Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

The function of creaky voice in South Korean: A perception study

Patrik Hrabánek, Michaela Watkins, Silke Hamann

University of Amsterdam, The Netherlands

patrik.hrabanek99@gmail.com, m.m.watkins@uva.nl, silke.hamann@uva.nl

Abstract

This study examines whether creaky voice serves as a perceptual cue for the fortis-lenis word-initial stop distinction in Standard Seoul Korean. Twenty-nine native speakers completed an online forced-choice ABX task, where they determined whether a manipulated, artificially creaky lenis token (Sound X) with ambiguous values of VOT and F0 resembled more closely a natural fortis (Sound A) or a natural lenis token (Sound B). Results showed considerable inter-speaker variation: some rarely categorized creaky lenis tokens as fortis, while others did so up to 15% of the time. Statistical analysis showed that presence of creak in Sound X did not change the perceptual classification significantly, suggesting creak is not a primary cue alongside VOT and F0. Word-specific effects also emerged, highlighting the complexity of perceptual cues in Korean stop contrasts.

Index Terms: voice quality, creaky voice, Seoul Korean, plosives, speech perception

1. Introduction

Korean has a typologically uncommon three-way laryngeal contrast among word-initial stop consonants. In Standard Seoul Korean (SSK), these stops are aspirated, fortis (also called tense or constricted), and lenis (or lax) [1]. All three are voiceless in word as well as phrase-initial position.

Table 1: *Overview of Korean stops.*

Labial	Aspirated	Fortis	Lenis
Labial	ㅍ /pʰ/	ㅃ /p*/	ㅂ /p/
Coronal	ㅌ /tʰ/	ㄸ /t*/	ㄷ /t/
Velar	ㅋ /kʰ/	ㄲ /k*/	ㄱ /k/

Phonetically, these stops differ in multiple ways. Early studies emphasized the difference in voice-onset time (VOT), with fortis stops having the shortest, aspirated stops the longest (i.e., strong aspiration), and lenis stops in-between values (weak aspiration) [2, 3, 4]. Later research found that younger speakers (born post-1965) enlarged the VOT values for lenis stops, leading to an overlap with aspirated stops [5, 6]. These speakers primarily distinguish aspirated from lenis stops through differences in fundamental frequency (F0) of the following vowel, with aspirated and fortis stops having a high F0 in the vowel, and lenis stops having a low F0 [4, 7, 8, 9]. This shift suggests a (quasi-)tonogenetic change in SSK [10, 11, 12], though debate remains on whether it is true transphonologization or an effect of prosodic prominence [11]. Perception studies confirm that listeners of all ages prioritize F0 over VOT for the lenis-aspirated contrast [13, 14, 15].

Studies on glottal behavior reveal that aspirated and lenis stops exhibit a wider glottal aperture than fortis stops, which feature nearly complete vocal fold contact [2, 7, 8]. These studies also link the fortis stop to glottal tension, often leading to creaky voice in the following vowel. This role of voice quality has been attested in several acoustic studies: vowels following tense stops in SSK often display a creaky voice quality, whereas lax and aspirated stops often show breathy voice [1, 16, 17].

Creak following fortis stops in SSK differs from 'prototypical creak' [18] - such as that in Mandarin third tone [19, 20] - as it does not seem to be caused by a low F0: fortis stops consistently instigate higher F0 than lenis stops [6, 8, 10]. However, post-fortis vowels in SSK exhibit multiple pulsed voice [21], where alternating pulse lengths and amplitudes create two simultaneous periodicities [18, 19]. This manifests as two F0s, one low and one about an octave higher, persisting through part of the vowel before fading into modal phonation [21]. Such multiple pulsed voice can cause issues for pitch trackers relying on raw autocorrelation, potentially leading to pitch jumps. However, [21] suggest that rather than being tracking errors, these jumps can naturally occur in vowels following SSK fortis stops. By setting octave jump cost in Praat [22] to zero, such jumps can be visualised in the F0 contour. In an analysis of fortis tokens by female speakers in the Seoul Corpus [23], [21] found that pitch jumps upward occur frequently and roughly 1/3 of the way into the vowel in SSK, demonstrating a transition from creaky to modal voice.

