

JOURNAL ARTICLE

# Korean laryngeal contrast revisited: An electroglottographic study on denasalized and oral stops

Jiayin Gao<sup>1,2</sup>, Jihyeon Yun<sup>1,3</sup> and Takayuki Arai<sup>1</sup>

<sup>1</sup> Sophia University, Tokyo, JP

<sup>2</sup> University of Edinburgh, Edinburgh, UK

<sup>3</sup> Chungnam National University, Daejeon, KR

Corresponding author: Jiayin Gao ([jiayin.gao@ed.ac.uk](mailto:jiayin.gao@ed.ac.uk))

In several Korean dialects, domain-initial nasal onsets undergo denasalization as a recent sound change. Nasal stops may be realized as prevoiced or even devoiced stops. This makes it necessary to examine the interplay of phonetic properties of the denasalized and the three oral stop series as a whole, in synchrony and diachrony. What are their concomitant and conflicting properties? Our study provides a bigger picture of the laryngeal contrast in Seoul and Gyeonggi Korean by examining the acoustic distributions related to the laryngeal properties of the four stop series, using acoustic and electroglottographic data. VOT and  $f_0$  play important roles in the distinction of the four stop series, in line with previous studies. While the contribution of voice quality is relatively minor, we show that it plays an essential role of disambiguation when the VOT- $f_0$  space gets crowded: When lenis stops can be confused with other stops, there is an enhancement of breathy voice. Finally, we discuss stop variation according to prosodic contexts. We highlight the basis of both syntagmatic variation and paradigmatic contrast in their phonetic implementations. They illustrate a constant reorganization to reconcile contrast maintenance with constraints from articulatory and perceptual systems, as well as language-specific structures.

**Keywords:** Korean; electroglottograph; laryngeal contrast; denasalization; voice quality; breathy voice; VOT;  $f_0$

## 1. Introduction

### 1.1. Korean oral stops: The role of voice quality

The laryngeal contrast in Korean has been extensively studied in the past few decades (for a review, see Lee, Holliday, & Kong, 2020, and references therein). The three oral stop series—commonly labeled as ‘aspirated,’ ‘fortis,’ ‘lenis’—shed light on typological patterns on both synchronic and diachronic levels. First, these stops are all voiceless in phrase-initial position, which is typologically uncommon among the languages with a three-way laryngeal contrast. Second, in Seoul Korean, phrase-initially, the VOT (Voice Onset Time) difference between two of the three stop series has decreased in the course of the last century, while the following  $f_0$  (fundamental frequency) difference has increased (Bang, Sonderegger, Kang, Clayards, & Yoon, 2018; Kirby, 2013; Silva, 2006, to cite just a few). This phonologization of  $f_0$  has been argued to be incorporated into the Korean intonational phonology (Jun, 1993, 1998) (see discussions in Choi, Kim, & Cho, 2020).

To a lesser extent, the voice quality differences have also been found to contribute to the laryngeal contrast in Korean. Cho, Jun, and Ladefoged (2002) concluded that the three

stop series can be “differentiated from each other by the voice quality of the following vowel.” However, results of spectral tilt measures from previous literature are variable (for a review, see Lee & Jongman, 2012, **Table 7**). Restricted to Seoul Korean, it has been consistently found that lenis stops are followed by breathy vowels, and fortis stops by modal or laryngealized vowels (Cho et al., 2002; K.-H. Kang & Guion, 2008; Yu, 2018). On the other hand, aspirated stops do not show clear-cut patterns. H. Ahn (1999) (cited in Cho et al., 2002) and Lee and Jongman (2012) found that H1–H2 (i.e., amplitude difference between the first and second harmonics) was higher when following aspirated than lenis stops, suggesting breathier voice with the former. This is different from Cho et al.’s. (2002) findings, showing that H1–H2 values following aspirated stops were higher than fortis stops but lower than lenis ones, which according to the authors suggested modal voice with aspirated stops, breathy voice with lenis ones, and laryngealized voice with fortis ones. K.-H. Kang and Guion (2008) found yet another pattern: H1–H2 differed little between aspirated and lenis stops. Gender and age differences have also been reported: In M.-R. Kim (2014)’s study on 13 speakers in their 20–30s, H1–H2 was higher at vowel midpoint after lenis than aspirated stops only for females. Similarly, in Yu’s (2018) study on 11 older and 17 younger speakers (gender-balanced), H1–H2 and H1-A2 (i.e., amplitude difference between the first harmonic and the second formant) were higher when following lenis than aspirated stops only for young female speakers. Note, however, that this pattern (i.e., lenis breathier than aspirated) was found with older male speakers in Cho et al. (2002).

The aforementioned studies reported spectral tilt measures at fixed time points (generally at the vowel onset or midpoint) or averaged over the entire vowel. While  $f_0$  has been reported to be relatively stable over the entire vowel in Korean, we do not fully understand how voice quality evolves in the course of the vowel. This could potentially explain the contradictory results: It might be the case that the vowel following aspirated stops is breathier due to the aspiration, thus near the onset, but becomes less breathy during the latter part of the vowel. In the current study, we collected electroglottographic (EGG) data for the purpose of illustrating more accurately the time course of both  $f_0$  and voice quality thanks to the high temporal resolution of glottal variation reflected in the EGG signals (see the principle of the EGG method in sections 2.3 and 2.4). Finally, it should also be noted that a broad voice quality category, such as breathy voice, may correspond to different glottal mechanisms in terms of control of intrinsic laryngeal muscles as well as timing, speed, and magnitude of glottal opening/closing, airflow, and air pressure, as demonstrated in electromyographic, articulatory, and aerodynamic studies on Korean oral stops (e.g., Hirose, Lee, & Ushijima, 1974; H. Kim, Maeda, Honda, & Crevier-Buchman, 2018; Lee & Jongman, 2012).

Given the inconclusive results with respect to voice quality, the first question of this study is: How does voice quality on the following vowel differ between aspirated and lenis stops? Is the discrepancy in previous findings due to the time course of the voice quality over the entire vowel?

## **1.2. Denasalized and oral stops: A four-way contrast**

Independently, some attention has been drawn to another interesting sound pattern in several Korean varieties including Seoul and Gyeonggi dialects: denasalization. In these dialects, domain-initial positions have a stronger tendency for denasalization than domain-medial positions: For example, phrase-initial nasal onsets are often produced with weak nasality, and sometimes with no nasality, or even become devoiced (M.-J. Ahn, 2013; M. Chen & Clumeck, 1975; Y. S. Kim, 2011; Yoo, 2015; Yoshida, 2008), and they are likely to be perceived as oral voiced stops by nonnatives (M. Chen & Clumeck, 1975; Y. S.

Kim, 2011). In gestural terms, a real-time MRI study has shown that nasals in Korean are subject to greater variability of intergestural timing in syllable-onset than syllable-coda position (Oh, Byrd, Louis, & Narayanan, 2020). Finally, based on a recent apparent-time corpus study, denasalization is very likely to be an ongoing sound change in the process of stabilization (Yoo & Nolan, 2020).

Phrase-initial denasalization has rarely been analyzed together with the three oral stop series (but see a perception study from M.-J. Ahn, 2019, summarized in Section 1.3). In phrase-initial position, the weakening or absence of nasality makes the prototypical form of a nasal stop similar—although not identical to—a prevoiced oral stop (Y. S. Kim, 2011). As a consequence, together with the three voiceless oral stops in this position, speakers of these Korean varieties are most often producing four *oral* stop series on the surface level. The realizations of the nasal and oral stops are, however, conditioned by prosodic positions (Cho & Keating, 2001) and will be discussed in Section 4.3. The current study is mainly restricted to the phrase-initial position. The four-way contrast makes it necessary to examine the interplay of multiple phonetic properties in the production of the four stop series *as a whole*. How are the four stop series distinguished one from another? What are concomitant and conflicting properties of each stop series?

First, for the three voiceless oral stops, the positive range of VOT is divided into two spaces, with fortis (short-lag) on one side, lenis and aspirated (long-lag) on the other side. Now, the inclusion of nasal stops means that the negative VOT range is occupied as well. Second,  $f_0$  is lower at the vowel midpoint after lenis and nasal stops than aspirated and fortis stops (Y. Kang, 2014). In Korean intonation, aspirated and fortis stops and affricates trigger a high boundary tone while all the other onsets (including zero onsets, i.e., vowel-only syllables, and sonorant onsets) trigger a low boundary tone (Jun, 1993, 1998). Putting the four stop series altogether, the  $f_0$  space is divided into two, with fortis and aspirated in the high space, and lenis and nasal in the low space. Third, as summarized in Section 1.1, the vowel is breathier following lenis stops than fortis stops, while the voice quality pattern following aspirated stops is inconclusive. The inclusion of nasal stops would potentially add another complexity with its use of voice quality. Nasality and breathy voice are often tightly linked both synchronically and diachronically (Matisoff, 1975), partly motivated by their similar acoustic characteristics, such as the presence of anti-resonances (Ohala, 1975). Garellek, Ritchart, and Kuang (2016) reviewed the interactions between nasality and breathy voice, attributable to the misperception between the two properties (Arai, 2006; Ohala, 1975; Ohala & Busà, 1995) and/or mutual articulatory enhancement (e.g., Stevens & Keyser, 1989). In their study, they also reported that vowels were breathier when following nasal than oral consonants in three Yi (Loloish) languages. A possible scenario in Korean could be that the weakening of nasality is compensated by an articulatory enhancement of breathy voice on the following vowel. As Garellek et al. have explained, in cases where nasality may be involved, given the acoustic similarities between breathy and nasal voice, acoustic signals are less reliable for spectral tilt measures. Therefore, EGG signals have another advantage of not being interfered by supralaryngeal settings including nasality.

We attempt to use EGG data to look into the second question of this study: How would nasality/denasalization interact with breathiness and how is this interaction affected by the breathiness of lenis and aspirated stops?

### 1.3. Previous perception data

The output of this phonetic investigation would be crucial to understand the mapping of these cues into listeners' perceptual system. What are the acoustic properties present in the production of the four stop series that could be used as perceptual cues?

In a recent perception study on the four stop series, M.-J. Ahn (2019) examined how incongruent consonantal and vocalic cues affected 16 Seoul listeners' stop identification. The author concluded a critical role of VOT in the identification of fortis stops, and a major role of  $f_0$  in the distinction between nasal and oral stops as well as between aspirated and lenis stops. A detailed inspection of her data uncovers some asymmetries that are worth noting. First, the role of  $f_0$  is asymmetric in categorizing nasal and fortis stops: When  $f_0$  of the vowel following a nasal stop was artificially raised, the consonant led to 100% identification of a fortis stop, but when  $f_0$  of the vowel following a fortis stop was artificially lowered, the consonant led to 43.2% identification of a nasal stop, but also 37.5% and 19.3% identification of a fortis and a lenis stop, respectively (M.-J. Ahn, 2019, **Table 3**). This suggests that a high  $f_0$  on the following vowel is sufficient for cueing a fortis stop regardless of the nasal portion (although the nasal properties in the stimuli were not reported) or at least of the prevoiced portion, whereas a low  $f_0$  on the following vowel is necessary but insufficient for cueing a nasal stop. Second, the role of VOT is asymmetric in categorizing the three oral stops: A post-fortis vowel mismatched with a long VOT led to a nearly unanimous identification of an aspirated stop, while a post-aspirated vowel mismatched with a short VOT still led to 51% identification of an aspirated stop, followed by 40.1% identification of a fortis stop. That is, when the following vowel has a high  $f_0$ , a long VOT is sufficient but unnecessary for cueing an aspirated stop, while a short VOT is necessary but insufficient for cueing a fortis stop. Furthermore, a post-lenis vowel mismatched with a short VOT led to a predominant identification of a lenis stop (M.-J. Ahn, 2019, **Table 4**), suggesting that VOT alone is insufficient for cueing a lenis stop.

Other perception studies restricted to oral stops have also shown such perceptual asymmetries. Lee, Politzer-Ahles, and Jongman (2013) varied VOT orthogonally with  $f_0$  of the following vowel (12 VOT steps  $\times$  12  $f_0$  steps). Results were first presented along the two continua, and then broken down by the three most representative steps (min, mid, max) for the two continua (**Figures 7, 9, 11**), showing that stimuli with the highest  $f_0$  were perceived as a fortis stop in the short-mid VOT range and an aspirated stop in the mid-long VOT range but never as a lenis stop, whereas those with the lowest  $f_0$  were perceived as a lenis stop in the mid VOT range but more likely as an aspirated stop when VOT increased. These results replicated in a more refined way those from M. Kim (2004) with a similar experimental design. Therefore, the asymmetric roles of  $f_0$  and that of VOT were also exhibited with aspirated *versus* lenis stops: The bias from a high  $f_0$  or a long VOT yielding the identification of an aspirated stop was stronger than from a low  $f_0$  or a short VOT yielding the identification of a lenis stop.

