Comparing Tongue Positions of Vowels in Oral and Nasal Contexts

Takayuki Arai

Department of Electrical and Electronics Engineering
Sophia University, Tokyo, Japan
arai@sophia.ac.jp

Abstract

We studied tongue positions of vowels in oral and nasal contexts. In the previous study [Arai, J. Acoust. Soc. Am., 115, p.2541 (2004)], formant frequencies were measured and bidirectional formant shifts in F1 frequency were observed: increasing F1 for high vowels and decreasing F1 for low vowels. Then, we tried to answer to the next question, that is, whether or not speakers and/or listeners compensate for the formant shifts. The perceptual experiment by Arai (2004) showed that compensation occurs when an isolated vowel has nasalization and is accompanied by formant transitions. This result agreed with the findings of Krakow et al. [J. Acoust. Soc. Am. 83, 1146-1158 (1988)]. The goal of this study is to examine the compensation effect for the formant shifts in production. In the EMMA experiment, the measurement of the positions of the articulators showed almost no compensation except for the lowest vowel /ɑ/.

1. Introduction

The additional poles and zeros due to nasal coupling cause modifications in the vowel spectrum, such as reduction in amplitude of the first formant (F1), broadening the bandwidth of F1, shifting F1 upwards in frequency, and a relative strengthening of the spectrum in the vicinity around 250 Hz [1-7]. The higher frequencies may also be modified by nasal coupling. The main effect of nasalization, however, is the perturbation of the low-frequency spectrum by replacing the first formant with a shifted F1 (F1'), a nasal formant (Fn), and a nasal zero (Fz) [3,5,8]. As the cross-sectional area of the velopharyngeal opening is gradually increased, the spacing between the pole and zero introduced in the vicinity of the first formant increases, F1 frequency shifts, and F1 bandwidth increases.

Calculations of the acoustic consequences of nasal coupling predict distinctively different modifications depending on vowel identity [4,5]. For high vowels, the theory predicts F1 shifts upwards in frequency, and a nasal formant appears in the spectral valley between F1 and F2. For low vowels, F1 also shifts upwards in frequency. At the same time, F1, however, comes close to the zero, because the first pole-zero pair is lower in frequency than F1 in the corresponding oral vowel, and

as a result, F1 is weakened and seemingly split into two peaks [8]. At low degrees of coupling, the nasal pole (Fn) is almost canceled by the nasal zero (Fz), and in this case, the prominence of Fn is small. At higher degrees of coupling, on the other hand, Fn increases in prominence [5], and Fn could be identified as a formant [7].

Thus, nasal coupling can shift F1 frequency and this may affect perceived vowel height. Due to the upwards shift in F1 frequency, nasalization might be expected to lower perceived vowel height [9]. In addition, due to the prominence of Fn, nasalization might be expected to raise perceived vowel height for low vowels. This seems to be a possible explanation for the bidirectional shifts in perceived nasal vowel height [10,11] that were observed in the perceptual experiments of previous studies [12,13].

In the previous study [14], we confirmed through acoustic analyses the manner in which formant frequencies shift due to nasalization, especially the bidirectional shifts in F1 frequency. Even though the acoustic theory shows the bidirectional movements of the F1 frequency, there were not many references in the literature which showed clear acoustic evidence for the bidirectionality. Therefore, the acoustic evidence observed by Arai [14] is valuable in a sense that the F1 frequency tends to shift in a more central direction when nasalized based on several measurements of formant frequencies for various vowels.

In English and many other languages, vowels should be perceived as the same phoneme regardless of nasalization. In other words, a speaker and/or a listener might tend to compensate for such formant shifts. Krakow et al. [10] examined the hypothesis that perceived nasal vowel height is not entirely determined by the spectral shape of the nasal vowel, but rather the context in which the nasal vowel occurs can affect the way in which the nasalization of that vowel is perceived. This hypothesis was tested by comparing listeners' perception of nasal vowels in the presence and absence of an adjacent nasal consonant, and Krakow et al. [10] showed that nasal coupling does not necessarily lead to listener misperceptions of vowel quality when the vowel's nasality is coarticulatory in nature. Arai [14] has also discussed the compensation effect on the perception side.

The goal of this study is, therefore, to investigate whether or not compensation exists in production. We

measured the positions of the articulators, especially tongue height, and compared them in oral and nasal contexts using the Electromagnetic midsagittal articulometer (EMMA) system [15].

2. EMMA measurement

In the previous study [14], we observed bidirectional shifts in F1 frequency of nasalized vowels. It raised the following two questions: 1) does a speaker tend to compensate for these formant shifts? 2) does this F1 shift cause misperception of the vowel quality by a listener? In Arai [14], we have discussed the compensation effect on the perception side by a perceptual experiment. In this study, we will try to answer these questions by a measurement of the positions of the articulators.

To test the compensation on the production side, we measured the difference in tongue height of a vowel in oral and nasal contexts using the Electromagnetic midsagittal articulometer (EMMA) system [15].

