9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Effect of Prosodic Context on Lexical Access: An Investigation of Korean Denasalization

Seulgi Shin[1], *Annie Tremblay*[1]

[1]University of Kansas, USA
seulgi.shin@ku.edu, atrembla@ku.edu

Abstract

This study investigates the effect of prosodic context on listeners' interpretation of prosodically driven variations in lexical access. It does so by examining Korean listeners' processing of words that begin with a denasalized nasal. In Korean, nasal-initial words have their initial consonant denasalized at the beginning of the Accentual Phrase (AP) [1-2]. Participants completed cross-modal priming tasks where they saw Korean target words that began with a nasal (e.g., *noru* 'roe deer') (Experiment 1) or a plosive (e.g., *toru* 'stealing a base') (Experiment 2). The experimental auditory primes rhymed with the target but began with a denasalized nasal (e.g., /noru/); the control auditory primes were phonologically and semantically unrelated to the target (e.g., /tʃotɛ/ 'invitation'). The primes were heard in AP-initial or AP-medial position.

In Experiment 1, the denasalized primes facilitated the recognition of nasal-initial target words (compared to the control primes) in AP-initial position, but not in AP-medial position. In Experiment 2, there was no effect of prosodic context or priming condition, indicating that listeners interpreted denasalized nasals differently from plosives regardless of the prosodic context. These results suggest that listeners take prosodic context into account and are sensitive to prosodically driven fine-grained phonetic details in lexical access.

Index Terms: lexical access, prosodic context, Korean Accentual Phrase, denasalization, cross-modal priming

1. Introduction

A variety of factors influence spoken word recognition [3-7]. Phonetic/phonological context is one such factor. Researchers have found that when segmental ambiguity arises from phonological assimilation, the neighboring phonetic/ phonological context helps listeners interpret the assimilated segment and recover its underlying representation in lexical access [8-12].

Prosodic context can be considered as one type of phonetic/phonological context that influences spoken word recognition, as prosodic boundaries help listeners resolve segmentation ambiguities in the signal [13-15]. For example, using a word-monitoring task, Christophe et al. [13] showed that Phonological Phrase boundaries helped French listeners disambiguate target words (e.g., *chat* 'cat') from competitor words (e.g., *chagrin* 'sorrow'). Similarly, using a visual-world eye-tracking task, Tremblay et al. [15] showed that prosodic cues such as the fundamental frequency (F0) rise that occurs phrase-finally in French helps French listeners locate word-final boundaries and ultimately recognize phrase-final words in continuous speech. The findings of these studies demonstrate that acoustic cues to prosodic boundaries influence the word activation process and enhance listeners' segmentation of the speech signal.

What is less clear from existing research on prosodic effects, however, is whether (and if so, how) prosodic context also modulates listeners' interpretation of words that are potentially ambiguous at the segmental level. Domain-initial denasalization in Korean allows us to answer this question. Korean nasals undergo nasality weakening in domain-initial prosodic positions such as the left edge of the AP [1-2]. This denasalization can potentially create segmental ambiguity in the speech signal. For example, some researchers have described Korean denasalized nasals as voiced plosives [16-17]. (It should be noted that Korean lenis plosives are phonetically voiced in AP-medial, intervocalic position [18].)

Few studies have investigated the effect of Korean denasalization on word perception and word recognition. Using an offline phoneme identification task, Kim [19] showed that Korean listeners perceived denasalized nasals as nasals when the consonant was spliced into the initial position of CVCV sequences but as plosives when the consonant was spliced into the intervocalic position of VCV sequences. Additionally, Korean listeners perceived intervocalic plosives as nasals when the consonant was spliced into the initial position of CVCV sequences but as plosives when the consonant was spliced into the intervocalic position of VCV sequences. These results suggest that Korean listeners use the prosodic context to interpret the phonetic variations of the nasal consonant. However, it is still unclear whether the prosodic context also influences the interpretation of the phonetic variations in (higher-level) lexical access. Thus, it remains to be seen whether the word recognition system interprets these prosodically driven phonetic variations as a function of the prosodic context in which they are heard.