A new study by [24] expanded the original dataset to include fortis stops by male speakers. The procedure remained the same as in [21] and resulted in a total of 943 fortis tokens, of which 555 were by male and 388 by female speakers. The results found in [24] demonstrated that 36% of all female tokens contained a jump compared to 12% of male tokens, noting a considerable gender difference. Male speakers, if creaky voice was audibly present, appeared to maintain creak throughout and thus show a consistent low F0 contour with no jumps, whereas female speakers show the creaky to modality switch more frequently, which was picked up by the F0 ratio as in [21]. This finding by [24] suggests that pitch jumps are to some extent systematic in SSK, and important for identifying a voice quality transition.

However, it is unclear whether creak has any perceptual significance in SSK. Perception studies on Korean have given voice quality only limited attention [13, 25]. No prior perception studies have explicitly accounted for creaky voice in stimuli, though [25] noted low H1-H2 values in the fortis vowel of their stimuli, suggesting possible creak.

Anecdotal evidence suggests occasional fortis-lenis confusion. For example, in Yun et al.'s speech corpus [23], stop type labeling was sometimes found to be inconsistent between ortho-

 10.21437/Interspeech.2025-2179

graphic and phonemic tiers, with lenis at times being labeled as fortis phonemically. Similarly, [25] found fortis-lenis ambiguity at low VOT values, while fortis and aspirated stops remained distinct. If creaky voice is unique to fortis stops, could its presence enhance fortis-lenis discrimination? In order to respond to this inquiry, the present study aims to answer the following research question:

Does the presence of creak influence fortis-lenis stop type discrimination for South Korean listeners?

We hypothesize that creaky voice, as an exclusive feature of SSK fortis stops, will reinforce a fortis percept and potentially override other cues, particularly when VOT and F0 are ambiguous.

2. Method

Using a forced-choice ABX listening experiment, we investigate whether a stop is perceived as fortis or lenis when creak is present. In this task, listeners hear three sounds in succession (A, B, X) and must determine whether Sound X more closely resembles A or B. An ABX design was chosen as it does not require explicit use of orthographic labels and therefore does not draw participants' attention to the phonological categories in question.

2.1. Stimuli

To minimize semantic influence, nonce word stimuli were used. These were constructed according to Korean phonology, with each word being two syllables long. Words were produced with each stop type (fortis, lenis) at the beginning of a carrier sentence. To control for place of articulation effects [2], only words beginning with a coronal stop (/t*/, /t/) were used as experimental stimuli. The first syllable always ended in /a/, and the second syllable was either /sja/ or /dʑut/. The stimuli were recorded in a sound-attenuated booth with a Neumann TLM103 cardioid condenser microphone at a 44.1 kHz sampling rate by a 24-year-old female native speaker of Standard Seoul Korean. Almost half of the fortis tokens produced by this speaker included natural creak in the first part of the following vowel.

16 lenis (/tasja/, /tadʑut/) and 16 fortis tokens (/t*asja/, /t*adʑut/) were extracted from the recording. For half of the fortis tokens, naturally creaky ones were used, which were perceptually checked by the authors. To minimize the potential influence of the second syllable, one clearly articulated token of /sja/ and /dʑut/ were selected and used, i.e., the second syllable of all experimental tokens was the same /sja/ or /dʑut/. The intensity of all tokens was adjusted to 70 dB. The mean duration and F0 of the first vowel in these tokens was measured in Praat using pitch settings with an octave-jump cost of 0 and an octave cost of 0.1, see Table 2. As can be seen, the mean F0 for lenis tokens differs only marginally from that of the fortis tokens (in being 12-16 Hz lower, depending on the word pair).

These fortis tokens and lenis tokens were used as Sounds A and B. For Sound X, the same lenis tokens underwent further manipulations. For them, VOT was set to an ambiguous 40 ms by cutting out the middle section of the turbulent noise at zero crossings. These values were based on reports of lenis stops having VOTs of 30-70 ms and fortis stops of 10-50 ms, depending on place of articulation and sentence context [3, 14, 26]. F0 of the /a/ vowel remained unaltered as it was fairly ambiguous.