M. Kim and Lee et al.'s studies have shown another interesting result: Stimuli with very short VOTs were unambiguously perceived as fortis stops, regardless of  $f_0$ . However, an important detail is that these stimuli were constructed with a non-breathy vowel (i.e., extracted from a fortis-initial syllable). In contrast, M.-J. Ahn (2019) has shown that a short VOT was most often perceived as a lenis stop when it was cross-spliced with a breathy vowel (i.e., extracted from a lenis-initial syllable). These apparently contradictory results suggest that the combination of VOT and  $f_0$  might not be sufficient for categorizing the three oral stops and voice quality is likely to be involved. Indeed, M.-R. Kim, Beddor, and Horrocks (2002) concluded that while  $f_0$  was the most salient cue to lenis stops, a combination of VOT,  $f_0$ , and voice quality contributed to the distinction between fortis and aspirated stops. The effect of voice quality has also been studied and confirmed in Francis and Nusbaum (2002), where the perception of five Korean-speaking listeners was affected by multiple acoustic parameters including clarity of formant structure (a voice quality parameter) in natural stimuli. More recently, Schertz, Kang, and Han (2019) conducted a perception study which manipulated orthogonally VOT,  $f_0$ , and voice quality in two

other Korean dialects than Seoul Korean, and found that voice quality played a role in the fortis–aspirated and more substantially in the fortis–lenis distinction.

To summarize, previous perception data revealed perceptual asymmetries. The biases from both VOT and  $f_0$  are asymmetrical in terms of the direction: (a) a high  $f_0$  is sufficient for the identification of a fortis or aspirated stop, but a low  $f_0$  is necessary but insufficient for the identification of a nasal or lenis stop; (b) a long VOT is sufficient for the identification of an aspirated stop, but a short VOT does not exclude the possibility of identifying any of the four stop series. VOT and  $f_0$  seem insufficient for stop identification, thus other dimensions such as voice quality might come into play.

We thus raise the third question of this study: Can the production pattern explain the asymmetries in perception? First, do we observe asymmetrical distributions of VOT and  $f_0$  in production, mirroring the directional biases in perception? Second, given a short VOT and a vowel with low  $f_0$ , the consonant can be identified as a nasal, fortis, or lenis stop, calling for the need of another cue. Do we thus observe the use of voice quality in this VOT– $f_0$  range in particular? Finally, how about the nasal series? Few studies have investigated the perception of the nasal stop series alone. They will be summarized in Section 4.2 to further discuss the link between production and perception.

#### 1.4. The current study

The main purpose of this study is to re-explore the laryngeal properties by linking together the denasalized stop series with the three oral stop series in Korean, as illustrated by the quadruplet 불 /pul/ (lenis) ‘fire,’ 풀 /p<sup>h</sup>ul/ (aspirated) ‘grass,’ 뿔 /p’ul/ (fortis) ‘horn,’ and 물 /mul/ (nasal) ‘water.’ We aim to provide data on VOT distributions of the four stops,  $f_0$ , and voice quality evolution over the vowel following the four stops. In addition, we will briefly examine the acoustic realizations of nasal onsets. We will focus on phrase-initial position, but discussions will also be made on prosodic and positional variations. Throughout the paper, we still refer to the four stop series as ‘aspirated,’ ‘fortis,’ ‘lenis,’ and ‘nasal’ for simplicity purposes, even though some terms do not accurately describe the phonetic nature of the targeted sounds. (Fortunately enough, nasals are also stops, thus we do not go through further terminological choices.)

## 2. Method

### 2.1. Participants

Nine native speakers of Korean (5F, 4M) from Seoul and its surrounding area, Gyeonggi region, participated in the recording. They were aged 23 on average at the time of the recording (between 19 and 29). They were students at Sophia University, Tokyo, and none reported any speech or hearing disorder. Except for one speaker (M04) who lived in Japan for seven years at the time of the recording and rated his Japanese conversational skill at 5/5, the average length of stay in Japan of the other speakers was six months (from 3 to 18) and their average self-evaluation of Japanese conversational skill was 2.9/5 (from 1 to 4). Three additional female participants were recruited but their electroglottographic (EGG) signals could not be interpreted due to excessively high noise-to-signal ratios. They all received prepaid gift cards for their participation. The experiment was approved by the ethics committee of Sophia University (No. 2017-63). All participants signed a written informed consent form.

### 2.2. Speech materials

The CV target syllable consisted of a stop onset (nasal, lenis, fortis, aspirated), either labial or alveolar, followed by one of the base vowels in Korean /i, e, a, ʌ, o, u, i/. Velar onsets were not included because syllable-initial velar nasals are phonotactically prohibited (unless by resyllabification). This made a total of 56 target syllables (4 stop series × 2

places of articulation  $\times$  7 vowels). The target syllable was followed by a copula, /ta/ (‘-다’) (starting with a lenis stop). As a training before the recording of /CV-ta/ phrases, participants were instructed to read twice the seven base vowels followed by /ta/. The vowel data were also analyzed for  $f_0$  and voice quality. They will be referred to as zero-onset syllables in the following.

It is difficult to find real quadruplets with comparable morphosyntactic structures and frequencies. We were thus led to opt for the phrase /X-ta/. The most reasonable interpretation of this phrase would be “This is X,” as if one were spelling out a monosyllable in a carrier phrase. Since many monosyllables and disyllables are polysemic in Korean, participants might have certain semantic interpretations, or might have simply read them as syllable blocks in the Hangul chart. They were mostly capable of producing the phrase with the intended intonation, that is, the target syllable was focused and domain-initial followed by an unfocused /ta/. Occasionally, /ta/ was produced with a higher or unnatural pitch after an initial low boundary tone, possibly because the whole accentual group was analyzed as non-focused, undergoing an LH intonation pattern. In such cases, the first syllable would be unfocused, but we still expect it to maintain its pitch and voice quality pattern.

We also noticed that the phrase-medial /t/ in the copula /ta/ was produced predominantly without closure voicing by all female speakers and half of the male speakers. Despite general assumptions concerning the voicing of phrase-medial lenis stops, it has been empirically shown that these stops are often fully or partly voiceless with a slow speech rate (Jun, 1994). We thus do not attribute the voiceless realization of /t/ only to the use of our nonce phrase, and if anything, the general slow speech rate in producing such a phrase might contribute to more voiceless occurrences. In any event, there should be little effect on our target syllable. Each phrase /X-ta/ was presented on one slide in Hangul orthography in a fixed order of nasal, fortis, lenis, aspirated onsets, first labial, then alveolar for each onset series. We intended to give participants the impression that the task was to read a list of syllables so that they could avoid focusing on the semantic interpretations. Thus, we chose a fixed order so as to avoid misreading in such a task, and also to maximize their consistency in intonation throughout a series.

Participants differed in speech rate, thus the repetition number varied from two to four, making 1709 tokens of /CV/ syllables in total for the analysis. (The length of the recording was decided in advance due to the fixed time schedule for the experiment.) All the analyzed items are given in **Table 1**. In addition to phrase-initial position, the stop onsets were also recorded in phrase-medial position /a-CV-ta/, and fricative and liquid onsets were recorded in both positions, but they will not be presented in this study. Only phrase-medial nasals were used as a comparison with phrase-initial nasals (see Section 2.4).

**Table 1:** List of the target syllables used in the current study.

vowel	nasal		fortis		lenis		aspirated		
이	i	미 mi	니 ni	피 p'i	티 t'i	비 pi	디 ti	피 p <sup>h</sup> i	티 t <sup>h</sup> i
애	e	매 me	내 ne	페 p'e	테 t'e	베 pe	테 te	페 p <sup>h</sup> e	테 t <sup>h</sup> e
아	a	마 ma	나 na	파 p'a	타 t'a	바 pa	다 ta	파 p <sup>h</sup> a	타 t <sup>h</sup> a
어	ʌ	머 mʌ	너 nʌ	퍼 p'ʌ	터 t'ʌ	버 pʌ	더 tʌ	퍼 p <sup>h</sup> ʌ	터 t <sup>h</sup> ʌ
오	o	모 mo	노 no	포 p'o	토 t'o	보 po	도 to	포 p <sup>h</sup> o	토 t <sup>h</sup> o
우	u	무 mu	누 nu	푸 p'u	투 t'u	부 pu	투 tu	푸 p <sup>h</sup> u	투 t <sup>h</sup> u
으	ɨ	미 mi	니 ni	피 p'ɨ	티 t'ɨ	비 pi	디 ti	피 p <sup>h</sup> ɨ	티 t <sup>h</sup> ɨ

### 2.3. Recording

Participants were recorded individually in a sound-proof room at Sophia University in summer 2019. Speech materials were presented on a laptop computer. An electroglottograph (Glottal Enterprises EG2-PCX2) and an electret condenser microphone (Sony ECM-MS957) were connected to another laptop computer via an audio interface (Edirol UA-25EX). Acoustic and EGG signals were recorded simultaneously using Audacity at a 44.1 kHz sampling rate. Each recording session lasted up to one hour, including equipment testing and EGG signal detection. Oral and written instructions were given in Korean.

EGG is a non-invasive technique which records variations of vocal fold contact with time, by means of two electrodes placed symmetrically around the speaker's neck on each side of the thyroid cartilage (for a review, see Henrich, d'Alessandro, Doval, & Castellengo, 2005). The electrical impedance varies with the opening and closing of the glottis, which is reflected in the EGG signals. When signal-to-noise ratios are high, EGG signals give a reliable estimation of the degree of breathiness with a high temporal resolution, with little interference from other sound sources than the glottal source.

### 2.4. Data processing and analyses

**Segmentations.** Recordings were segmented manually in Praat version 6.1.09 (Boersma & Weenink, 2020) by the second author, a native speaker of Korean, and checked at random by the first author. Within the target syllable, the first ascending zero-crossing point on the waveform at F1 onset was determined as the vowel onset, and the disappearance of the formants above F2 as the vowel offset. VOT intervals were marked manually for the four stop series. For voiceless stops, VOT intervals corresponded to the onset of the following vowel and the onset of the release burst; for voiced stops (i.e., most nasal stops), VOT intervals corresponded to the time lag between the onset of regular glottal pulses during the closure—based on both acoustic and EGG signals—and the onset of the release burst when it was visible on the spectrogram. When the release burst was not visible, the segment was excluded from VOT analyses.

**Voice quality.** *Praatdet*, a suite of Praat scripts, was used to estimate  $f_0$  and glottal open quotient ( $Oq$ ) values over the vowel intervals from the EGG signals (Kirby, 2017).  $Oq$  corresponds approximately to the ratio between the glottal open phase and the glottal cycle, which correlates with the degree of breathiness during phonation and the steepness of spectral tilt in the acoustic signal.  $Oq$  generally increases from laryngealized voice to modal voice, then to breathy voice, on a simplified glottal openness continuum. An accurate method of estimating  $f_0$  values is based on the derivative of the EGG signals (dEGG) by the detection of closing peaks corresponding to glottal closing instants (Childers, Hicks, Moore, & Alsaka, 1986; Childers, Hicks, Moore, Eskenazi, & Lalwani, 1990).  $Oq$  can be reliably estimated by the detection of closing and opening peaks on the dEGG signals, but the detection of opening peaks can be less accurate. Threshold methods can also be used to detect opening and closing instants. A combination of the dEGG method and the threshold method has been proposed by D. M. Howard (Howard, 1995; Howard, Lindsey, & Allen, 1990). For a review of different methods, see Henrich et al. (2005). *Praatdet* computes  $f_0$  and  $Oq$  based on dEGG and Howard's methods. The two methods yielded highly consistent results on our data, with a mean difference at 0.006. We discarded 427 data points out of 17676 (2.4%), for which the difference between the two methods was above two standard deviations from the mean—as this might suggest a poor detection of glottal opening instants—and adopted the dEGG method in our analysis. Finally,  $f_0$  and  $Oq$  values were duration-normalized to nine equidistant time points over the vowel based on linear interpolation and averaging.

**Nasality.** While we lack aerodynamic data to address more directly the nature of nasality—which is not the main purpose of the current study—we will provide some descriptive data based on acoustic measurements. Acoustic duration was measured. As for nasality, the detection of nasal formants and anti-formants proving difficult (e.g., Tabain, Butcher, Breen, & Beare, 2016), we decided to use spectral energy as an indication. If phrase-initial nasals are produced with weakened nasality which make them similar to prevoiced stops, they are expected to have weaker spectral energy—especially in the high-frequency range—than phrase-medial nasals where nasality is reported to be better preserved. For this purpose, we measured the following parameters for nasal stops at both positions, excluding devoiced nasal stops (13% of the phrase-initial nasals): (a) the intensity difference between the nasal onset and the following vowel; (b) spectral moments (Center of Gravity, standard deviation, skewness, kurtosis), based on a 30 ms Gaussian windowed Fast Fourier Transform (FFT) with a 0–5 kHz pass-band filter computed in Praat (power set to 2.0 and smoothing to 100 Hz) (cf. Tabain et al., 2016, for the use of the first two spectral moments in nasal consonants). Measurements in (a) and (b) were taken at the segment midpoint using a 30 ms window length, thus, we excluded from intensity and spectral analyses all tokens with an onset duration below 30 ms.