The present study thus investigates how Korean listeners resolve potential lexical ambiguities driven by denasalization, and ultimately assesses the role of prosodic context in lexical access. We hypothesize that the potential segmental ambiguity caused by denasalization will be resolved with the help of the prosodic context. If so, Korean listeners should process denasalized nasals as nasals in licit prosodic contexts (i.e., in AP-initial position) and as (lenis) plosives in illicit prosodic contexts (i.e., in AP-medial position), consistent with Kim's results [19]. Alternatively, if denasalized nasals are sufficiently different from voiced plosives in their acoustic realization, listeners may not process denasalized nasals as plosives in either prosodic context; such results would differ from those of Kim [19].

These hypotheses were tested using two cross-modal priming tasks with lexical decision where auditory primes that

 10.21437/SpeechProsody.2018-83

contained denasalized nasals were presented in AP-initial or in AP-medial position and where visual targets were orthographic words that began with a nasal (Experiment 1) or a plosive (Experiment 2).

2. Methods

2.1. Participants

A total of 36 native Korean listeners participated in this study (mean age: 24.1, standard deviation: 3.99, 13 females). All the participants were tested at a university in Seoul, South Korea.

2.2. Materials

2.2.1. Stimuli

A total of 32 /n/-/t/-initial disyllabic minimal pairs of Korean words were selected for Experiments 1 and 2. The two words in the minimal pairs differed only in whether the onset of the first syllable was /n/ or /t/ (lenis plosive), and were not semantically related (e.g., /noru/ 'roe deer' vs. /toru/ 'steal in a baseball game'). Across pairs, /n/- and /t/-initial words were matched in token frequency [20].

Table 1 shows the manipulation of prime-target relatedness and prosodic position in the two experiments. In Experiment 1, the auditory prime word was a denasalized /n/-initial word (e.g., /noru/ 'roe deer') (experimental condition) or a word matched in length and frequency to the target but phonologically and semantically unrelated to it (e.g., /tʃʰotɛ/ 'invitation') (control condition). The prime words were presented in AP-initial or AP-medial position in carrier sentences (i.e., [onɯl pɛun pʰjohjəni # __(n)ɯn ɑnipnitɑ] or [onɯl pɛun pʰjohjən # i__(n)ɯn ɑnipnitɑ], respectively). The visual target words were the same /n/-initial disyllabic words as the experimental prime words (e.g., *noru*) in both the experimental and control conditions. Experiment 2 differed from Experiment 1 only in its visual target words, which were /t/-initial rather than /n/-initial (e.g., *toru*).

Table 1: *Schematized example of stimuli for Experiments 1 and 2.*

		AP-initial [pʰjohjəni]AP#[Noru] "expression # a deer"		AP-medial [pʰjohjən]AP#[iNoru] "expression # this deer"	
Exps	Conditions	Prime	Target	Prime	Target
Exp1	experimental	[Noru]	noru	[Noru]	noru
	control	[tʃʰotɛ]	noru	[tʃʰotɛ]	noru
Exp2	experimental	[Noru]	toru	[Noru]	toru
	control	[tʃʰotɛ]	toru	[tʃʰotɛ]	toru

Note. N marks a denasalized nasal

The experiments also included 64 filler trials. Among the filler trials, 16 pairs of auditory prime and visual target words consisted of real words in Korean, of which 8 were minimal pairs similar to the experimental item pairs but with different word-initial consonants (e.g., /kɑlku/ 'longing' - /sɑlku/ 'apricot') and 8 were phonologically and semantically unrelated words similar to control item pairs (e.g., /jənpʰil/ 'pencil' – /sɑtʰɑŋ/ 'candy').

The auditory sentences were each recorded three times by a Seoul Korean female speaker.

2.2.2. Resynthesis

The stimuli were manipulated so that the experimental prime word would begin with a denasalized nasal in both the AP-initial and the AP-medial conditions. As illustrated in Figure 1, the NV portion of the prime words was spliced from one recording of an AP-initial token into another recording of an AP-initial token (identity-spliced) or into a recording of the corresponding AP-medial token (cross-spliced). The splicing was done in Praat [21]. The pitch and duration of the spliced vowel were manipulated to mimic those of the prosodic position into which the NV sequences were spliced so that the target word would have the pitch and duration of the prosodic context in which it was heard. As a result of the resynthesis, the denasalized nasals were licit in the AP-initial condition but illicit in the AP-medial condition.

