

# The level of perception of illusory vowels

Jiwon Hwang and Ellen Broselow  
Stony Brook University

## 1. Background

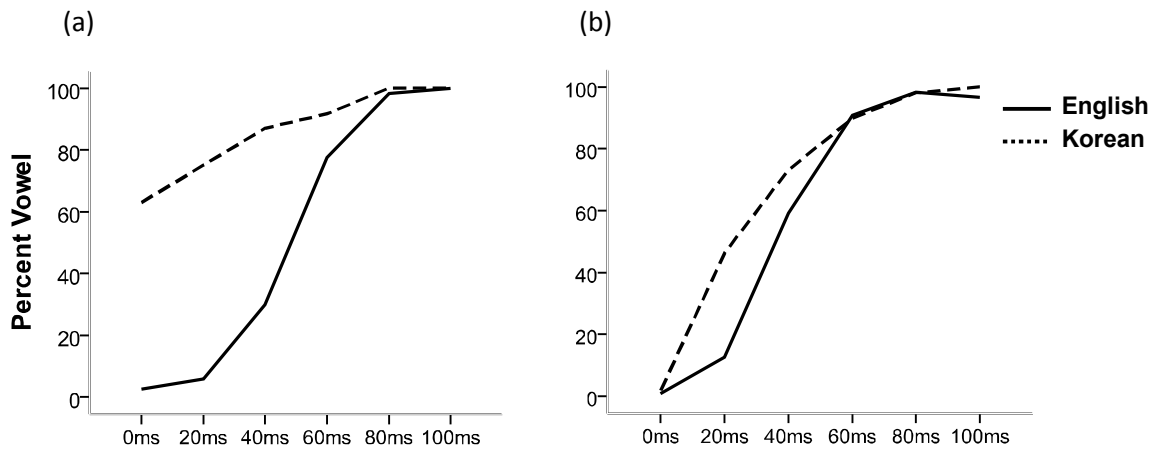
Korean has no sequences of stop followed by nasal, and a stop is nasalized before a nasal (kukmul → kuŋmul ‘soup’). However, L2 learners of English pronouncing English nonsense words containing stop-nasal sequences (e.g., *tegnal*, *sebnan*) sometimes nasalized the stops (sebnan → semnan), but more frequently inserted a vowel between the stop and the nasal (tegnal → teginal), considerably more often in voiced stop-nasal sequences than in voiceless stop-nasal sequences (Hwang, 2010). Hwang (2010) argued, based on identification and discrimination experiments, that the insertion of vowels after voiced consonants is an effect of misperception--since in Korean voiced obstruents occur only intervocalically (as an allophone of voiceless obstruents), Korean listeners tend to hear a vowel following a voiced obstruent.

This paper reports on both behavioral and Event-Related Potential (ERP) experiments designed to examine whether Korean listeners’ misperception of voiced stop-nasal sequences results from the failure to hear the acoustic differences between voiced stop-nasal sequences and voiced stop-V-nasal sequences, or simply from their failure to realize that these differences are relevant to linguistic categorization. Changes in an acoustic signal elicit a response in the electrical activity of the brain known as the MMN (mismatch negativity) response (Näätänen, Gaillard, & Mäntysalo, 1978), typically evoked by contrasts that are linguistically significant in the listener’s native language, even in the absence of conscious attention.

## 2. Categorization Experiment

11 Korean and 10 English listeners participated in a categorization task, in which they were asked to categorize in terms of absence or presence of a vowel in the ‘stop-nasal’ stimuli along a continuum that ranged from no vowel to a full vowel (*igna-igina* and *ikna-ikina*). The materials were produced by a Korean-English bilingual speaker who could pronounce the Korean epenthetic vowel [i] while maintaining the English phonology. The mean percentages of ‘vowel’ responses as a function of vowel length and language for each continuum are plotted in Figure 1. As illustrated in Figure 1, Korean listeners reported they heard a vowel about 60% of the time even when there was no vowel at all, but **only** when the stop was voiced. However, English-speaking listeners began to report hearing a vowel in [ignə] when the vowel was longer than 20ms. An ANOVA showed a significant three-way interaction between Language (Korean vs. English), Voicing (g vs. k) and Vowel length (0 through 100ms) ( $F(3.18, 54.07) = 10.14, p < .001$ ), which indicates that Korean participants heard a vowel even for the stimuli that had a shorter vowel or no vowel, especially when the continuum was [gn-gin]. The categorization experiment in Hwang (2010) was basically replicated, showing a robust language-specific effect of voicing in perception of stop-nasal sequences.