In half of the lenis tokens for Sound X, artificial creak was created. This was done in two steps. First, a copy of the sound file was created and in its sound wave, every second periodical

Table 2: *Mean vowel duration and F0 of /a/ in non-manipulated sound A (/t*a/) and sound B (/ta/) tokens per word.*

	/t*asja/	/tasja/	/t*adʑut/	/tadʑut/
Duration (ms)				
Mean	117	104	124	109
SD	4	7	7	10
F0 (Hz)				
Mean	193	181	199	183
SD	29	6	20	8

cycle of the vowel from its onset up to the fourth even cycle (typically around 50-60% of the vowel's duration) was set to zero. Then, the amplitude of the original sound file was set to 60% of its original level, and the amplitude of the copy was set to 40% of its original level, and the two sounds were added onto each other. Since these percentages add up to 100%, the resulting sound file matched the original sound file in all areas except where the amplitude was set to zero in the copy. In these zero-amplitude sections, only the original sound file at 60% of its original amplitude was included, resulting in an amplitude-modulated sound wave. See Figure 1 for an example illustration. The resulting sound files contained audible creak (as attested by the authors).

Each experimental ABX set contained as Sound A a natural fortis token (with natural creak present in 50% of the tokens), as Sound B a natural lenis token, and as Sound X a manipulated lenis token with ambiguous VOT and F0, and artificial creak in 50% of the times. Two nonce words were used (tasja, tadʑut), and sets were presented in both ABX and reversed BAX order, yielding a total of 64 experimental sets (8 sets * 2 conditions * 2 words * 2 orders).

Additionally, 128 filler sets were included, from recorded tokens of all three stop types with varying places of articulation. In filler sets, the first syllable of A and B differed in place or

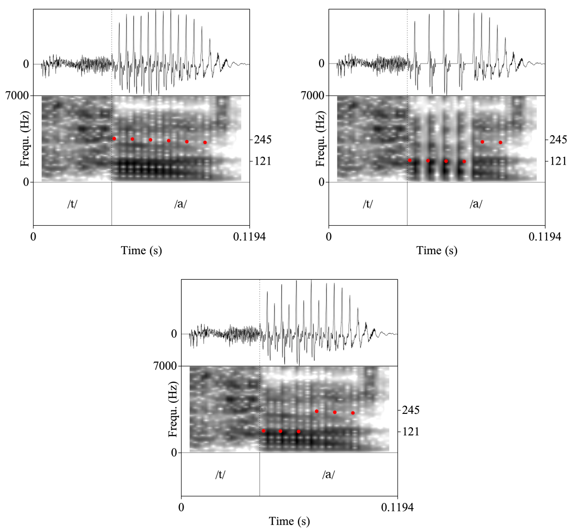

Figure 1: *Example lenis stop syllable /ta/ with original audio file (top left) and a copy with every other cycle of the vowel portion set to zero (top right). The resulting audio file (bottom) from adding the original sound file with 60% amplitude and the copy with 40% amplitude.*

manner of articulation (e. g., /kaju/ - /paju/). 32 of the filler sets were catch trials where A and B were clearly different words, and one was always a real Korean word (e. g., /kasja/ - /nanɯn/).

The sounds within a set had an inter-stimulus interval of 900 ms.

2.2. Participants

A total of 29 native Korean speakers (21 female, 8 male, mean age = 33.66 years, SD = 8.35) were recruited via *Prolific* [27]. At the beginning of the experiment, participants answered demographic questions, including age, gender, and region of origin (Seoul, Gyeonggi, Other). Since the Seoul and Gyeonggi dialects form a continuous speech area and have long been influenced by the social and cultural dominance of the capital, no significant differences were expected between participants from these two regions. The remaining Korean dialects were grouped into the Other category as some differences in perception might have been expected, however, very few participants belonged to this category. Our three-way distinction was thus made to exploratively examine potential dialectal differences rather than address a known disparity.

2.3. Experimental design

The experiment was set-up with Experiment Designer (ED; [28]) and conducted online. Participants were instructed as follows: if Sound X resembled the first sound more, they had to press the key 'Z'; if it resembled the second sound more, they had to press 'M'. They were advised to wear headphones to minimize background noise. A training phase with five filler sets preceded the main task.

The experiment consisted of 192 sets (64 experimental + 128 filler). Each sound was played in stereo. A fixation cross appeared before each set, and response instructions were only displayed after all three sounds had played. Participants had 4 seconds to respond before the next set began. Set order was randomized per participant. A self-timed break was included at the midpoint. The experiment lasted approximately 20 minutes, depending on break duration.