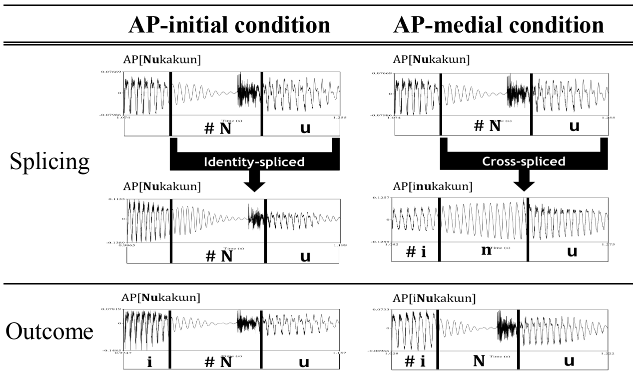

Figure 1: *Resynthesis of the stimuli for the AP-initial and AP-medial conditions.*

2.2.3. Acoustic analysis

Figure 2 compares the waveforms and spectrograms of a naturally produced (AP-medial) nasalized nasal, (AP-initial) denasalized nasal, AP-medial (voiced) plosive, and AP-initial (voiceless) plosive from the raw recordings. The plosives are lenis. (Although plosives were not heard as primes in the experiment, they were audio recorded for the purpose of comparing them to nasals in the acoustic analyses.)

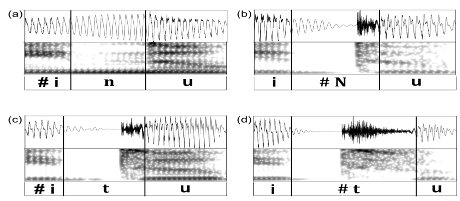

Figure 2: *Waveforms and spectrograms for a naturally produced (AP-medial) nasalized nasal (a), (AP-initial) denasalized nasal (b), AP-medial (voiced) plosive (c), and AP-initial (voiceless) plosive (d). # refers to the AP boundary. The words were /nukak/ or /tukak/.*

The acoustic properties of denasalized nasals were analyzed to compare them with those of both nasalized nasals and plosives. Table 2 presents a summary of the acoustic measurements of nasalized nasals, denasalized nasals, AP-medial plosives, and AP-initial plosives. Table 3 shows the results of paired-samples *t*-tests comparing denasalized and nasalized nasals, and comparing denasalized nasals and AP-medial and AP-initial plosives. All statistical analyses and comparisons were conducted on the consonant measurements relative to those of the following vowel (i.e., the values in parentheses in Table 2.).

Table 2: *Summary of mean acoustic measurements of (AP-initial) denasalized nasals, (AP-medial) nasalized nasals, AP-medial (voiced) plosives, and AP-initial (voiceless) plosives. The mean ratio of the measurements of the nasal relative to that of the following vowel are in parentheses.*

	Nasalized nasal	**Denasalized nasal**	**AP-medial plosive**	**AP-initial plosive**
RMS amplitude (dB)	57.51 (0.85)	44.77 (0.67)	24.59 (0.35)	26.44 (0.38)
Consonant duration (ms)	64.32 (1.33)	63.51 (0.66)	64.13 (0.72)	108.30 (0.99)
RMS burst energy (dB)	n/a	61.43 (0.93)	62.03 (0.87)	59.85 (0.85)
Burst duration (ms)	n/a	8.90 (0.14)	13.78 (0.21)	64.66 (0.60)

Table 3: *Summary of paired-samples t-tests comparing nasalized and denasalized nasals, and denasalized nasals and AP-medial and AP-initial plosives. DNs refers to denasalized nasals, NNs to nasalized nasals, MPls to AP-medial plosives, and IPls to AP-initial plosives.*

	Nasalized vs. denasalized nasals	**Denasalized nasals vs. AP-medial/initial plosives**
Relative RMS amplitude	DNs < NNs***	DNs > MPls *** DNs > IPls***
Relative consonant duration	DNs = NNs $^{n.s.}$	DNs = MPls $^{n.s.}$ DNs < IPls***
Relative RMS burst energy	n/a	DNs > MPls*** DNs > IPls***
Relative burst duration	n/a	DNs < MPls*** DNs < IPls***

Note. α=.05; ** = p <.01; *** = p <.001; $^{n.s.}$ = not significant

As illustrated in Figure 2, an important difference between denasalized and nasalized nasals was the presence or absence (respectively) of a burst. Another difference was in the Root Mean Square (RMS) amplitude of the nasal: As shown in Table 2, relative RMS amplitude was significantly lower for denasalized nasals than for nasalized nasals (t[93]=12.909, p<.001). These results indicate that Korean nasals are realized differently in different prosodic positions.