**Statistical models.** Linear mixed-effect (LME) models were built using the *lmer* function of the *lmerTest* package version 3.1-1 (Kuznetsova, Brockhoff, & Christensen, 2017) in R version 3.6.2 (R Core Team, 2019). The selection of the model was based on AIC, BIC, and likelihood ratio tests. Post-hoc pairwise comparisons were made using the *emmeans* package version 1.4.3.01 (Lenth, 2019) in R. Finally, a classification tree analysis was used to assess the relative importance of each property and will be presented in detail in Section 3.5. The datasets as well as the R code to reproduce analyses and plots can be found in Additional Files. Recordings are accessible upon request.

## 2.5. Limitations

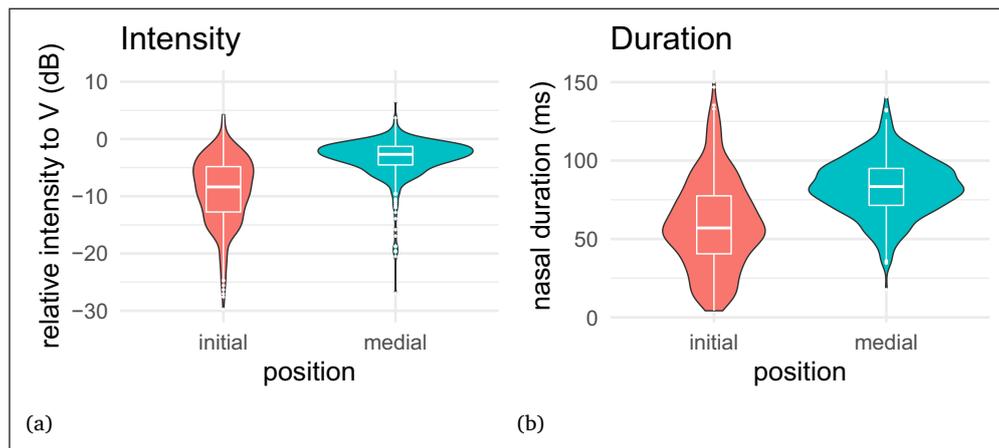
Methodological limitations of this study include the use of nonce phrases alone in the speech materials and the bilingual living environment of our participants. We also acknowledge the low statistical power of the study with the relatively small number of participants, which could result in failure to detect an existing significant difference (e.g., Kirby & Sonderegger, 2018). Hence, we avoid over-interpreting our results when no significant difference is found. Finally, glottal open quotient is one useful indication of breathiness but should not be taken as a parameter which tells everything about the complex voice quality mechanism.

## 3. Results

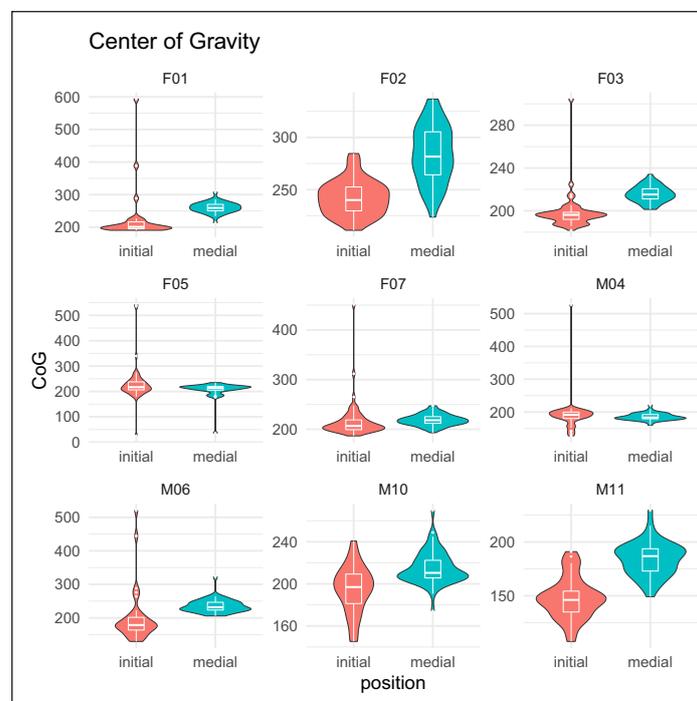
### 3.1. Nasal stops: Variation in nasal weakening

Y.S. Kim (2011) has reported that phrase-initial nasals are often similar, but not identical to word-medial lenis stops that are prevoiced in Korean. While her use of spontaneous speech may have led to frequent medial voicing, closure voicing in medial position is uncommon in our controlled speech (in line with Jun, 1994, as explained in Section 2.2), making it difficult to compare denasalized stops with oral prevoiced stops. For this reason, we will compare phrase-initial with phrase-medial nasal stops, the rationale being that phrase-initial nasals will show lower energy than their phrase-medial counterparts, especially in the high-frequency range, which could be indicated by lower intensity (normalized to vowel), lower Center of Gravity (CoG), greater skewness, and/or greater kurtosis.

**Figure 1a** shows the intensity differences between nasal onsets and their following vowel by position. (Three outliers of less than  $-30$  dB are not shown for aesthetic purposes.) As expected, the relative intensity of nasals to vowels is overall lower in phrase-initial than



**Figure 1:** Violin plot with embedded boxplot by position for nasals **(a)** C-V intensity difference in dB; **(b)** Nasal duration.



**Figure 2:** Violin plot with embedded boxplot by position and speaker for Center of Gravity of nasal onsets.

medial position. The distribution of the intensity is more dispersed for initial position, due to a large speaker variability: Some speakers have distinct patterns between the two positions, while others show a nearly complete overlap. **Figure 1b** shows the acoustic duration of nasal onsets by position, including all phonetic variants. The duration is again overall shorter and more variable in phrase-initial than medial position. And again, speaker variability is observed: At least M10 and M11 do not show a clear duration difference between the two positions. Individual plots for intensity and duration are shown in Appendix A.

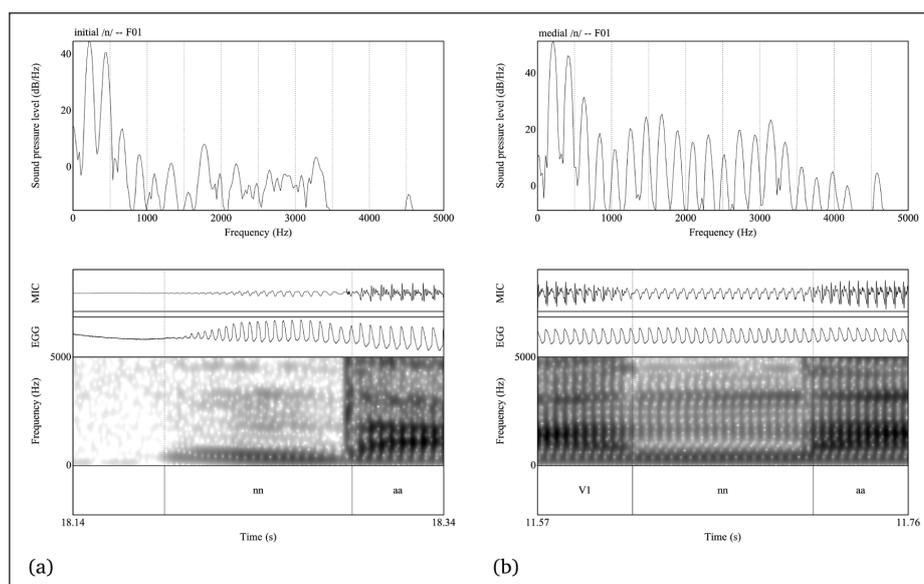
The CoG data are less clear-cut when pooled across speakers. However, as shown by the individual plot in **Figure 2**, most speakers show lower CoG for phrase-initial than medial nasals, which could suggest that phrase-initial nasals had less energy in the high-frequency range. The plots for the other three spectral moments are shown in Appendix A.

The spectral distribution for phrase-initial nasals compared with medial ones has the following patterns for most speakers: overall less dispersed (i.e., with lower standard deviation), similarly skewed but less flat (i.e., with higher kurtosis), suggesting that low-frequency energy is more prominent.

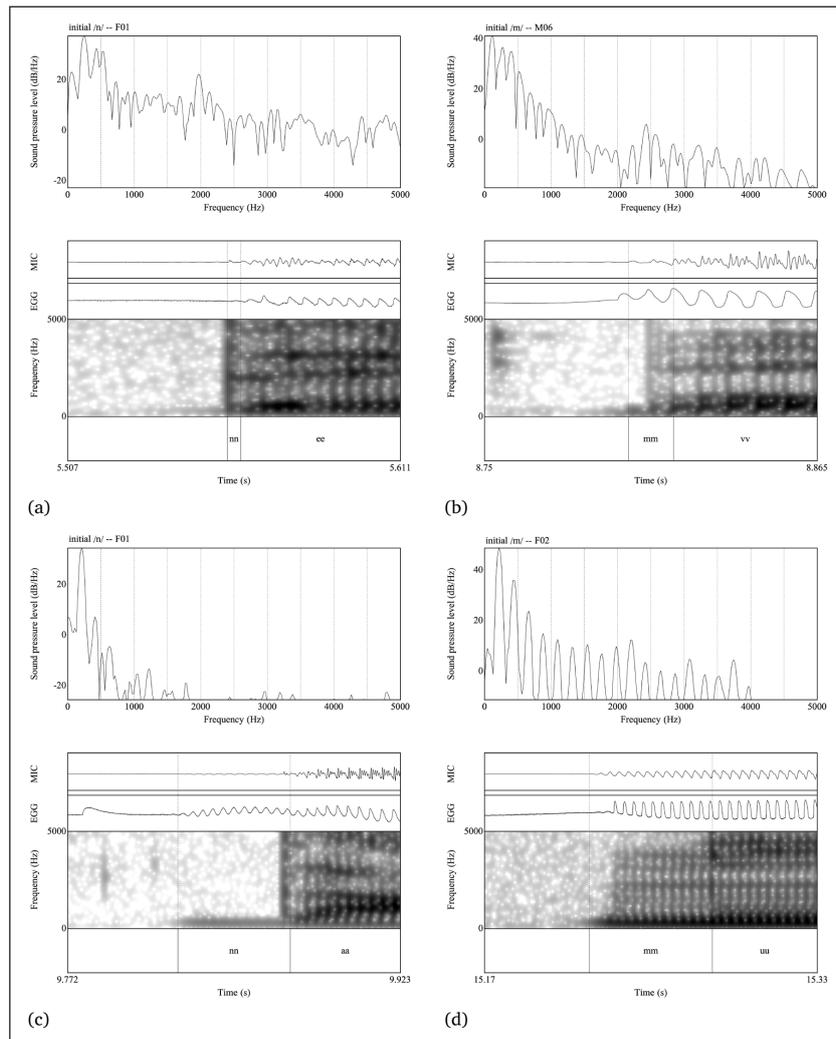
In addition to speaker variability, variable realizations of phrase-initial nasals may also occur within one speaker. They can be a devoiced stop, a prevoiced stop (sometimes with very short voice lead) with no audible nasality, a prevoiced stop with audible nasality (stronger or weaker), or a plain nasal stop. A prevoiced stop with nasality in our data, which is the most typical realization, is characterized by a strong voice bar and a release burst and relatively low energy in higher-frequency regions, as shown in **Figure 3a** (speaker F01). The spectral structure looks similar to an illustration of Spanish prevoiced stop with nasal leak (**Figure 3d** in Solé, 2018). (Note, in passing, that nasal leak is often observed with prevoiced oral stops across languages.) The intensity of the murmur is relatively low but constant, which is different from a typical prenasalized stop with its drop in intensity during the closure signaling the transition from nasal to oral murmur (e.g., Burton, Blumstein, & Stevens, 1992). As a comparison, phrase-medial nasals produced by the same speaker have higher intensity and richer formant structures indicating stronger nasality (**Figure 3b**). **Figure 4** illustrates other realizations of phrase-initial nasals: devoiced, prevoiced with short voice lead, prevoiced with no audible nasality, and plain nasal stop.