The experiment was approved by the ethics committee of the University of Amsterdam. All participants provided informed consent before beginning the experiment, and were informed that they could withdraw from the study at any point without penalty. No personally identifiable information was collected, and all data were stored securely and anonymized prior to analysis.

3. Results

All participants scored at least 80% accuracy on catch trials, and therefore the data of no participant had to be excluded. The dataset was analyzed in RStudio [29], and statistical analyses were performed using the lme4 package [30].

A generalized linear mixed-effect model was applied, with participants' responses (fortis or lenis) as the dependent variable (*Choice*), and with *Condition* (presence or absence of artificial creak in Sound X), *FortisCreak* (fortis Sound A contained creak or no creak), *Order* (ABX or BAX), *Gender* (male or female), *Word* (/tasja/ or /tadʑut/), and *Region* (Seoul, Gyeonggi, Other). Interaction effects between *Condition*, *Creak* and all other predictors were included. The random-effects structure was built with random intercepts for *Participant*, but no random slopes were included due to convergence problems. The resulting model structure was:

$$Choice \sim (Condition * FortisCreak) * (Order + Gender + Word + Region) + (1 \mid Participant)$$

The intercept (transformed from log-odds to odds) showed that an average participant's odds of choosing 'fortis' were 14.37 times smaller (95% CI = [9.15, 22.57]; $p < 0.001$) than the odds for a 'lenis' choice. The choices made by each participant are shown in Figure 2.

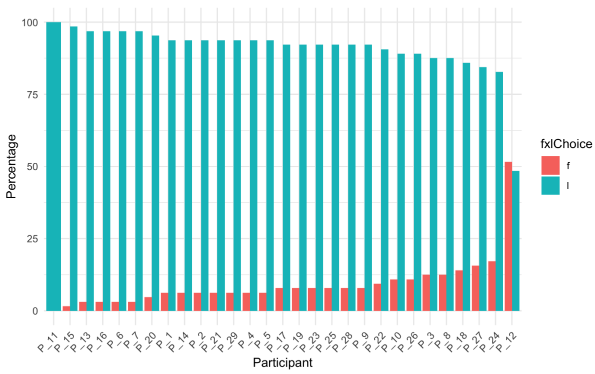

Figure 2: *Percentages of 'fortis' (f) vs. 'lenis' (l) choices for all participants in all trials, ordered from lowest to highest percentage of 'fortis'.*

Every participant made a proportionately higher amount of lenis choices, with fortis choices being quite rare for many listeners, although three participants consistently made fortis choices more than 10% of the time. There was a statistically significant effect of *Word*, whereby the word /tadʑut/ instigated higher odds of making a fortis choice than the word /tasja/ (estimated odds ratio = 2.47, 95% CI = [1.72, 3.53]; $p < 0.001$). Furthermore, there was an interaction effect between *FortisCreak* and *Order* (estimated odds ratio = 2.58, 95% CI = [1.30, 5.12]; $p < 0.006$), indicating that fortis choices were more likely when the fortis token included natural creak and was presented second ($BA_{creak}X$ set). No other effects or interactions reached statistical significance (all $p > 0.05$).

Crucially, Condition was non-significant, suggesting participants' performance remained fairly stable regardless of whether Sound X was manipulated to include creak or not.

4. Discussion

Our results indicate that the presence of artificial creak in the lenis X sound did not significantly trigger fortis perception of the word-initial stop, fortis choices remained generally low for most speakers.

4.1. Non-significance of Condition

The aforementioned absence of a significant result of Condition could mean that creak is not a robust perceptual cue for stop type distinction, with participants primarily relying on VOT and F0, as supported by prior research (e.g., [14, 31]). Alternatively, our amplitude-based creak manipulation may not have been perceptually salient enough or sufficiently naturalistic, as natural creak involves the interplay of multiple acoustic features together (e. g., inter-cycle differences in amplitude and frequency, jitter, shimmer etc.), though they sounded sufficiently creaky and natural to all authors.

Despite the lack of a significant effect, this finding carries theoretical implications for creak in SSK. While creak cooccurs with fortis stops in production, this does not guarantee its relevance in perception. Our results suggest its role may be secondary.

Inter-speaker variation might also be a factor, with some individuals performing vastly different than others in our task. Notably, some participants (P_27, P_24, P_12) tended to categorize both creaky and non-creaky lenis tokens as fortis more consistently than others, indicating that other mechanisms might be at play for certain speakers, and warranting further study.