As can be seen in Table 2, denasalized nasals also had a relatively higher RMS amplitude than both AP-medial and AP-initial plosives (t[93]=20.926, p<.001; t[93]=16.271, p<.001, respectively) and a relatively shorter duration than AP-initial plosives (t[73]=8.25, p<.001). Furthermore, denasalized nasals had a relatively higher RMS burst energy than both AP-medial and AP-initial plosives (t[93]=8.154, p<.001; t[93]=4.538, p<.001, respectively), and a relatively shorter burst duration than both AP-medial and AP-initial plosives (t[84]=5.997, p<.001; t[71]=25.73, p<.001, respectively). The results of this acoustic analysis suggest that denasalized nasals are acoustically different from both AP-medial and AP-initial plosives.

2.3. Procedures

Each participant completed two cross-modal priming experiments with at least a two-day gap between the experiments. A participant never heard or saw the same prime or visual target twice across the two experiments. The lexical items in AP-initial and AP-medial conditions was counterbalanced across the experiments.

The experiments were built and run with PsyScope [22]. In each trial of the experiments, participants were asked to listen to sentences and decide as quickly as possible whether the visually presented word, which appeared 50 ms after the offset of the prime word, existed in Korean. Participants were given 1,500 ms to respond; if they did not respond within 1,500 ms, another trial started.

2.4. Data analysis

Linear mixed-effects models were conducted on the log-transformed response times (RTs) in Experiments 1 and 2. In the models, prosodic context (i.e., AP-initial vs. AP-medial positions), prime type (i.e., experimental vs. control conditions), and their interaction were included as fixed effects. The baseline was listeners' RTs in the AP-initial context of the experimental condition. Participants and items were included as crossed random effects. Model comparisons were first made between the most complex model (including prosodic context, prime type, and their interaction) and a model that differed from the most complex model by only one term. When the best model included a significant interaction between prosodic context and prime type (i.e., the most complex model), subsequent mixed-effects models tested the effect of prime type separately for the AP-initial and AP-medial items, using an adjusted alpha level of .025 (Bonferroni correction).

3. Results

3.1. Experiment 1

Figure 3 shows participants' log-transformed RTs for the experimental and control conditions in the AP-initial and AP-medial contexts. The linear mixed-effects model with the best fit on the log-transformed RTs, presented in Table 4, included a significant effect of prosodic context, a significant effect of prime type, and a significant interaction between prosodic context and prime type.

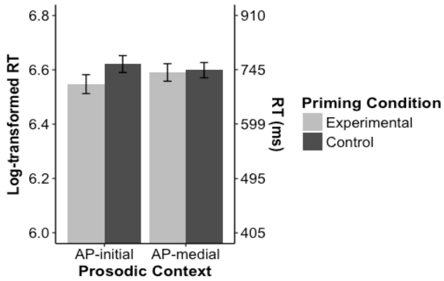

Figure 3: *Log-transformed RTs and corresponding RTs (ms) of the experimental and control conditions in AP-initial and AP-medial contexts for Experiment 1.*

The model in Table 4 yielded a significant effect of prosodic context, with the RTs in the experimental condition being slower in the AP-medial context than in the AP-initial context. The effect of prime type was also significant, with the RTs in the AP-initial context being significantly slower in the control condition than in the experimental condition.

Table 4: *Results of linear mixed-effects model with best fit on log-transformed reaction*

	Estimate	Std. Error	*t*	*p*
(Intercept)	6.551	0.027	239.71	<.001
Prosodic context (AP-medial)	0.051	0.018	2.77	<.001
Prime type (control)	0.091	0.023	3.95	<.001
Prosodic context x Prime type	-0.078	0.027	-2.92	<.001

Note. α=.05

Given the significant interaction between prosodic context and prime type, models were conducted on the RTs separately for the AP-initial and AP-medial contexts. As shown in Table 5, in the AP-initial context (where the prime was prosodically licit), the RTs were slower in the control condition than in the experimental condition. By contrast, in the AP-medial context (where the prime was prosodically illicit), the effect of prime type did not improve the model, with the RTs not differing as a function of prime type.

These results indicate that the prime word that began with a denasalized nasal facilitated the recognition of the corresponding visual target only when it occurred at the beginning of an AP, suggesting that denasalized nasals were processed as nasals only in a licit prosodic context.

