar ways by coarticulatory cues to the place of articulation of an upcoming consonant and by coarticulatory cues to the nasality of an upcoming consonant, however, makes it likely that this is the case. Especially in light of the results reported in [7] who found facilitation in lexical access due to coarticulatory place of articulation cues.

Overall, the results in this study contradict claims that listeners ignore non-contrastive vowel nasality [10], [2], and [3]. However, to what extent the phonological status of vowel nasalization affects lexical access remains a question for future research, and can only be answered by cross-linguistic studies.

References

[1] Allopenna, P.D., J. Magnuson & M.K. Tanenhaus 1998. Tracking the time course of spoken word recognition using eye movements. *Journal of Memory and Language* 38, 419-439.

[2] Beddor, P. & R. Krakow. 1999. Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *JASA*. 106, 2868-2887.

[3] Beddor, P., R.Krakow & S.Lindemann 2001. Patterns of perceptual compensation. *The Role of Speech Perception in Phonology*. ed. Hume & Johnson, Academic Press.

[4] Berger, M. 2008. Measurement of vowel nasalization by multi-dimensional acoustic analysis. *WPLS: UR*. Winter 2008, 4.1.

[5] Chen, M.Y. 1997. Acoustic correlates of English and French nasalized vowels. *JASA* 102, 2360–2370.

[6] Cohn, A. 1990. *Phonetic and phonological rules of nasalization*, Ph.D. thesis, UCLA.

[7] Dahan, D., J.S. Magnuson, M.K. Tanenhaus & E.M. Hogan. 2001. Subcategorical mismatches and the time course of lexical access. *Language and Cognitive Processes* 16, 507-534.

[8] Gow, D. W. Jr. 2003. Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, 65, 575-590.

[9] Hawkins, S. & Stevens, K.N. 1985. Acoustic and perceptual correlates of the nonnasal-nasal distinction for vowels, *JASA* 77, 1560–1575.

[10] Kawasaki, H. 1986. Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. In: J.J. Ohala & J. J. Jaeger (Eds.) *Experimental Phonology*, pp. 81-103. Orlando: academic Press.

[11] Lahiri, A. & Marlsen-Wilson, W. 1992. Lexical processing and phonological representation, in *Papers in Laboratory Phonology II*: ed Docherty & Ladd, 229–260

[12] Liberman, A. M., Delattre, P. C., & Cooper, F.S. 1954. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. Psychological Monographs, 68, 1-13.

[13] McClelland, J. L. & Elman, J. L. 1986. The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.

[14] McMurray, B. Clayards, M. A., Tanenhaus, M. K. & Aslin, R. N. 2008. Tracking the time course of phonetic cue integration during spoken word recognition. Psychonomic Bulletin & Review, 15, 1064-1071.

[15] Norris, D. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.

[16] Ohala, J.J. & Ohala, M. 1995. Speech perception and lexical representation: The role of vowel nasalization in Hindi and English in *Papers in Laboratory Phonology IV*, ed Connell & Arvaniti, 41–60.

[17] Praat. 2008 Boersma & Weenik.

[18] Solé, M. 1992. Phonetic and phonological processes: case of nasalization. *Language and Speech* 35, 29–43.