4.2. The potential interplay of creak and perceived pitch

Our results contained a significant interaction of FortisCreak and Order, meaning that the presence of creaky voice in the fortis Sound A affected participants' performance towards more frequent fortis choices in the reversed BAX sets. However, this interaction was irrespective of the manipulation of Sound X, meaning that these sets instigated more consistent fortis percepts regardless of the presence of creak in X.

Creaky voice can influence perceived F0, as suggested by prior research on voice quality and pitch perception (e. g., [19, 20, 32, 33]). Certain subtypes of creaky voice are associated with lower perceived pitch ([34] on Mandarin and Black Miao; [35] on Cantonese) and even lower produced F0 in pitch-imitation tasks ([20] on Mandarin). While much of this research focuses on Mandarin, the effect appears crosslinguistic [20]. Given creak's link to low F0, listeners might naturally associate creak with low pitch, potentially as part of a natural psychoacoustic mechanism [33], at least for speakers of languages where creaky voice occurs.

The amplitude manipulation we used to create creaky lenis tokens resembled that of Huang [20], whereby the amplitude of every other glottal cycle was decreased to 60% (ratio of ∼1/1.67) of its original amplitude. Huang's study showed that at comparable levels (between the ratios of 1/1.6 and 1/1.8) there was a noticeable increase of responses corresponding to a low tone, as opposed to a nearly consistent high tone response at a ratio of 1/1.4. Since both Mandarin and English speakers exhibited this effect, it is likely our modulation lowered perceived F0, promoting lenis choices, as lenis stops are followed by lower-F0 vowels (e. g., [13, 14]). Due to these reasons, we can hypothesize that amplitude modulation might have had an effect on the perceived F0.

If creak indeed lowered perceived vowel pitch, it would favor lenis perception. Therefore, it is unclear why our participants were more likely to favor a fortis percept in these sets, especially given the fact that every X stimulus contained a lenis burst.

4.3. Word-specific effects

Though not expected, a statistically significant difference in performance between our two words was found, with /tadʑut/ resulting in more fortis categorizations than /tasja/. This could have been caused by slight differences in duration of the /a/ vowel of the first syllable in our data. The word /tadʑut/ had a slightly longer vowel duration, which would aid the categorization of a fortis vowel, since post-fortis vowels in our data were found to have longer durations than post-lenis vowels, cf. Table 2. An effect of intonation of the second syllable can be excluded as potential explanation, as /dʑut/ did not have higher pitch than /sja/ or a tonal pattern different from /sja/ that reinforced fortis perception.

4.4. Limitations of the present study

The exploratory nature of this study brings inherent limitations. First, since our ABX task omitted explicit labels, fortis choices should be interpreted as judgments that creaky voice aligns more with fortis than lenis rather than direct category selection. Providing orthographic labels might have affected participants' performance due to possible preconceived notions of stop types. Furthermore, the absence of fine-grained VOT manipulation limited our ability to assess creak's strength as a cue across varying VOT values. Additionally, despite standardizing VOT to 40 ms in our X stimuli, the burst remained the naturally occurring lenis burst, with only the middle part of the turbulent noise removed. Finally, duration of the /a/ vowel in our stimuli was not standardized.

5. Conclusion

Our study demonstrates that creaky voice - or its perceptual instigation through amplitude modulation - can alter stop perception in certain contexts. While production studies frequently measure voice quality such as creak via spectral slope (e. g., H1-H2) and other parameters (e. g., Harmonics-to-Noise Ratio (HNR), jitter, flutter, period doubling), perception studies manipulating these parameters remain scarce in research of Korean, but see [13, 25] who cross-spliced vowels following fortis consonants onto non-fortis plosives.

Further research on creak strength is needed, as our study only maintained a fixed 60% amplitude modulation across all tokens. Investigating various levels of amplitude modulation and other modulation types for successful creak perception would further advance our knowledge on the role of creak as perceptual cue.

Our results raise the question whether speakers who relied on creak perceptually also utilize it more in production. A follow-up study analyzing production data from participants who exhibited a stronger overall tendency for fortis perception (e. g., P_27, P_24, P_12) could provide valuable insights into the interaction between perception and production in stop type distinction.