Table 5: *Results of linear mixed-effects models on log-transformed reaction times for AP-initial condition*

		Estimate	Std. Error	*t*	*p*
AP-initial condition	(Intercept)	6.554	0.029	221.94	<.001
	Prime type (control)	0.094	0.026	3.65	<.001

Note. α=.025

3.2. Experiment 2

Figure 4 shows the log-transformed RTs for the experimental and control conditions in AP-initial and AP-medial contexts.

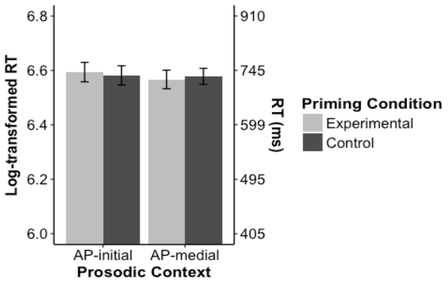

Figure 4: *Log-transformed RTs and corresponding RTs (ms) of the experimental and control conditions in AP-initial and AP-medial contexts for Experiment 2.*

Linear mixed-effects models conducted on the log-transformed RTs revealed that the best model was the model without any fixed effect. That is, neither prosodic context nor prime type improved the model, indicating that RTs did not differ across conditions. In other words, the denasalized nasal-initial prime did not facilitate the recognition of the plosive-initial disyllabic word that rhymed with it regardless of whether the prime word was presented in AP-initial or AP-medial context. This suggests that the denasalized nasals were not interpreted as plosives.

4. Discussion and Conclusion

The results of Experiment 1 showed that prosodic context affects listeners' interpretation of denasalized nasals in lexical access. While denasalized nasals are interpreted as nasals in a licit prosodic context such as in AP-initial position, facilitating the recognition of /n/-initial targets (compared to the control condition), they do not appear to be interpreted as nasals in an illicit prosodic context such as in AP-medial position, resulting in no facilitation of the recognition of /n/-initial targets. These results suggest that acoustic cues are interpreted based on the prosodic context in which they are heard. In other words, denasalized nasals are not interpreted as one of the acoustic variations of Korean nasals in lexical processing when the prosodic context does not license this variation.

What remains unclear from Experiment 1, however, is how native Korean listeners interpreted denasalized nasals in AP-medial position. We hypothesized that denasalized nasals would be interpreted as voiced plosives rather than as nasals. However, Experiment 2 revealed that listeners did not interpret denasalized nasals as plosives despite the fact that denasalized nasals showed some acoustic resemblance to plosives in AP-medial position, as shown in our acoustic analysis of denasalized nasals and plosives in different prosodic contexts. From the results of Experiment 2, we conclude that acoustic differences between denasalized nasals and plosives may have prevented denasalized nasals from being mapped onto plosives. These results would suggest that listeners are sensitive to the fine-grained acoustic differences between denasalized nasals and plosives, even in AP-medial context, but further research is necessary to explicitly examine listeners' use of these fine-grained acoustic cues in lexical access.

Our results differ from those of previous studies that used offline tasks such as a phoneme identification task, where denasalized nasals in illicit prosodic position were identified as plosives [19]. Even though such a task may lead listeners to focus on acoustic details, it also taps into processes that can be influenced by listeners' explicit and metalinguistic knowledge about a language. More importantly, such a task does not necessarily reflect the more implicit processes that take place in lexical access, including the use of fine-grained acoustic information in the activation of lexical candidates.

The present findings suggest that, in lexical processing, acoustic information interacts with prosodic context: Listeners do not interpret a potentially ambiguous segment in a word based only on the acoustic information they hear; they also take the prosodic context into account.

5. Acknowledgements

We would like to express our gratitude to Dr. Taehong Cho for making the Hanyang Phonetics and Psycholinguistics Lab available for the data collection, and to the members of LING 850 at KU for their valuable feedback.

6. References

[1] T. Cho and P. Keating, “Articulatory and acoustic studies on domain-initial strengthening in Korean,” *Journal of Phonetics,* vol. 29, no. 2, pp. 155-190, 2001.

[2] K. Yoshida, “Phonetic implementation of Korean “denasalization” and its variation related to prosody,” *IULC Working Papers Online. Vol. 8*. Bloomington: Department of Linguistics, Indiana University, 2008.