### 3.2. VOT distributions

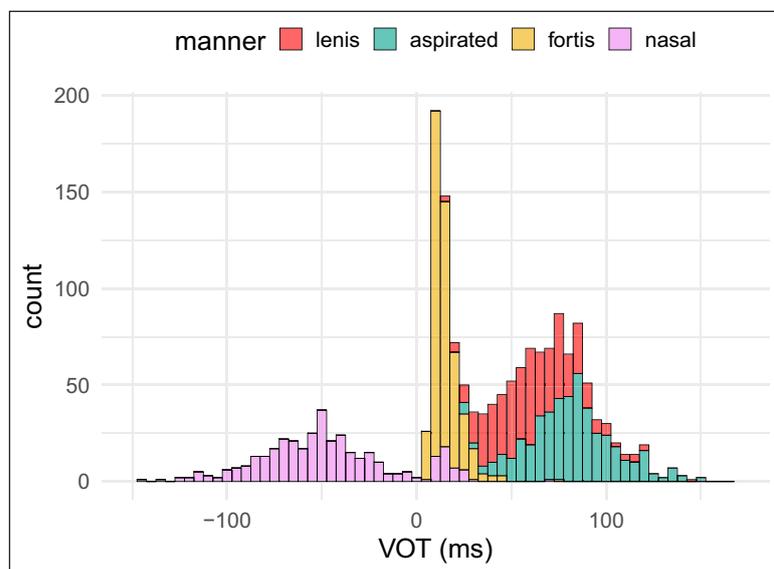
**Figure 5** shows the distributions of VOT of the four stop onsets. (All distribution plots in this article are stacked rather than overlaid, so as to make overlapped bars more visible. For example, the bars of lenis stops are piled up on top of aspirated stops on the rightmost side and on top of the other three stops in the VOT range between 10 and 50 ms. For aesthetic purposes, one outlier with a VOT at 251 ms is not shown in the VOT plots.) Nasal stops are clearly separated from the other three stop series by their negative VOTs, but 13% (48 out of 361 occurrences) of them have positive VOTs. The distributions are approximately trimodal, with strong overlap between lenis and aspirated stops, and some overlap between fortis and the two other oral stop series.



**Figure 3:** From top to bottom: FFT spectrum at nasal midpoint using a 30 ms Gaussian window, waveform, EGG signal, spectrogram of the acoustic waveform. (Amplitude scale: [-1, 1]; dynamic range: 60 dB). **(a)** Phrase-initial /na/; **(b)** Phrase-medial /na/.



**Figure 4:** From top to bottom: FFT spectrum at nasal midpoint in phrase-initial position using a 30 ms Gaussian window, waveform, EGG signal, spectrogram of the acoustic waveform. (Amplitude scale: [-1, 1]; dynamic range: 60 dB). **(a)** Devoiced (speaker F01); **(b)** Prevoiced with short voice lead (speaker M06); **(c)** Prevoiced with no audible nasality (speaker F01); **(d)** Plain nasal (speaker F02). (Intensity and spectral analyses were excluded for (a) and (b)).

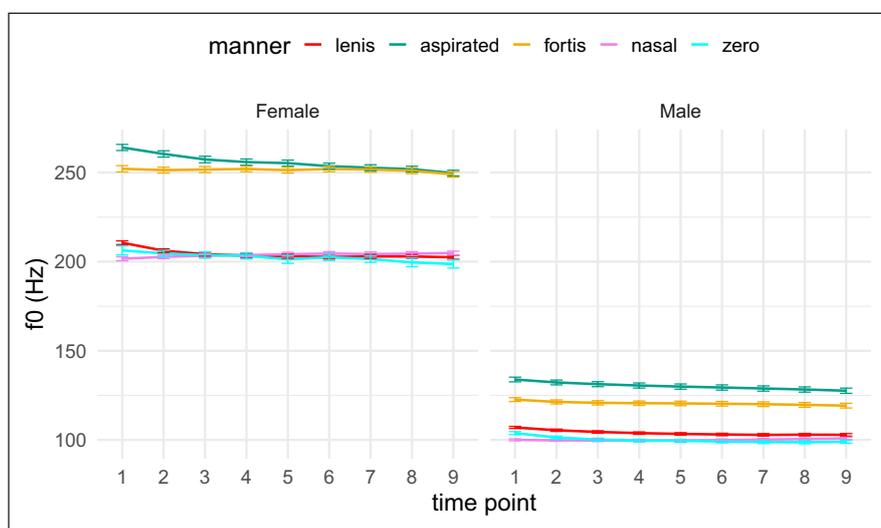


**Figure 5:** VOT distributions (stacked) by onset manner.

An LME model was fitted to the raw VOT data, with sum-coded predictors MANNER, PLACE, and VOWEL, and random intercepts for speakers and repetitions, which all improved the fitness of the model. SEX was not included as it did not improve the model. Estimated marginal means for VOTs of each stop series are shown in **Table 2**. Pairwise comparisons confirm that the contrast between each pair of stops is at  $p < .0001$  level. The summary of the full model and pairwise comparisons is given in Appendix B (Tables B.1 and B.2).

### 3.3. $f_0$ curves

**Figure 6** shows the averaged  $f_0$  curves in Hz over the vowel following each stop series and zero onset (i.e., vowel-only) by sex group. The separation between fortis/aspirated and lenis/nasal/zero onsets is clear and extended over the entire vowel. **Table 3** shows the  $f_0$  values in Hz averaged over all time points after each stop and zero onset. The  $f_0$  difference is about 50 Hz higher after aspirated than lenis stops for females, and about 30 Hz higher for males.



**Figure 6:** Aggregated  $f_0$  curves in hertz (with standard errors) by onset manner and sex group.

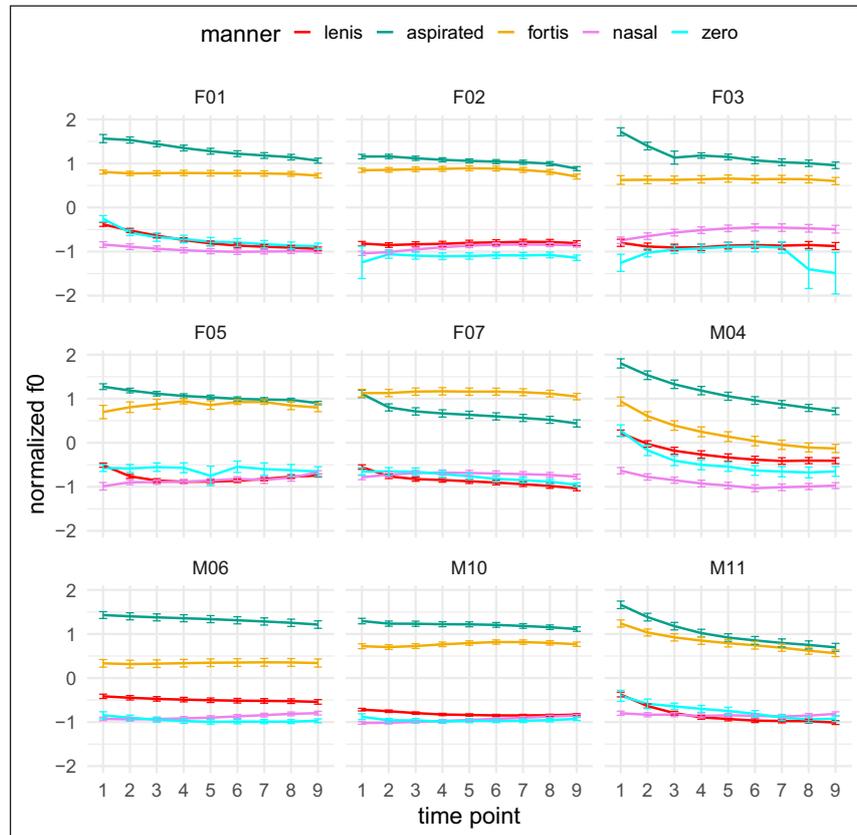
**Table 2:** Estimated marginal means of VOTs of each stop series.

manner	emmean	SE	df	lower.CL	upper.CL
place = labial					
lenis	62.66	3.01	11.08	56.05	69.27
aspirated	82.16	3.00	10.93	75.56	88.76
fortis	16.47	3.01	11.09	9.86	23.08
nasal	-46.96	3.05	11.73	-53.62	-40.30
place = alveolar					
lenis	60.07	3.00	10.99	53.47	66.68
aspirated	79.57	3.01	11.06	72.96	86.18
fortis	13.88	3.00	11.00	7.28	20.48
nasal	-49.55	3.05	11.78	-56.21	-42.88

Results are averaged over the levels of: vowel.  
 Degrees-of-freedom method: kenward-roger.  
 Confidence level used: 0.95.

**Table 3:** Averaged  $f_0$  in Hz (with SD) over all timepoints by onset manner and sex group.

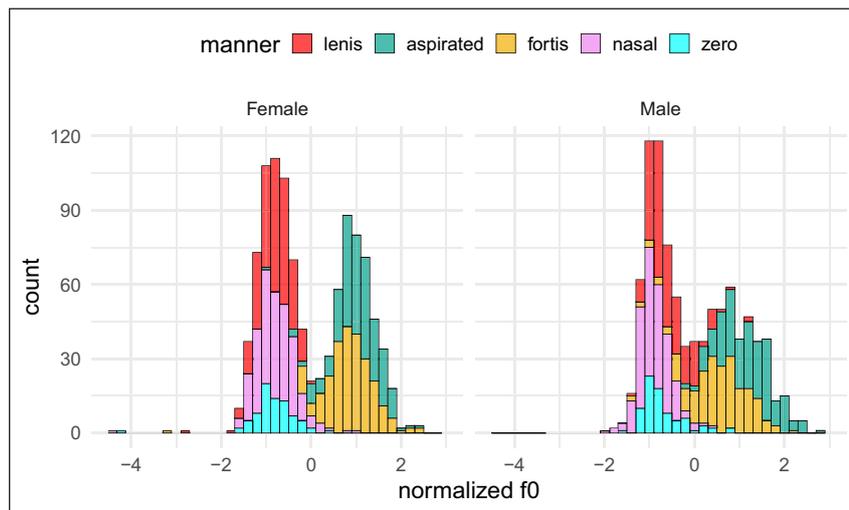
	aspirated	fortis	lenis	nasal	zero
female	251 (26)	253 (25)	203 (17)	203 (17)	204 (17)
male	132 (21)	115 (18)	105 (10)	99 (9)	101 (7)

**Figure 7:** Normalized  $f_0$  curves (with standard errors) by onset manner and speaker.

To visualize the individual variability within the  $f_0$  range of each speaker, **Figure 7** shows the individual plots of normalized  $f_0$  curves scaled and centered by speaker (mean at 0 and SD at 1 for all normalizations in this study), using the `scale_by` function of the `standardize` package (Eager, 2017) in R. The pattern is relatively homogeneous across female speakers, with a clear boundary between the high  $f_0$  zone and the low  $f_0$  zone. While two male speakers follow this pattern, the other two males show a more continuous  $f_0$  (aspirated > fortis > lenis > nasal stops)—their lengths of stay in Japan happened to be the longest (M04: 7 years; M06: 1.5 years).

To visualize the distinctness of  $f_0$  among stop and zero onsets, **Figure 8** shows the distributions of normalized  $f_0$  by sex group, averaged over all time points, as  $f_0$  height is found to be the crucial property, rather than  $f_0$  contour over time. Bimodal distributions are observed both for female and male speakers, with more across-categories overlap for males than females.

An LME model was fitted to the normalized  $f_0$  data averaged over all time points. We excluded the zero onset and focused on the stop series. The predictors, sum-coded, included the maximal factors which improved the model: MANNER, PLACE, and VOWEL. The interactions between MANNER and PLACE and between MANNER and SEX were also added and improved the model. Random intercepts were included for speakers and for



**Figure 8:** Distributions (stacked) of normalized  $f_0$  averaged over all timepoints by onset manner and sex group.

**Table 4:** Pairwise comparisons of estimated marginal means of normalized  $f_0$ .

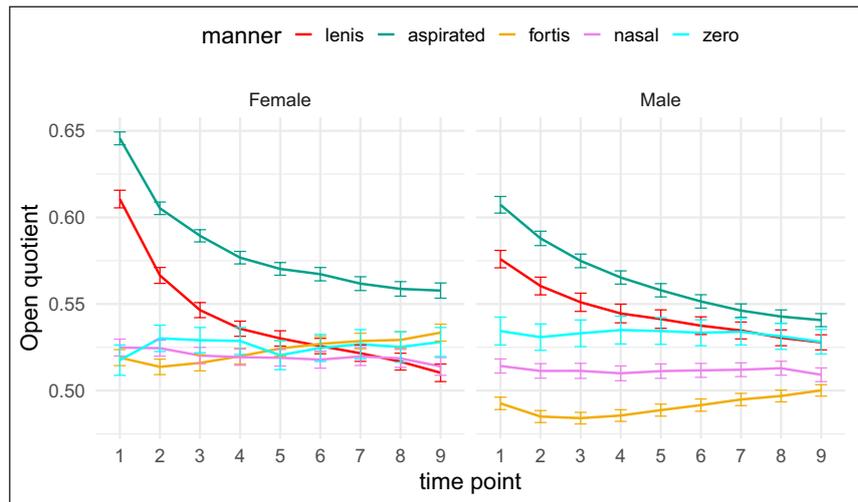
contrast	estimate	SE	df	t-ratio	p-value
sex = F					
lenis – aspirated	-1.83	0.04	1771.23	-44.193	<.0001
lenis – fortis	-1.68	0.04	1772.09	-40.712	<.0001
lenis – nasal	-0.01	0.04	1771.26	-0.181	0.9979
aspirated – fortis	0.15	0.04	1771.89	3.606	0.0018
aspirated – nasal	1.82	0.04	1771.18	44.195	<.0001
fortis – nasal	1.67	0.04	1771.99	40.704	<.0001
sex = M					
lenis – aspirated	-1.77	0.04	1771.41	-40.229	<.0001
lenis – fortis	-1.10	0.04	1771.20	-24.995	<.0001
lenis – nasal	0.31	0.04	1771.20	7.126	<.0001
aspirated – fortis	0.67	0.04	1771.20	15.215	<.0001
aspirated – nasal	2.09	0.04	1771.20	47.486	<.0001
fortis – nasal	1.42	0.04	1771.12	32.198	<.0001

Results are averaged over the levels of: vowel, place.  
 Degrees-of-freedom method: kenward-roger.  
 P-value adjustment: tukey method for comparing a family of 4 estimates.