Figure 1. Mean percentages of ‘vowel’ response of English and Korean in the two continua: (a) in [gn-*g*in] and (b) in [kn-*k*in].



### 3. ERP Experiment

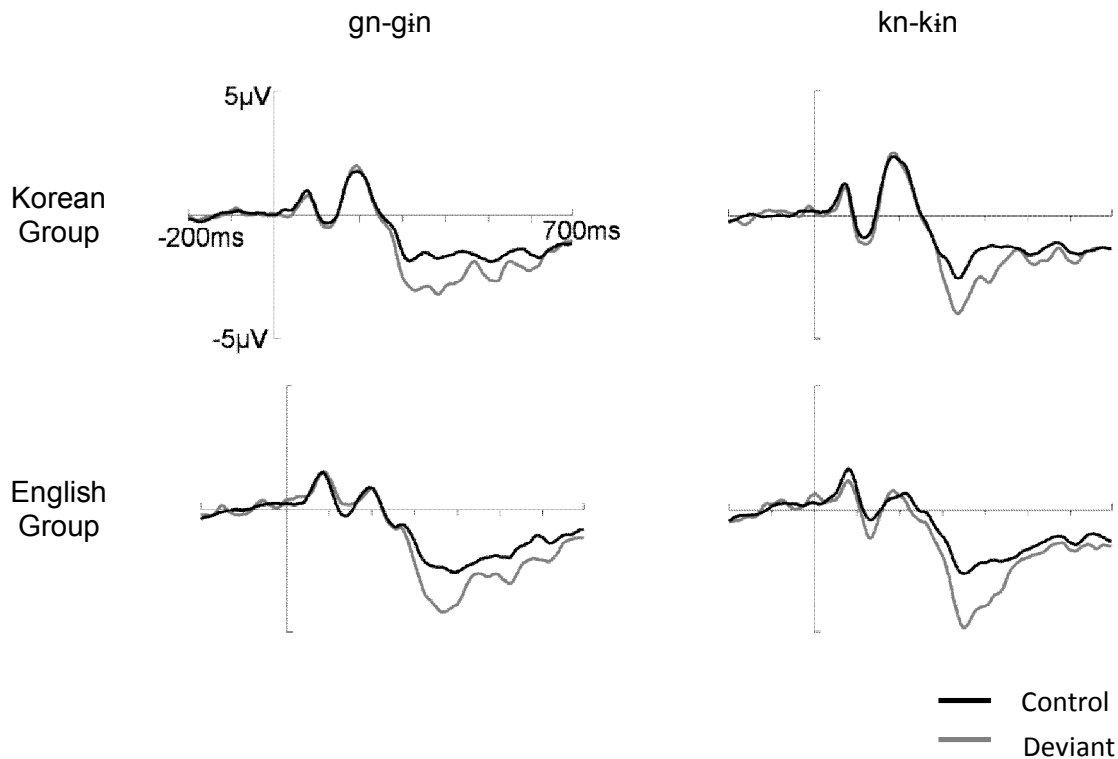
While the categorization task taps into phonological processing, an MMN displays brain responses to acoustic change, indicating pre-attentive discriminability of acoustic signal. The ERP experiment was conducted with the same subjects before the categorization task. The stimuli were the end point items on the continua used in the categorization task. During the electroencephalogram (EEG) recording, the subjects sat in a sound attenuating chamber and watched a silent movie while the auditory stimuli were presented binaurally via headphones (75dB). Continuous EEG was recorded using a 64 (Ag/AgCl) electrode cap and a Neuroscan Inc. data acquisition system with a fronto-central electrode as ground and electronically linked mastoid electrodes as reference. All recorded signals were bandpass filtered on-line between 0.1 Hz and 30 Hz, digitized at 500 Hz and amplified with a gain of 1000. Non-typical extreme movements were inspected visually and removed from continuous EEG data. Then, components indicating typical eye artifacts were identified and extracted via independent component analysis (ICA) algorithm (runica) of EEGLAB (Delorme & Makeig, 2004) for each participant. The corrected data were epoched from -200 ms to 700 ms and baseline-corrected. Epoched data were separately averaged for standard and deviant stimuli in each pair. Difference waveforms were obtained by subtracting the responses to the standard stimuli from the responses to those elicited by the identical stimuli when they were the deviant stimuli in other blocks. From those individual difference waveforms, mean MMN amplitudes were calculated using a 100 ms window centered at the negative peak latency, based on the grand average difference waveforms of each condition for each group. The presence of the MMN in each condition/group was tested by two-tailed one-group *t*-tests of the difference mean amplitude calculated at Fz. The mean amplitude of MMN at Fz are reported in Table 1. Surprisingly, both stimulus pairs (gn-*g*in or kn-*k*in) elicited a significant MMN in both subject groups regardless of the voicing of the stop.