[3] P. D. Allopenna, J. S. Magnuson, and M. K. Tanenhaus, “Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models,” *Journal of Memory and Language*, vol. 38, no. 4, pp. 419-439, 1998.

[4] S. D. Goldinger, P. A. Luce, and D. B. Pisoni, “Priming lexical neighbors of spoken words: Effects of competition and inhibition,” *Journal of Memory and Language,* vol. 28, no. 5, pp. 501-518, 1989.

[5] J. L. McClelland and J. L. Elman, “The TRACE model of speech perception,” *Cognitive Psychology,* vol. 18, no. 1, pp. 1-81, 1986.

[6] J. M. McQueen, D. Norris, and A. Cutler, “Lexical influence in phonetic decision making: Evidence from subcategorical mismatches,” *Journal of Experimental Psychology: Human Perception and Performance,* vol. 25, no. 5, pp. 1363-1389, 1999.

[7] M. S. Vitevitch and P. A. Luce, “Probabilistic phonotactics and neighborhood activation in spoken word recognition,” *Journal of memory and language,* vol. 40, no. 3, pp. 374-408, 1999.

[8] M. G. Gaskell and W. D. Marslen-Wilson, “Phonological variation and inference in lexical access,” *Journal of Experimental Psychology: Human Perception and Performance,* vol. 22, no. 3-4, pp. 144-158, 1996.

[9] M. G. Gaskell and W. D. Marslen-Wilson, “Lexical ambiguity resolution and spoken word recognition: Bridging the gap,” *Journal of Memory & Language,* vol. 44, no. 3, pp. 325-349, 2001.

[10] D. W. Gow and B. McMurray, “Word recognition and phonology: The case of English coronal place assimilation,” in J. S. Cole & J. H. Hualdo (Eds.), *Papers in Laboratory Phonology 9*, New York: Mouton de Gruyter, 2000, pp. 173-200.

[11] D. W. Gow, “Does English coronal place assimilation create lexical ambiguity?” *Journal of Experimental Psychology: Human Perception & Performance,* vol. 28, no. 1, pp. 163-179, 2002.

[12] D. W. Gow, “Feature parsing: Feature cue mapping in spoken word recognition,” *Perception & Psycholinguistics,* vol. 65, no. 4, pp. 575-590, 2003.

[13] A. Christophe, S. Peperkamp, C. Pallier, E. Block, and J. Mehler, “Phonological phrase boundaries constrain lexical access: I. Adult data,” *Journal of Memory and Language,* vol. 51, no. 4, pp. 523–547, 2004.

[14] A. P. Salverda, D. Dahan, and J. M. McQueen, “The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension,” *Cognition,* vol. 90, no. 1, pp. 51-89, 2003.

[15] A. Tremblay, M. Broersma, C. E. Coughlin, and J. Choi, “Effects of native language on the use of fundamental frequency in non-native speech segmentation,” *Frontiers in Psychology, 7: Phonology in the Bilingual and Bidialectal Lexicon*, article 985, 2016.

[16] M. Chen and H. Clumeck, “Denasalization in Korean: A search for universals,” in C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, Stanford, CA: Stanford University Linguistics Dept, 1975, pp. 125-131.

[17] J. J. Ohala and M. Ohala, M, “The phonetics of nasal phonology: theorems and data,” in M. K. Huffman & R. A. Krakow (Eds.), *Phonetics and Phonology: Nasals, Nasalization, and the Velum*. London: Academic Press Limited, 1993, pp. 225-249.

[18] S-A. Jun, “The status of lenis stop voicing rule in Korean,” in Dikran Karaguezian (Eds.), *Theoretical issues in Korean linguistics*, CSLI, Stanford University Press, 1994, pp. 101-114.

[19] Y. Kim*, An Acoustic, Aerodynamic and Perceptual Investigation of Word-Initial Denasalization in Korean.* PhD Dissertation. University of College London, 2011.

[20] H. Kim, *Modern Korean Usage Frequency Report*. National Institute of Korean Language, 2005.

[21] P. Boersma and D. Weenink, “Praat: doing phonetics by computer” [Computer program]. Version 6.0.30, retrieved 22 July 2017 from http://www.praat.org, 2017.

[22] J. D. Cohen, B. MacWhinney, M. Flatt, and J. Provost, “PsyScope: A new graphic interactive environment for designing psychology experiments,” *Behavioral Research Methods, Instruments, and Computers,* vol. 25, no. 2, pp. 257-271, 1993.