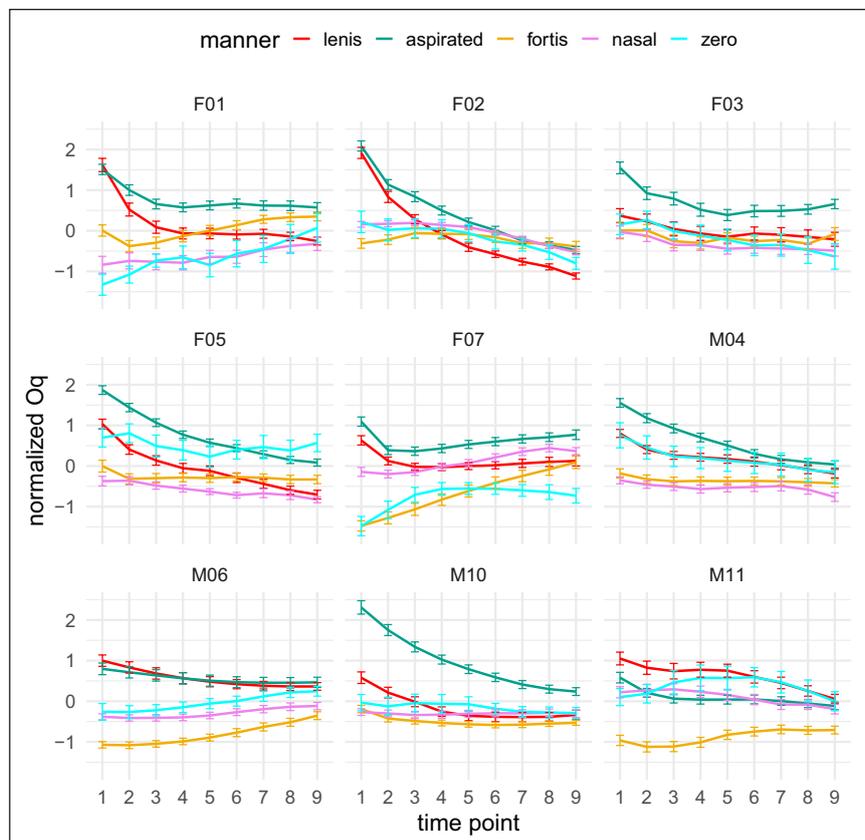
repetitions. MANNER|SPEAKER random slope led to singular fit and thus was excluded. Pairwise comparisons (Table 4) showed a large difference of  $f_0$  between aspirated/fortis stops and lenis/nasal stops and a relatively minor difference between other pairs: aspirated > fortis > lenis ≥ nasal (no lenis-nasal difference for female speakers). Summary of the full model is given in Appendix B (Table B.3).

### 3.4. Oq curves

Figure 9 shows the averaged Oq curves over the vowel following each stop series and zero onset by sex group. Oq is higher after aspirated than after lenis stops, extending over almost the entire vowel, and it is higher after lenis stops than the other stops and zero



**Figure 9:** Aggregated *Oq* curves (with standard errors) by onset manner and sex group.



**Figure 10:** Normalized *Oq* curves (with standard errors) by onset manner and speaker.

onsets, the difference diminishing from vowel onset to offset. *Oq* is on average higher for female than male speakers. The pattern between fortis, nasal, and zero onsets differs between males and females. However, this difference is more likely due to individual than sex-based variability, as shown in **Figure 10**. Three speakers out of nine clearly show higher *Oq* after nasal than fortis stops, contrary to one speaker who clearly shows the reverse pattern, while the remaining five speakers show similar *Oq* curves after fortis and nasal stops. *Oq* is more variable after zero onset, but patterns with nasal stops most frequently.

As pointed out in Section 1.1, previous studies reported inconsistent results with respect to the voice quality difference between lenis and aspirated stops. Our individual data of normalized  $Oq$  (scaled and centered by speaker) (**Figure 10**) show higher  $Oq$  for aspirated than lenis stops for most speakers (7 out of 9) mostly over the entire vowel. One speaker (M11) shows the reverse pattern, and one speaker (M06) shows similar  $Oq$  curves after the two stop series.

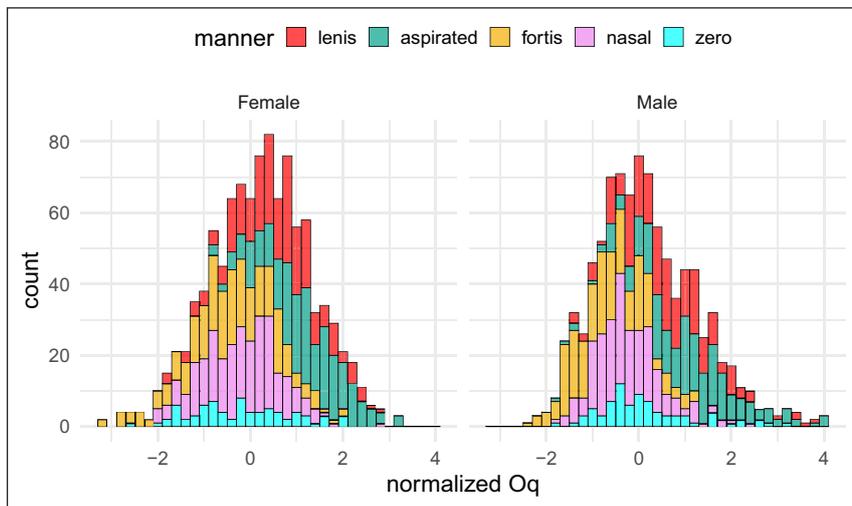
To visualize the distinctness of  $Oq$  among stop and zero onsets, **Figure 11** shows the distributions of normalized  $Oq$  by sex group. Since the  $Oq$  difference across all onset types is shown notably at the first part of the vowel, only the first three time points are taken into account, which correspond to the first third of the entire vowel. Unlike the trimodal VOT distributions or the bimodal  $f_0$  distributions, the  $Oq$  distributions are strongly overlapped between the onset types.

Restricted to the first three time points, an LME model was fitted to the normalized  $Oq$  data (time point  $\leq 3$ ), excluding the zero onset. The predictors, sum-coded, included the maximal factors which improved the model: MANNER and VOWEL. The interaction between MANNER and VOWEL was also added and improved the model. Random intercepts were included for speakers, and by-speaker random slopes were included for MANNER, which accounted for variable effects of MANNER across speakers. (Random intercepts for repetitions were excluded since the model was overfitted.) Summary of the full model is given in Appendix B (Table B.4). Pairwise comparisons (**Table 5**) show no difference for

**Table 5:** Pairwise comparisons of estimated marginal means of normalized  $Oq$  for all speakers.

contrast	estimate	SE	df	t-ratio	p-value
lenis – aspirated	-0.51	0.21	8.00	-2.415	0.1513
lenis – fortis	1.13	0.19	7.99	6.064	0.0014
lenis – nasal	0.81	0.12	7.98	6.576	0.0008
aspirated – fortis	1.63	0.10	7.95	16.909	<.0001
aspirated – nasal	1.32	0.21	8.00	6.163	0.0012
fortis – nasal	-0.32	0.20	8.00	-1.544	0.4579

Results are averaged over the levels of: vowel.  
 Degrees-of-freedom method: kenward-roger.  
 P-value adjustment: tukey method for comparing a family of 4 estimates.



**Figure 11:** Distributions (stacked) of normalized  $Oq$  averaged over the first three timepoints by onset manner and sex group.

the lenis–aspirated pair and the fortis–nasal pair, but a difference between all other pairs. However, based on **Figure 10**, the same analysis (but including random intercepts for repetitions) without speakers M06 and M11 shows that  $Oq$  is higher for aspirated than lenis stops (**Table 6**).

### 3.5. Classification tree analysis

The above analyses show that all three properties contribute to some degree to the distinction between certain members of the four stop series. What is, then, the relative importance of each property in classifying each stop series?

A classification and regression tree (CART) analysis was used to determine the predictors of each category, that is, the produced stop series, using the `rpart` package version 4.1-15 (Therneau, Atkinson, & Ripley, 2019) in R. The analysis explores the predictor used in each binary partition and the cut-off point which maximizes the homogeneity of two split groups, recursively until the maximal homogeneity within each group. This analysis has been used in van Alphen and Smits (2004) for assessing the ranking of acoustic and perceptual cues in the classification of Dutch initial plosives, and in Brunelle (2009) for assessing the ranking of perceptual cues in the classification of Vietnamese tones. The resulting tree gives a visual representation of the relative importance of each predictor in classifying a homogeneous group. However, as warned in Brunelle (2009), it should *not* be taken that such recursive decision-making process is used by speakers in their production or perception of a phonological category.

In our analysis, the response variable was stop category. The three properties, VOT (in ms),  $f_0$  (in semitone (st), averaged over the vowel),  $Oq$  (normalized, of time points 1–3), were included as numerical predictors, and MANNER, PLACE, VOWEL, and SEX as categorical predictors. Complexity parameter (CP, i.e., the minimum improvement of the model at each node) was set to 0.01. No post-pruning was needed as the analysis already had the lowest prediction error rate (7.11%) in cross-validation (CP table and confusion matrix in Appendix C).

**Figure 12** shows the classification tree plotted by the `rpart.plot` package version 3.0.8 (Milborrow, 2019) in R. Only VOT and  $f_0$  are used in the tree construction. VOT first divides the four stop series into two: When VOT is longer than or equal to 30 ms, it is most probably a lenis or an aspirated stop; otherwise, it is most probably a fortis or a nasal stop. Within the  $VOT \geq 30$  ms group,  $f_0$  further divides the stops into two groups: It is most probably a lenis stop when  $f_0 < 0.15$  st (mean  $f_0$  is at 0.09 st) and an aspirated stop, otherwise. Within the  $VOT < 30$  ms group, when VOT is less than 3 ms, it is certainly a nasal stop; otherwise,  $f_0 < -0.75$  st predicts very well a fortis stop, and  $f_0 < -0.75$  st predicts reasonably well a nasal stop.

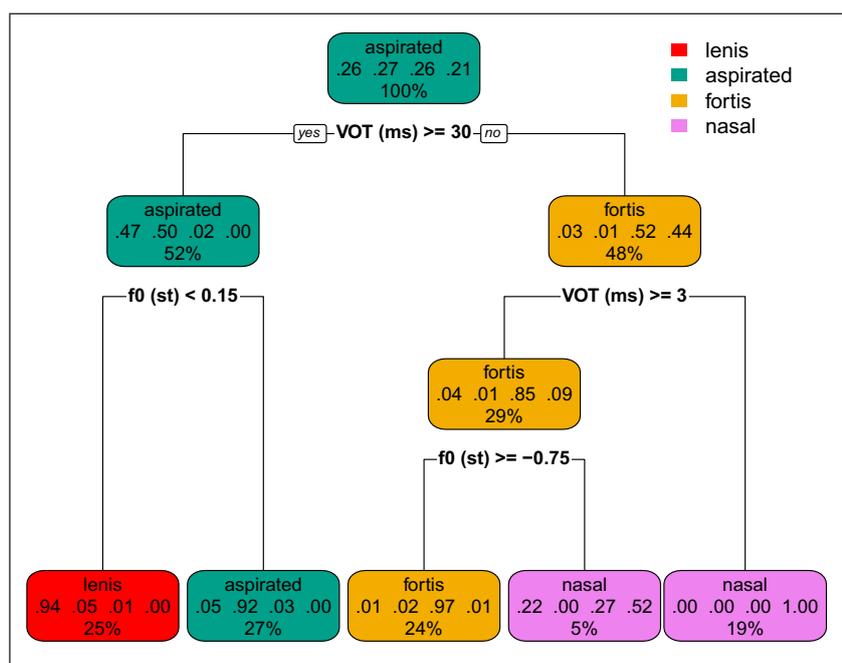
**Table 6:** Pairwise comparisons of estimated marginal means of normalized  $Oq$  excluding speakers M06 and M11 who show different patterns.

contrast	estimate	SE	df	t-ratio	p-value
lenis – aspirated	–0.75	0.17	5.99	–4.397	0.0179
lenis – fortis	0.90	0.14	5.98	6.229	0.0033
lenis – nasal	0.78	0.14	5.98	5.609	0.0056
aspirated – fortis	1.65	0.12	5.96	13.901	<.0001
aspirated – nasal	1.53	0.17	5.99	8.845	0.0005
fortis – nasal	–0.11	0.19	5.99	–0.593	0.9306

Results are averaged over the levels of: vowel.