6. Acknowledgments

The authors would like to thank Paul Boersma for devising an efficient method for generating amplitude-reduced cycles, and Dirk Vet from the Speech Lab at the University of Amsterdam for his help with creating and setting up the experimental environment.

7. References

[1] Y. Kang, J. Schertz, and S. Han, "The phonology and phonetics of Korean stop laryngeal contrasts," in *The Cambridge Handbook of Korean Linguistics*, 1st ed., S. Cho and J. Whitman, Eds. Cambridge University Press, Aug. 2022, pp. 215–247. [Online]. Available: https://www.cambridge.org/core/product/identifier/9781108292351%23CN-bp-8/type/book_part

[2] E. Abberton, "Some laryngographic data for Korean stops," *Journal of the International Phonetic Association*, vol. 2, no. 2, pp. 67–78, Dec. 1972. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0025100300000517/type/journal_article

[3] L. Lisker and A. S. Abramson, "A cross-language study of voicing in initial stops: Acoustical measurements," *WORD*, vol. 20, no. 3, pp. 384–422, Jan. 1964. [Online]. Available: http://www.tandfonline.com/doi/full/10.1080/00437956.1964.11659830

[4] R. S. Weitzman and M. S. Han, "Acoustic features in the manner differentiation of Korean stop consonants," *The Journal of the Acoustical Society of America*, vol. 40, no. 5, pp. 1272–1273, Nov. 1966. [Online]. Available: https://pubs.aip.org/jasa/article/40/5_Supplement/1272/684884/Acoustic-Features-in-the-Manner-Differentiation-of

[5] Y. Kang, "Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study," *Journal of Phonetics*, vol. 45, pp. 76–90, Jul. 2014. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0095447014000357

[6] D. J. Silva, "Acoustic evidence for the emergence of tonal contrast in contemporary Korean," *Phonology*, vol. 23, no. 02, pp. 287–308, Aug. 2006. [Online]. Available: http://www.journals.cambridge.org/abstract_S0952675706000911

[7] R. Kagaya, "A fiberscopic and acoustic study of the Korean stops, affricates and fricatives," *Journal of Phonetics*, vol. 2, no. 2, pp. 161–180, Mar. 1974. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S009544701931191X

[8] C.-W. Kim, "On the autonomy of the tensity feature in stop classification (with special reference to Korean stops)," *WORD*, vol. 21, no. 3, pp. 339–359, Dec. 1965. [Online]. Available: http://www.tandfonline.com/doi/full/10.1080/00437956.1965.11435434

[9] H. Kim, S. Maeda, and K. Honda, "Invariant articulatory bases of the features [tense] and [spread glottis] in Korean plosives: New stroboscopic cine-MRI data," *Journal of Phonetics*, vol. 38, no. 1, pp. 90–108, Jan. 2010. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0095447009000205

[10] H.-Y. Bang, M. Sonderegger, Y. Kang, M. Clayards, and T.-J. Yoon, "The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis," *Journal of Phonetics*, vol. 66, pp. 120–144, Jan. 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0095447017302188

[11] S.-A. Jun, "Studies in Korean prosody: Current issues and future prospects," 2020.

[12] X. Ren and P. Mok, "Tonogenesis in Seoul Korean and L3 production of Korean stops by Cantonese- English bilinguals," in *6th International Symposium on Tonal Aspects of Languages (TAL 2018)*. ISCA, Jun. 2018, pp. 192–196. [Online]. Available: https://www.isca-archive.org/tal_2018/ren18_tal.html

[13] K.-H. Kang, "Generational differences in the perception of Korean stops." *Phonetics and Speech Sciences*, vol. 2, no. 3, pp. 3–10, 2010.

[14] M.-R. Kim, P. S. Beddor, and J. Horrocks, "The contribution of consonantal and vocalic information to the perception of Korean initial stops," *Journal of Phonetics*, vol. 30, no. 1, pp. 77–100, Jan. 2002. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0095447001901524

[15] E. J. Kong and H. Lee, "Attentional modulation and individual differences in explaining the changing role of fundamental frequency in Korean laryngeal stop perception," *Language and Speech*, vol. 61, no. 3, pp. 384–408, Sep. 2018. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0023830917729840

[16] T. Cho, S.-A. Jun, and P. Ladefoged, "Acoustic and aerodynamic correlates of Korean stops and fricatives," *Journal of Phonetics*, vol. 30, no. 2, pp. 193–228, Apr. 2002. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0095447001901536

[17] H. J. Yu, "Tonal development and voice quality in the stops of Seoul Korean," *Phonetics and Speech Sciences*, vol. 10, no. 4, pp. 91–99, Dec. 2018. [Online]. Available: http://www.eksss.org/archive/view_article?doi=10.13064/KSSS.2018.10.4.091

[18] P. Keating, M. Garellek, and J. Kreiman, "Acoustic properties of different kinds of creaky voice," in *In Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, Glasgow, 2015, pp. 2–7.