The speech samples were mono-syllabic nonsense words that were a subset of the samples used in Arai [14], that is, /bVb/ and /bVm/ with the vowel V of /i/, /ɪ/, /ɛ/, /ʌ/, /æ/, and /ɑ/. A male native speaker of American English uttered the target words embedded in the carrier phrase "Say _____, again." All 12 combinations were repeated five times in a random order (60 utterances, in total). The movements of the articulators, such as the upper lip (UL), the lower lip (LL), the tongue tip (TT), the tongue body (TB), the tongue dorsum (TD), and the lower incisor (LI), were recorded by the EMMA system simultaneously.

3. Results and discussions

Figure 1 shows the comparisons of the tongue position (especially, TB and TD transducers) between each minimal pair during the vowels in oral and nasal contexts (the blue/dark and red/light dots represent the tongue positions in oral and nasal contexts, respectively) . In this figure, the x-coordinate runs from back to front and the y-coordinate runs from bottom to top on the midsagittal plane, and the origin is located at the upper incisor.

For /i/, /æ/ and /ʌ/, the two distributions in oral and nasal contexts are fairly overlapped in both TB and TD cases. For /ɪ/ and /ɛ/, the distributions of the y-coordinate of the TB and TD transducers are overlapped, whereas the x-coordinates are somewhat separated (i.e., tongue positions are more advanced in nasal context). Because we are interested in tongue height, however, we can conclude that there is no major difference in y-coordinate between the two contexts. For /ɑ/, the distributions of the x-coordinate of the TB and TD transducers are fairly overlapped. There is, however, a difference between the two distributions in the y-coordinate. The difference of the x-coordinates of the TD transducer is approximately 2-3 mm.

As a result, the distributions between oral and nasal contexts were overlapped each other by a fair amount, and the difference was not significant except the lowest vowel /ɑ/. In case of the vowel /ɑ/, the result shows that the tongue dorsum is lower in nasal context (i.e., /bɑm/) than in oral context (i.e., /bɑb/). In other words, this speaker might have tried to make a more extreme /ɑ/ to compensate for the F1 shift due to nasalization.

If a speaker knows that F1 frequency shifts when nasalized, there might be a chance that the speaker tries to compensate for the shift during the production of speech. Furthermore, we might be able to observe the compensatory correlation, or "motor equivalence," between velar and tongue heights. However, it could be only true in languages that distinguish nasal and oral vowels phonemically (such as French, Portuguese, and Hindi). American English, on the other hand, does not make such distinction, and there might be no compensation on the production side.

4. Conclusions

In this study, tongue positions of vowels in oral and nasal contexts were investigated. From the acoustic measurements in Arai [14], bidirectional formant shifts in terms of F1 frequency were observed as predicted by the acoustic theory. From measurements using the EMMA system, speakers of American English tend not to compensate for such an F1 frequency shift by adjusting the tongue height except for the lowest vowel /ɑ/. This might be an indication that a speaker is not compensating his/her tongue height except the extreme one.
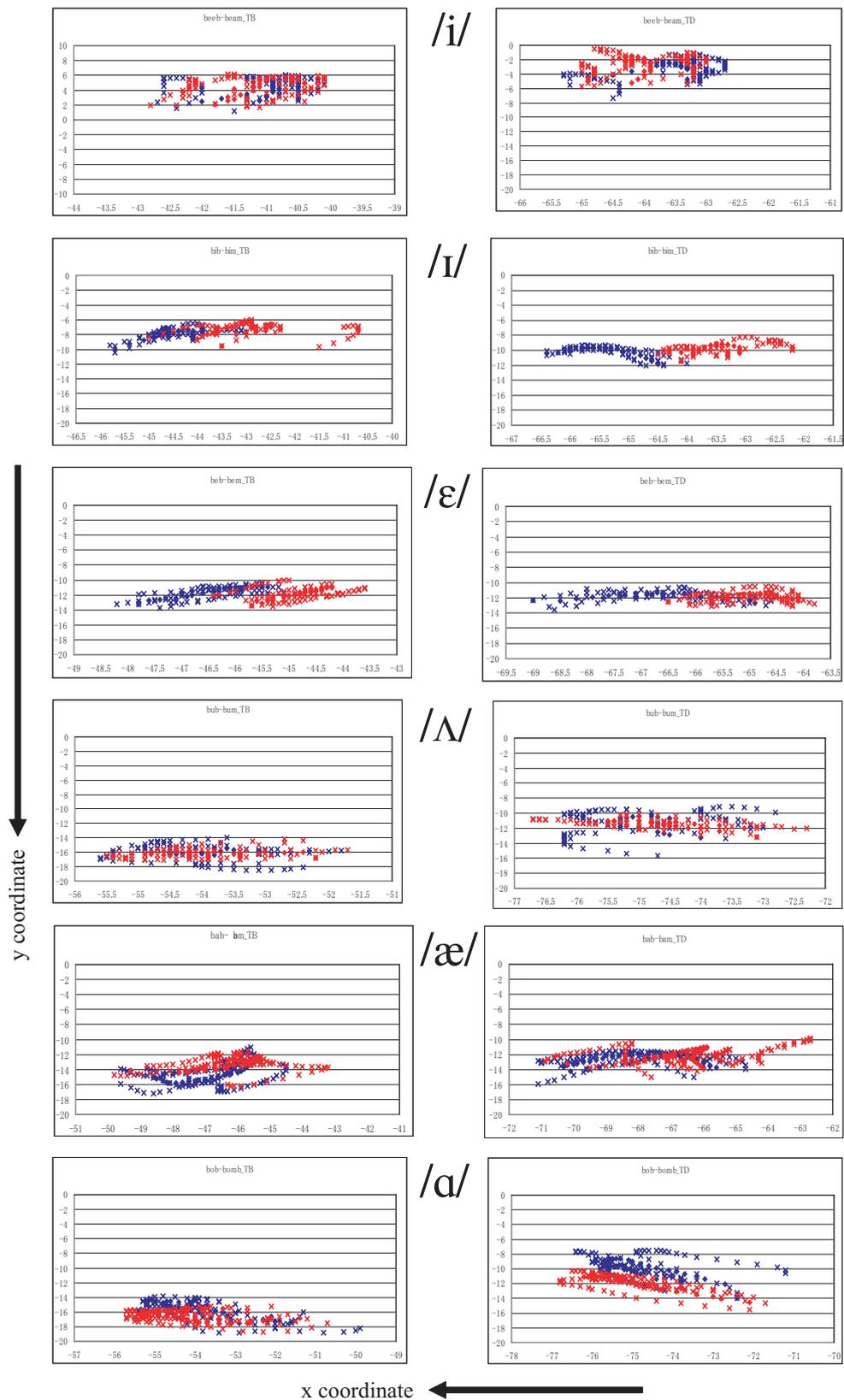
5. Acknowledgments