Degrees-of-freedom method: kenward-roger.

P-value adjustment: tukey method for comparing a family of 4 estimates.



**Figure 12:** Classification tree of the stop categories in production. For each split, the left branch ('yes') satisfies the rule based on which the split is made and the right branch ('no') does not. Below each node, the predicted probability of each response category in the order 'lenis,' 'aspirated,' 'fortis,' and 'nasal' is printed, followed by the percentage of their observations out of the entire dataset.

## 4. General discussion

### 4.1. Phonetic properties of the four stop series

This study examined the phonetic properties of the three oral stop series and the nasal stop series in phrase-initial position in Seoul and Gyeonggi Korean. We first briefly reported the nasal or denasalized property of nasal stops based on our acoustic data. We have shown an overall weaker nasality and shorter duration of nasal stops in phrase-initial compared to phrase-medial position, but also important variability, both between- and within-speaker, in line with previous studies. It is also possible that denasalization occurs less often in our read speech than in spontaneous speech as reported in Y. S. Kim (2011), but we will not make any comparisons due to the methodological differences between the two studies.

More crucially, we examined the laryngeal properties of the four stop series in phrase-initial position, namely VOT,  $f_0$ , and voice quality as estimated by glottal open quotient. Our VOT results of the (voiceless) oral stops are highly consistent with previous reports: shortest for fortis, longest for aspirated, and intermediate for lenis stops; lenis and aspirated stops overlap with each other and they both remain well distinguished from fortis stops. Of the nasal stops, 13% are devoiced, but in the majority of the cases, they can be distinguished from the other three stop series based on VOT alone. Our  $f_0$  results are also in line with previous findings, with a clear bimodal distribution: Aspirated and fortis stops are followed by high  $f_0$ , while lenis, nasal stops, and zero onsets are followed by low  $f_0$ , and this difference is observed over the entire vowel. For two male speakers who have stayed in Japan longer than the others,  $f_0$  following fortis and lenis stops is in an intermediate zone between aspirated and nasal stops, but no causal link can be established by our limited data.

Voice quality, on the other hand, is less straightforwardly conditioned by the stop onset. Overall, lenis and aspirated stops are followed by breathier vowels than fortis and nasal stops, especially at the beginning of the vowel. We raised two more specific questions in Sections 1.1 and 1.2:

- (1) How does voice quality on the following vowel differ between aspirated and lenis stops? Is the discrepancy in previous findings due to the time course of the voice quality over the entire vowel?
- (2) Is the weakening of consonantal nasality compensated by an enhancement of breathier voice on the following vowel (due to its acoustic similarity with nasality)?

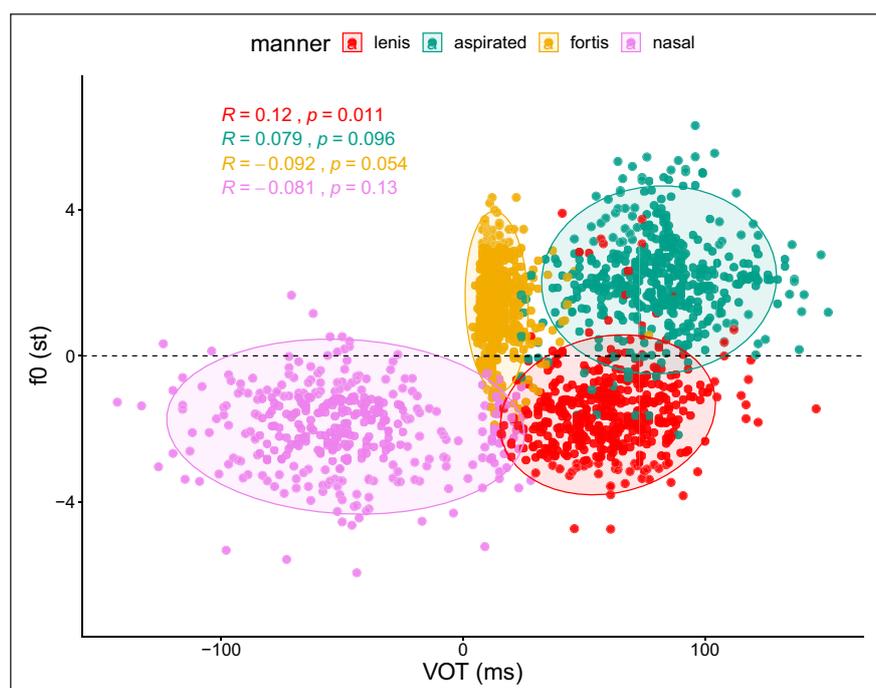
For question (1), our results did not show any compelling patterns regarding the time course after aspirated and lenis stops, but they demonstrated speaker variability which could also be a reason why previous studies had different results. Nevertheless, for seven out of nine speakers, the vowel was breathier following aspirated than lenis stops. For question (2), no difference of breathiness was found between nasal and fortis stops. But this result needs to be interpreted with caution due to the low statistical power. Note, however, that **Figure 10** shows that for three out of nine speakers,  $Oq$  was higher after nasal than fortis stops, which could suggest some articulatory enhancement with a breathier voice after nasal stops, or a more laryngealized voice with their production of fortis stops.

More globally, the contribution of voice quality to the overall distinction between the four stop series is less robust than VOT or  $f_0$ , as indicated by the strong overlap of  $Oq$  after the four stop onsets and important individual variations, and further confirmed by our CART analysis.

#### 4.2. Mismatch between production and perception

Another issue we raised was whether and how the production data could explain the two intriguing issues in perception with respect to VOT and  $f_0$ . VOT and  $f_0$  from our production data are displayed in a scatterplot (**Figure 13**). Let's examine them in relation to the observations on previous perception data (see Section 1.3).

- (1) Can the distributions of  $f_0$  and VOT in production explain the asymmetrical perceptual biases in terms of the direction, that is, the perceptual bias from high  $f_0$  or long VOT is stronger than from low  $f_0$  or short VOT?



**Figure 13:** Scatterplot of  $f_0$  (st) against VOT (ms). (Ellipse type: multivariate t-distribution).

- (2) Given a short VOT and a vowel with low  $f_0$ , the consonant can be identified as a nasal, fortis, or lenis stop. Are there any other cues that might be involved to distinguish between them in this VOT- $f_0$  range?

**The perceptual bias from high  $f_0$ .** Concerning the production of  $f_0$ , we could expect the distributions of aspirated and fortis stops to be skewed into the low  $f_0$  range (i.e., negatively skewed), but to a lesser degree in the other direction for the other two stop series, so that listeners could be less used to the co-occurrence between high  $f_0$  and nasal/lenis stops. However, this is not exactly the case. As shown in **Figure 13**, the observed numbers of intrusions are comparable between the high and low  $f_0$  range approximately divided by the 0 semitone line. Median skewness was measured on the  $f_0$  distributions after the four stop series, showing a symmetric distribution for the aspirated series (-0.01)—contrary to the expectation—and skewed distributions of a similar degree for the other three stop series (fortis: -0.27, lenis: 0.26, nasal: 0.22). Thus, the directional bias of  $f_0$  does not find a clear answer in the production pattern. We then propose a tentative explanation which is perception-based: Korean-speaking listeners are more sensitive to high  $f_0$  than to low  $f_0$ , which could have a language-independent and/or a language-specific basis. First, the human perceptual system, notably the attentional system, is found to be more sensitive to raised than lowered  $f_0$  in speech (Hsu, Evans, & Lee, 2015, and references therein). Second, the increase of the  $f_0$  difference in Korean might be more attributable to the  $f_0$  raising after aspirated and fortis stops than the  $f_0$  lowering after the other onset series, as suggested by the  $f_0$  enhancement in clear speech of young speakers by raising the high  $f_0$  rather than lowering the low  $f_0$  (K.-H. Kang & Guion, 2008, **Figure 1c**).

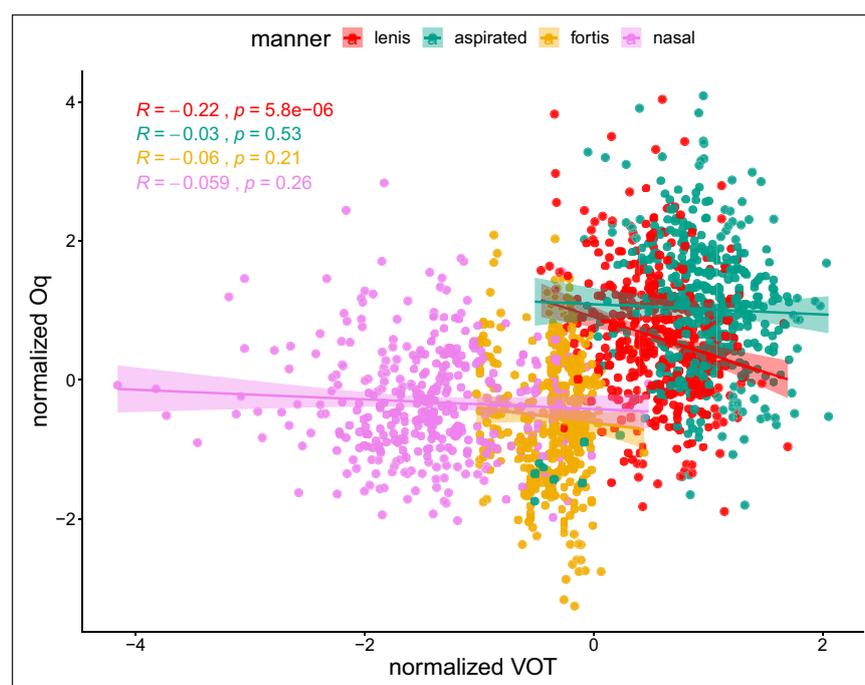
**The role of breathiness in production and perception.** Concerning the production of VOT, aspirated stops are much more likely to occur within the very long VOT range than lenis stops (which also reflects the historical change of VOT shortening of aspirated stops). This probably leads to a strong bias towards the identification of an aspirated stop when the stimulus has a very long VOT even combined with a low  $f_0$ . In the other direction, aspirated stops are rarely produced with a very short VOT, thus we might also expect that short VOTs in the stimuli would exclude the identification of an aspirated stop in perception. However, previous studies have shown that, when followed by a breathy vowel, a short VOT may still be perceived as an aspirated stop. Therefore, we conjecture that aspiration is not a necessary cue in perception when breathiness is involved. It is possibly a perceptual bias that a breathy vowel may be re-interpreted as aspiration in perception. This perceptual bias has probably served as a phonetic basis for a sound change from breathy voice to aspirated onsets, which has occurred in a number of languages, including Central Thai (e.g., Abramson & Erickson, 1992), Tamang (Tibeto-Burman) (Mazaudon, 1978) (see also a Tai dialect in Pittayaporn & Kirby, 2016, and their review of literature). On the other hand, in producing an aspirated stop, aspiration appears as an active gesture, suggested by a long VOT, which corroborates previous physiological studies showing higher glottal opening peak and airflow peak for aspirated than lenis stops (H. Kim et al., 2018).

In the low  $f_0$  range, our production data indicate that short VOT + low  $f_0$  can be mapped into three stop categories: nasal, fortis, and lenis, although neither nasal stops nor fortis stops are prototypically produced in this zone. This mirrors exactly the perceptual confusion in this VOT- $f_0$  zone. We will come to nasal stops in the following. How, then, do speakers and listeners distinguish lenis from fortis stops? Again, voice quality may come into play. In our production data, an overall difference has been found between lenis and fortis stops, in that lenis stops are produced by a breathier voice than fortis stops. If it is indeed the case that voice quality is a rescue option when the other two cues

are ambiguous, we could expect speakers to actively control voice quality, for example, by enhancing the voice quality difference in this ambiguous VOT- $f_0$  zone. Closer scrutiny into the VOT- $Oq$  relation does suggest this trend. **Figure 14** plots normalized  $Oq$  against normalized VOT (scaled and centered by speaker to minimize inter-speaker variability), showing a negative correlation, though weak, only for lenis stops ( $R = -0.22$ ,  $p < .0001$ ). This suggests that, when VOT of a lenis stop decreases, the following vowel is breathier. In contrast, our production data do not show such a correlation for aspirated stops, suggesting that breathiness on the vowel might be a consequence of aspiration rather than an intentional laryngeal maneuver to produce a breathy vowel.