Table 1. Mean amplitude ( $\mu\text{V}$ ) of MMN at Fz and p-value from a one-way *t*-test

	gn-gin (0ms vs. 100ms)	kn-kin (0ms vs. 100ms)
English group	-1.45 $p < .005$	-1.73 $p < .005$
Korean group	-1.43 $p < .001$	-1.15 $p < .005$

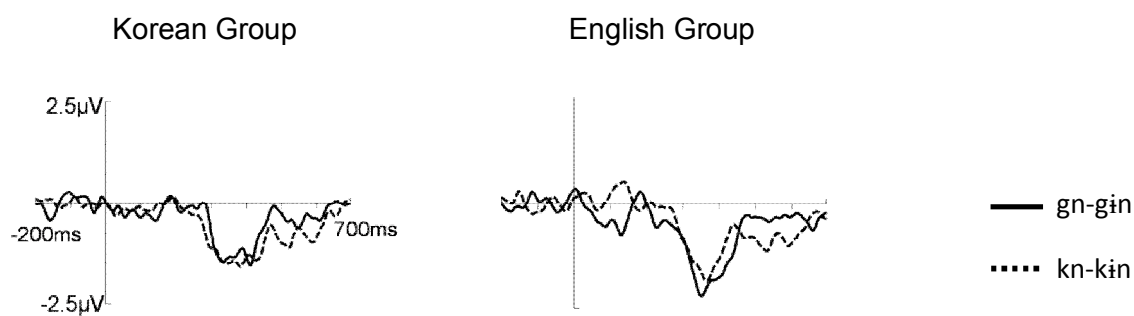
The grand-averaged ERPs are plotted in Figure 2. From the ANOVAs, a significant main effect of Anteriority was confirmed ( $F(1, 19)=51.06, p < .001$ ), consistent with the typical MMN topography of the observed MMNs in the literature, where frontally recorded responses are stronger than parietally recorded ones. However, that was the only significant effect found in the statistical analyses. There was no significant effect of Language (Korean vs. English) or significant interactions between Language and Voicing (g vs. k), indicating that the two groups' MMNs were very similar regardless of the stimulus type.

Figure 2. ERPs of controls and deviants in *gn-gin* and *kn-kin* stimulus pairs (0ms vs. 100ms) in Korean group and English group



Difference waveforms are presented in Figure 2. As shown in grand-averaged ERPs, the difference waveforms neither show significant differences between language groups or stimulus types, suggesting that voicing of the stop did not affect Korean listeners' pre-attentive discrimination of the pair.

Figure 3. Difference waveform (grand-average deviant-minus-control difference waves) elicited by a 100ms vowel duration difference in *ignə-iginə* and *iknə-ikinə* stimulus pairs in Korean group and English group



#### 4. Discussion

English native speakers displayed significant MMNs in all conditions, as expected. However, Korean speakers also displayed large MMNs even when the stimulus pair was *igna-igina*, suggesting that voicing of the stop did not affect their pre-attentive discrimination of the pair, unlike their categorization. This means that Korean and English participants' preattentive detection of acoustic differences between no vowel and vowel in stop-nasal sequences was fairly accurate for both language groups, regardless of the voicing of the stop in the stimulus pairs. What is striking then is that in their categorization of vowel vs. no vowel, language background played a significant role. The discrepancy between MMN and categorization responses shows that misperception of coda voiced stop as a sequence of the stop and a vowel is not rooted in audibility, but rather in the categorization process. We argue that a perception grammar mapping acoustics to phonological representations directs the auditory input [gn] to the phonological representation /kin/ (therefore [gin] in the production grammar) in Korean because voiced stops can occur only before a vowel in Korean. The results of the current experiments provide evidence that language-specific perception of the absence/presence of a vowel in stop-nasal sequences takes place not at the pre-attentive auditory level but rather at the level of phonological categorization.

#### References

- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134, 9-21.
- Hwang, J. (2010). *The nature of inserted vowels in L2 learners' production*. Paper presented at the the 84th meeting of the Linguistics Society of America.
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42, 313-329.