Fig. 1. The x and y coordinates of the TB and TD transducers between each minimal pair during the vowels in oral and nasal contexts (left column: TB; right column: TD). The blue (dark) and red (light) dots represent the tongue positions in oral and nasal contexts, respectively.

6. References

[1] House, A.S., and Stevens, K. N. (1956). "Analog studies of the nasalization of vowels," J. Speech Hear. Disord. 21, 218-232.

[2] Hattori, S., Yamamoto, K., and Fujimura, O. (1958). "Nasalization of vowels in relation to nasals," J. Acoust. Soc. Am. 30, 267-274.

[3] Fant, G. (1960). Acoustic Theory of Speech Production (Mouton, The Hague).

[4] Fujimura, O. (1960). "Spectra of nasalized vowels," Res. Lab. Electron. Q. Prog. Rep. No. 58, MIT, 214-218 (July 15, 1960).

[5] Fujimura, O., and Lindqvist, J. (1971). "Sweep-tone measurements of vocal-tract characteristics," J. Acoust. Soc. Am. 49, 541-558.

[6] Hawkins, S. W., and Stevens, K. N. (1985). "Acoustic and perceptual correlates of the non-nasal--nasal distinction for vowels," J. Acoust. Soc. Am. 77, 1560-1575.

[7] Maeda, S. (1993). "Acoustics of vowel nasalization and articulatory shifts in french nasal vowels," in Phonetics and Phonology: Nasals, Nasalization, and the Velum 5, edited by M. K. Huffman and R. A. Krakow (Academic Press, San Diego, CA), pp. 147-167.

[8] Stevens, K. N., Fant, G., and Hawkins, S. (1986). "Some acoustical and perceptual correlates of nasal vowels," in In Honor of Ilse Lehiste, edited by R. Channon and L. Shokey (Foris, Dordrecht, Holland), 241-254.

[9] Ohala, J. J. (1986). "Phonological evidence for top-down processing in speech perception," in Invariance and Variability of Speech Processes, edited by J. S. Perkell and D. H. Klatt (Erlbaum, Hillsdale, NJ), 386-401.

[10] Krakow, R. A., Beddor, P. S., Goldstein, L. M., and Fowler, C. A. (1988). "Coarticulatory influences on the perceived height of nasal vowels," J. Acoust. Soc. Am. 83, 1146-1158.

[11] Stevens, K. N. (1998). Acoustic Phonetics (MIT Press, Cambridge, MA).

[12] Wright, J. T. (1975). "Effects of vowel nasalization on the perception of vowel height," in Nasalfest: Papers from a Symposium on Nasals and Nasalization, edited by C. A. Ferguson, L. M. Hyman, and J. J. Ohala (Language Universals Project, Stanford University, Stanford, CA), 373-388.

[13] Wright, J. T. (1986). "The behavior of nasalized vowels in the perceptual vowel space," in Experimental Phonology, edited by J. J. Ohala and J. J. Jaeger (Academic, Orlando, FL), 45-67.

[14] Arai, T. (2004). "Formant shift in nasalization of vowels," J. Acoust. Soc. Am., 115, 2541.

[15] Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., and Jackson, M. (1992). "Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements," J. Acoust. Soc. Am. 92, 3078-3096.