To sum up, breathiness seems to be adopted in production as a strategy for the disambiguation of a lenis stop in perception. It is interesting to note that breathiness has been found to trade with multiple parameters perceptually, including nasality (see Section 1.2), aspiration (see this section), low  $f_0$  (Gao, Hallé, & Draxler, 2020; Kuang & Liberman, 2018), and consonantal voicing (Kingston & Diehl, 1994). In Korean lenis stops, breathiness possibly trades with both aspiration and low  $f_0$ . On the other hand, breathiness might not be *actively* produced in the production of an aspirated stop, but it plays a role in perception when it is artificially created in the stimuli.

**The perceptual use of nasality.** Although nasality is weak and inconsistent in production, two of our previous studies on Korean nasal stops show an interesting paradox in its perception: Nasality contributes to the identification of a nasal stop, but not necessarily to its goodness rating. In one study (Yun et al., 2020), Korean-speaking listeners were presented with isolated /Ca/ stimuli along the oral-nasal continuum created by a mechanical vocal-tract model and were asked to choose between a nasal-onset and an oral-onset syllable. Stimuli with stronger nasal coupling categorically increased the nasal response rate. In the other study with isolated Klatt-synthesized /Ca/ stimuli (Yun & Arai, 2020), longer nasal portion increased the nasal judgement by Korean-speaking listeners, however, two-thirds of the listeners judged oral stimuli as better /na/ syllables than nasal stimuli in a /na/-



**Figure 14:** Scatterplot of normalized  $Oq$  (first 3 timepoints) against normalized VOT, with fitted regression lines and coefficients.

goodness rating test. These results suggest that nasality plays an important role in the nasal stop identification, and strong nasality is very likely to be a sufficient cue to a nasal stop. However, when categorized in advance as a nasal stop, nasal stimuli were mostly judged as less prototypical of a nasal stop than oral stimuli. In the following section 4.3, we will discuss why nasality, still useful in perception, is weakened in production.

#### **4.3. Syntagmatic variation and paradigmatic contrast**

Several proposals have related the synchronic pattern of nasal ‘weakening’—despite this term—to the domain-initial ‘strengthening’ effect which accounts for variations in segmental properties according to their prosodic hierarchy (Cho & Keating, 2001; Yoshida, 2008). Crosslinguistically, a weaker nasal energy in domain-initial than medial nasal stops is also found in French, as evidenced by their results from nasal flow measures averaged across four participants (Fougeron, 2001), although probably to a much lesser degree than in Korean and never fully denasalized. In terms of the duration differences, Fougeron reported that both articulatory and acoustic durations of nasals were longer in domain-initial than medial positions, representing a form of domain-initial strengthening. In English, on the other hand, while the articulatory duration of nasals is longer in utterance-initial than other positions, their acoustic duration is shorter (Fougeron & Keating, 1997). Cho and Keating (2001) also reported such a pattern for three speakers of Korean (although they excluded quite a few utterance-initial nasal stops with very short articulatory duration due to sampling limitations with electropalatography). Their acoustic duration results are corroborated by ours. Therefore, if we assume that denasalization is a form of domain-initial strengthening, it is clear that its phonetic manifestation is different across languages: In Korean, the acoustic duration and the nasal articulation are in fact both reduced. Prosodically driven enhancement is often found to strengthen language-specific segmental features (Cho & McQueen, 2005) and suprasegmental/tonal features (Y. Chen & Gussenhoven, 2008), maximizing paradigmatic contrasts (Georgeton & Fougeron, 2014). However, if we assume that nasals in Korean have the [+nasal] feature, following the analysis of regressive nasal assimilation rule in Yoo (2016), this feature is in fact weakened in prosodically stronger positions. For the present, our reasoning is based on several assumptions, hence we must leave this debate to future research on the interface between segmental and prosodic features (see relevant discussions in Yoo & Nolan, 2020, Section 4.6). On the perceptual side of prosodic units, denasalization may serve as a cue in prosodic parsing. This possibility has been hinted at in Cho and Keating (2001) and Y. S. Kim (2011, p. 142), then supported by empirical evidence from a perceptual study (Yoo, 2018). Plausibly, Y. S. Kim has also pointed out that this perceptual cue is probably made more important by the regressive nasal assimilation process in Korean which may give rise to sequences of nasals at syllable boundaries, calling for a need to demarcate initial boundaries from non-initial nasals.

Another possibility, which might be hidden in plain sight, is that rather than strengthened, nasal stops undergo nasal weakening as a form of hypoarticulation. The motivation of this hypoarticulation could include speech style, lexical frequency, and predictability (Lindblom, Guion, Hura, Moon, & Willerman, 1995), which needs to be addressed in further research. In the meantime, the syntagmatic variation related to prosody/position should not compromise the maintenance of the paradigmatic contrast (when the functional load is high). Thus, nasal stops might *resist* hypoarticulation in domain-medial position, which would cause perceptual confusion with a lenis stop, while their hypoarticulation would not have this effect in domain-initial position. Although in our read speech, phrase-medial lenis stops are rarely fully voiced, they are more often so in spontaneous speech (e.g., Y. S. Kim, 2011). Even devoiced lenis stops

are almost twice shorter in closure duration than fortis stops (e.g., H. Kim et al., 2018), which could lead to a percept of a voiced stop. Therefore, denasalization is avoided to maintain the paradigmatic contrast between nasal and lenis stops. In our previous study on the position-dependent devoicing and  $f_0$  enhancement in Tokyo Japanese, we have explained this different phenomenon using the same motivation: positional variation and the maintenance of paradigmatic contrast (Gao & Arai, 2019; Gao, Yun, & Arai, 2019).

The realization of nasal stops is also highly variable even in phrase-initial position. This is in line with what is observed in corpus data (Yoo & Nolan, 2020) which, based on their apparent-time analysis, further suggests that denasalization is a sound change in progress. This turns our attention to the diachronic change of the nasal series. Denasalization has been reported in a few languages as an ongoing or established sound change, such as the ones reviewed in Y. S. Kim (2011, p. 44). In Karitiana spoken in Brazil, nasal stops are produced as nasal stops by older speakers, but as oral stops by young speakers in word-initial position followed by an oral vowel, and some form of denasalization is also realized in non-initial position (Storto, 1999; Storto & Demolin, 2012). In a number of Chinese dialects, various degrees of denasalization are observed (Hu, 2007), notably in Southern Min dialects (Norman, 1988, p. 235f). What is common to these languages as well as Korean is that they all lack(ed) prevoiced stops in their phonemic inventory. Finally, in some descriptions of several Pacific Northwest languages in North America, nasal stops are also impressionistically transcribed as oral stops or intermediate sounds between oral and nasal stops (Kinkade, 1985; Thompson & Thompson, 1972). One note is of particular interest to us: The phenomenon of denasalization in Pacific Northwest languages is not spread to the varieties that already have an oral voiced stop series (Kinkade, 1985, p. 480). In summary, it seems highly plausible that denasalization of the nasal stops fits better in a phonological system which lacks prevoiced stops, thus not threatening any paradigmatic contrast.

It is also likely that denasalization is more common than generally described, but remains underdocumented crosslinguistically. One of the reasons is that native linguists are ‘deaf’ to it due to their top-down biases, as argued in Y. S. Kim (2011), and acoustic evidence is usually not sufficient to account for nasal properties. Nonetheless, by observing how nasal/denasalized stops interact with other stops in different contexts and positions, we may gain insight into how a language may reach equilibrium between prosodic/positional variations and paradigmatic contrast.

In conclusion, the four stop series in Korean illustrate dynamic clashes between articulatory constraints, perceptual biases, and language-specific syntagmatic and paradigmatic structures, resulting in an ongoing reorganization.

### Additional files

An Open Science Framework project page (<https://doi.org/10.17605/OSF.IO/QUZK4>) is dedicated to this study. We additionally provide the following files:

- **Appendix A.** Additional plots for phrase-initial nasals for Section 3.1. DOI: <https://doi.org/10.5334/labphon.277.s1>
- **Appendix B.** Summaries of the linear mixed-effect models used in Sections 3.2 to 3.4. DOI: <https://doi.org/10.5334/labphon.277.s2>
- **Appendix C.** Complexity parameter table and confusion matrix of the classification tree analysis used in Section 3.5. DOI: <https://doi.org/10.5334/labphon.277.s3>
- **Supplementary Material.** A ZIP file containing the datafiles and the R script to reproduce the plots and statistics. DOI: <https://doi.org/10.5334/labphon.277.s4>

## Acknowledgements

This research was made possible by a Grant-in-Aid (Kakenhi) 17F17006 to the third and first authors by Japan Society for the Promotion of Science (JSPS), a JSPS postdoctoral fellowship to the first author, and a MEXT (Monbukagakusho) scholarship to the second author. We thank anonymous reviewers for constructive comments on the contextualization of the study and requests for clarification, Lisa Davidson (Associate Editor) for thoughtful suggestions on various aspects including sentence structure and word choices, Donna Erickson for proofreading and feedback, and James Kirby for suggestions on the re-organization of an earlier draft and helpful discussions on relevant topics.

## Competing Interests

The authors have no competing interests to declare.

## Author Contributions

The first and second authors contributed to the conceptualization, design, and execution of the experiment, as well as statistical modeling. The second author performed the processing and analysis of the data. The first author drafted and revised the manuscript, with the contribution of the second author. The third author supervised the study. All authors revised and approved the final version for submission.

## References

- Abramson, A. S., & Erickson, D. M.** (1992). Tone Splits and Voicing Shifts in Thai: Phonetic Plausibility. *Haskins Laboratories Status Report on Speech Research, SR-109/110* (pp. 255–262).
- Ahn, H.** (1999). *Post-release phonatory processes in English and Korean: Acoustic correlates and implications for Korean phonology* (Unpublished doctoral dissertation). University of Texas.
- Ahn, M.-J.** (2013). Acoustic duration of Korean nasals. *Studies in Phonetics, Phonology, and Morphology, 19*(3), 411–431. DOI: <https://doi.org/10.17959/sppm.2013.19.3.411>
- Ahn, M.-J.** (2019). Effects of F0 and VOT on the identification of Korean wordinitial oral and nasal stops. *The Journal of Linguistics Science, 90*, 203–218. DOI: <https://doi.org/10.21296/jls.2019.9.90.203>
- Arai, T.** (2006). Cue parsing between nasality and breathiness in speech perception. *Acoustical Science and Technology, 27*(5), 298–301. DOI: <https://doi.org/10.1250/ast.27.298>
- Bang, H. Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T. J.** (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics, 66*, 120–144. DOI: <https://doi.org/10.1016/j.wocn.2017.09.005>
- Boersma, P., & Weenink, D.** (2020). Praat: Doing phonetics by computer [computer program] [Computer software manual]. Retrieved from <http://www.praat.org/> (version 6.1.09).
- Brunelle, M.** (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics, 37*(1), 79–96. DOI: <https://doi.org/10.1016/j.wocn.2008.09.003>
- Burton, M. W., Blumstein, S. E., & Stevens, K. N.** (1992). A phonetic analysis of prenasalized stops in Moru. *Journal of Phonetics, 20*, 127–142. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30243-8](https://doi.org/10.1016/S0095-4470(19)30243-8)
- Chen, M., & Clumeck, H.** (1975). Denasalization in Korean: A search for universals. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a symposium on nasals and nasalization* (pp. 125–131). Stanford University Linguistics Department.

- Chen, Y., & Gussenhoven, C.** (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36, 724–746. DOI: <https://doi.org/10.1016/j.wocn.2008.06.003>
- Childers, D. G., Hicks, D. M., Moore, G. P., & Alsaka, Y. A.** (1986). A model for vocal fold vibratory motion, contact area, and the electroglottogram. *The Journal of the Acoustical Society of America*, 80(5), 1309–1320. DOI: <https://doi.org/10.1121/1.394382>
- Childers, D. G., Hicks, D. M., Moore, G. P., Eskenazi, L., & Lalwani, A. L.** (1990). Electroglottography and vocal fold physiology. *Journal of Speech, Language, and Hearing Research*, 33(2), 245–254. DOI: <https://doi.org/10.1044/jshr.3302.245>
- Cho, T., Jun, S.-A., & Ladefoged, P.** (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193–228. DOI: <https://doi.org/10.1006/jpho.2001.0153>
- Cho, T., & Keating, P. A.** (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*. DOI: <https://doi.org/10.1006/jpho.2001.0131>
- Cho, T., & McQueen, J. M.** (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33, 121–157. DOI: <https://doi.org/10.1016/j.wocn.2005.01.001>
- Choi, J., Kim, S., & Cho, T.** (2020). An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account. *PLoS ONE*. DOI: <https://doi.org/10.1371/journal.pone.0240682>
- Eager, C. D.** (2017). ‘standardize’ package [Computer software manual]. Retrieved from <https://github.com/CDEager/standardize> (R package version 0.2.1).
- Fougeron, C.** (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29(2), 109–135. DOI: <https://doi.org/10.1006/jpho.2000.0114>
- Fougeron, C., & Keating, P. A.** (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, 101(6), 3728–3740. DOI: <https://doi.org/10.1121/1.418332>
- Francis, A. L., & Nusbaum, H. C.** (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 349–366. DOI: <https://doi.org/10.1037/0096-1523.28.2.349>
- Gao, J., & Arai, T.** (2019). Plosive (de-)voicing and f<sub>0</sub> perturbations in Tokyo Japanese: Positional variation, cue enhancement, and contrast recovery. *Journal of Phonetics*, 77. DOI: <https://doi.org/10.1016/j.wocn.2019.100932>
- Gao, J., Hallé, P., & Draxler, C.** (2020). Breathy voice and low-register: A case of trading relation in Shanghai Chinese tone perception? *Language and Speech*, 63(3), 582–607. DOI: <https://doi.org/10.1177/0023830919873080>
- Gao, J., Yun, J., & Arai, T.** (2019). VOT-F<sub>0</sub> coarticulation in Japanese: Production-biased or misparsing? *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*, 0619.1–5.
- Garellek, M., Ritchart, A., & Kuang, J.** (2016). Breathy voice during nasality: A crosslinguistic study. *Journal of Phonetics*, 59, 110–121. DOI: <https://doi.org/10.1016/j.wocn.2016.09.001>
- Georgeton, L., & Fougeron, C.** (2014). Domain-initial strengthening on French vowels and phonological contrasts: Evidence from lip articulation and spectral variation. *Journal of Phonetics*, 44(1), 83–95. DOI: <https://doi.org/10.1016/j.wocn.2014.02.006>
- Henrich, N., d’Alessandro, C., Doval, B., & Castellengo, M.** (2005). Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal

- intensity, and fundamental frequency. *The Journal of the Acoustical Society of America*, 117(3), 1417–1430. DOI: <https://doi.org/10.1121/1.1850031>
- Hirose, H., Lee, C., & Ushijima, T.** (1974). Laryngeal control in Korean stop production. *Journal of Phonetics*, 2(2), 145–152. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31189-1](https://doi.org/10.1016/S0095-4470(19)31189-1)
- Howard, D. M.** (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice*, 9(2), 163–172. DOI: [https://doi.org/10.1016/S0892-1997\(05\)80250-4](https://doi.org/10.1016/S0892-1997(05)80250-4)
- Howard, D. M., Lindsey, G. A., & Allen, B.** (1990). Toward the quantification of vocal efficiency. *Journal of Voice*, 4(3), 205–212. DOI: [https://doi.org/10.1016/S0892-1997\(05\)80015-3](https://doi.org/10.1016/S0892-1997(05)80015-3)
- Hsu, C. H., Evans, J. P., & Lee, C. Y.** (2015). Brain responses to spoken F0 changes: Is H special? *Journal of Phonetics*, 51, 82–92. DOI: <https://doi.org/10.1016/j.wocn.2015.02.003>
- Hu, F.** (2007). Post-oralized nasal consonants in Chinese Dialects — Aerodynamic and acoustic data. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)*, 1405–1408.
- Jun, S.-A.** (1993). *The phonetics and phonology of Korean prosody* (Doctoral dissertation, The Ohio State University). DOI: <https://doi.org/10.4324/9780429454943>
- Jun, S.-A.** (1994). The status of the lenis stop voicing rule in Korean. In Y.-K. Kim-Renaud (Ed.), *Theoretical Issues in Korean Linguistics* (pp. 101–114). Center for the Study of Language.
- Jun, S.-A.** (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15(2), 189–226. DOI: <https://doi.org/10.1017/S0952675798003571>
- Kang, K.-H., & Guion, S. G.** (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, 124(6), 3909–3917. DOI: <https://doi.org/10.1121/1.2988292>
- Kang, Y.** (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45(1), 76–90. DOI: <https://doi.org/10.1016/j.wocn.2014.03.005>
- Kim, H., Maeda, S., Honda, K., & Crevier-Buchman, L.** (2018). The Mechanism and Representation of Korean Three-Way Phonation Contrast: External Photoglottography, Intra-Oral Air Pressure, Airflow, and Acoustic Data. *Phonetica*, 75, 57–84. DOI: <https://doi.org/10.1159/000479589>
- Kim, M.** (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates. In *Proceedings of Interspeech 2004 – 8th International Conference on Spoken Language Processing*.
- Kim, M.-R.** (2014). Ongoing sound change in the stop system of Korean: A three- to two-way categorization. *Studies in Phonetics, Phonology, and Morphology*, 20(1), 51–82. DOI: <https://doi.org/10.17959/sppm.2014.20.1.51>
- Kim, M.-R., Beddor, P. S., & Horrocks, J.** (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics*, 30(1), 77–100. DOI: <https://doi.org/10.1006/jpho.2001.0152>
- Kim, Y. S.** (2011). *An acoustic, aerodynamic and perceptual investigation of word-initial denasalization in Korean* (Unpublished doctoral dissertation). University College London.
- Kingston, J., & Diehl, R. L.** (1994). Phonetic Knowledge. *Language*, 70(3), 419–454. DOI: <https://doi.org/10.2307/416481>
- Kinkade, M. D.** (1985). More on nasal loss on the Northwest Coast. *International Journal of American Linguistics*, 51(4), 478–480. DOI: <https://doi.org/10.1086/465939>

- Kirby, J.** (2013). The role of probabilistic enhancement in phonologization. In A. Yu (Ed.), *Origins of sound change: Approaches to phonologization* (pp. 228–246). Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199573745.003.0011>
- Kirby, J.** (2017). Praatdet: Praat-based tools for EGG analysis [Computer software manual]. Retrieved from <https://github.com/kirbyj/praatdet> (version 0.1.1).
- Kirby, J., & Sonderegger, M.** (2018). Mixed-effects design analysis for experimental phonetics. *Journal of Phonetics*, 70, 70–85. DOI: <https://doi.org/10.1016/j.wocn.2018.05.005>
- Kuang, J., & Liberman, M.** (2018). Integrating voice quality cues in the pitch perception of speech and non-speech utterances. *Frontiers in Psychology*. DOI: <https://doi.org/10.3389/fpsyg.2018.02147>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B.** (2017). lmerTest package: Tests in linear mixed effects models. *Journal of statistical software*, 82(13). DOI: <https://doi.org/10.18637/jss.v082.i13>
- Lee, H., Holliday, J. J., & Kong, E. J.** (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*. DOI: <https://doi.org/10.1111/lnc3.12374>
- Lee, H., & Jongman, A.** (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungang dialects. *Journal of the International Phonetic Association*, 42(2), 145–169. DOI: <https://doi.org/10.1017/S0025100312000035>
- Lee, H., Politzer-Ahles, S., & Jongman, A.** (2013). Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics*, 41(2), 117–132. DOI: <https://doi.org/10.1016/j.wocn.2012.12.002>
- Lenth, R.** (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=emmeans> (R package version 1.4.3.01).
- Lindblom, B., Guion, S., Hura, S., Moon, S.-J., & Willerman, R.** (1995). Is sound change adaptive? *Rivista di Linguistica*, 7, 5–36.
- Matisoff, J. A.** (1975). Rhinoglottophilia: the mysterious connection between nasality and glottality. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a symposium on nasals and nasalization* (pp. 265–287). Stanford University Linguistics Department.
- Mazaudon, M.** (1978). Consonantal Mutation and Tonal Split in the Tamang Sub-family of Tibeto-Burman. *Kailash*, 6(3), 157–179. Retrieved from <https://www.repository.cam.ac.uk/handle/1810/227328>
- Milborrow, S.** (2019). ‘rpart.plot’ package [Computer software manual]. Retrieved from <https://cran.r-project.org/web/packages=rpart.plot> (R package version 3.0.8).
- Norman, J.** (1988). *Chinese*. Cambridge: Cambridge University Press.
- Oh, M., Byrd, D., Louis, G., & Narayanan, S. S.** (2020). Velum-oral timing and its variability in Korean nasal consonants. In *Poster presented at The 12th International Seminar on Speech Production*. Retrieved from [https://issp2020.yale.edu/S08/oh\\_08\\_13\\_135\\_poster.pdf](https://issp2020.yale.edu/S08/oh_08_13_135_poster.pdf)
- Ohala, J. J.** (1975). Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a symposium on nasals and nasalization* (pp. 289–316). Stanford University Linguistics Department.
- Ohala, J. J., & Busà, M. G.** (1995). Nasal loss before voiceless fricatives: A perceptual-based sound change. *Rivista di Linguistica*, 7, 125–144.

- Pittayaporn, P., & Kirby, J.** (2016). Laryngeal contrasts in the Tai dialect of Cao Bằng. *Journal of the International Phonetic Association*, 47(1), 65–85. DOI: <https://doi.org/10.1017/S0025100316000293>
- R Core Team.** (2019). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project.org/> (R version 3.6.2).
- Schertz, J., Kang, Y., & Han, S.** (2019). Sources of variability in phonetic perception: The joint influence of listener and talker characteristics on perception of the Korean stop contrast. *Laboratory Phonology*, 10(1). DOI: <https://doi.org/10.5334/labphon.67>
- Silva, D. J.** (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23(2), 287–308. DOI: <https://doi.org/10.1017/S0952675706000911>
- Solé, M. J.** (2018). Articulatory adjustments in initial voiced stops in Spanish, French and English. *Journal of Phonetics*, 66, 217–241. DOI: <https://doi.org/10.1016/j.wocn.2017.10.002>
- Stevens, K. N., & Keyser, S. J.** (1989). Primary features and their enhancement in consonants. *Language*, 65, 81–106. DOI: <https://doi.org/10.2307/414843>
- Storto, L. R.** (1999). *Aspects of a Karitiana grammar* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Storto, L. R., & Demolin, D.** (2012). The phonetics and phonology of South American languages. In L. Campbell & V. Grondona (Eds.), *The Indigenous Languages of South America* (pp. 331–390). DOI: <https://doi.org/10.1515/9783110258035.331>
- Tabain, M., Butcher, A., Breen, G., & Beare, R.** (2016). An acoustic study of nasal consonants in three Central Australian languages. *The Journal of the Acoustical Society of America*, 139, 890–903. DOI: <https://doi.org/10.1121/1.4941659>
- Therneau, T., Atkinson, B., & Ripley, B.** (2019). ‘rpart’ package [Computer software manual]. Retrieved from <https://cran.r-project.org/web/packages=rpart> (R package version 4.1-15).
- Thompson, L. C., & Thompson, M. T.** (1972). Language universals, nasals, and the Northwest coast. In M. E. Smith (Ed.), *Studies in Linguistics in Honor of George L. Trager* (pp. 441–456). The Hague: Mouton.
- van Alphen, P. M., & Smits, R.** (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics*, 32(4), 455–491. DOI: <https://doi.org/10.1016/j.wocn.2004.05.001>
- Yoo, K.** (2015). Domain-initial denasalisation in Busan Korean: A cross-generational case study. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, 0845.1–5.
- Yoo, K.** (2016). Can phonetically denasalised nasals trigger Korean regressive nasal assimilation? In *Poster presented at the BAAP Colloquium 2016*.
- Yoo, K.** (2018). The role of domain-initial denasalisation in prosodic parsing in Seoul, Busan and Ulsan Korean. In *Poster presented at the BAAP Colloquium 2018*.
- Yoo, K., & Nolan, F.** (2020). Sampling the progression of domain-initial denasalization in Seoul Korean. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11(1): 22. DOI: <https://doi.org/10.5334/labphon.203>
- Yoshida, K.** (2008). Phonetic implementation of Korean “denasalization” and its variation related to prosody. *IULC Working Papers*, 8(1).
- Yu, H. J.** (2018). Tonal development and voice quality in the stops of Seoul Korean. *Phonetics and Speech Sciences*, 10(4), 91–99. DOI: <https://doi.org/10.13064/KSSS.2018.10.4.091>

- Yun, J., & Arai, T.** (2020). Perception of synthesized /na/ by Korean listeners. *Acoustical Science and Technology*, 41(2), 501–512. DOI: <https://doi.org/10.1250/ast.41.501>
- Yun, J., Wong, J. W. S., Moore, J., Iwakami, J., Hui, C. T. J., Gao, J., & Arai, T.** (2020). Korean and Japanese listeners' perception of /ma/ produced by a mechanical vocal-tract model. *Acoustical Science and Technology*, 41(4), 697–700. DOI: <https://doi.org/10.1250/ast.41.697>

**How to cite this article:** Gao, J., Yun, J., & Arai, T. 2021 Korean laryngeal contrast revisited: An electroglottographic study on denasalized and oral stops. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 12(1):7, pp.1–29. DOI: <https://doi.org/10.5334/labphon.277>

**Submitted:** 03 June 2020      **Accepted:** 09 February 2021      **Published:** 19 March 2021

**Copyright:** © 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 