[19] Y. Huang, "Different attributes of creaky voice distinctly affect Mandarin tonal perception," *The Journal of the Acoustical Society of America*, vol. 147, no. 3, pp. 1441–1458, Mar. 2020. [Online]. Available: https://pubs.aip.org/jasa/article/147/3/1441/993966/Different-attributes-of-creaky-voice-distinctly

[20] ——, "Perception and imitation of period doubling," in *Proceedings of the 20th International Congress of Phonetic Sciences – ICPhS 2023*, Prague, 2023.

[21] M. Watkins, P. Boersma, and S. Hamann, "Revisiting pitch jumps: F0 ratio in Seoul Korean," in *Proceedings of Interspeech 2024*, 2024.

[22] P. Boersma and D. Weenink, "Praat: doing phonetics by computer." Mar. 2024. [Online]. Available: https://www.praat.org.

[23] W. Yun, K. Yoon, S. Park, J. Lee, S. Cho, D. Kang, K. Byun, H. Hahn, and J. Kim, "The Korean Corpus of Spontaneous Speech," *Phonetics and Speech Sciences*, vol. 7, no. 2, pp. 103–109, Jun. 2015. [Online]. Available: http://koreascience.or.kr/journal/view.jsp?kj=OMSOCX&py=2015&vnc=v7n2&sp=103

[24] M. Watkins, R. Puggaard-Rode, P. Boersma, and S. Hamann, "Robustness of F0 ratio as a diagnostic: comparing creaky voice in Danish and Seoul Korean," in *Proceedings of Interspeech 2025*, 2025.

[25] J. Schertz, Y. Kang, and S. Han, "Sources of variability in phonetic perception: The joint influence of listener and talker characteristics on perception of the Korean stop contrast," *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 10, no. 1, pp. 1–32, Jul. 2019. [Online]. Available: https://www.journal-labphon.org/article/10.5334/labphon.67/

[26] M.-R. C. Kim and A. Lotto, "An investigation of acoustic characteristics of Korean stops produced by non-heritage learners," *The Korean Language in America*, vol. 7, pp. 177–187, 2002.

[27] Prolific.ac.uk, "Prolific," 2025, last accessed 19/02/2025. [Online]. Available: https://www.prolific.com/

[28] D. J. Vet, "Ed: Experiment designer." 2025. [Online]. Available: https://www.fon.hum.uva.nl/dirk/ed.php

[29] R. Team, "RStudio: Integrated Development for R." Boston, MA, 2022. [Online]. Available: https://cran.rstudio.com

[30] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[31] H. Lee and A. Jongman, "Effects of sound change on the weighting of acoustic cues to the three-way laryngeal stop contrast in Korean: Diachronic and dialectal comparisons," *Language and Speech*, vol. 62, no. 3, pp. 509–530, Sep. 2019. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0023830918786305

[32] J. Kuang, "Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice," *The Journal of the Acoustical Society of America*, vol. 142, no. 3, pp. 1693–1706, Sep. 2017. [Online]. Available: https://pubs.aip.org/jasa/article/142/3/1693/613278/Covariation-between-voice-quality-and-pitch

[33] J. Kuang and M. Liberman, "The effect of vocal fry on pitch perception," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Shanghai: IEEE, Mar. 2016, pp. 5260–5264. [Online]. Available: http://ieeexplore.ieee.org/document/7472681/

[34] J. Kuang, "Phonation in tonal contrasts," PhD thesis, University of California Los Angeles, Los Angeles, CA, 2013.

[35] Y. Zhang and J. Kirby, "The role of F0 and phonation cues in Cantonese low tone perception," *The Journal of the Acoustical Society of America*, vol. 148, no. 1, pp. EL40–EL45, Jul. 2020. [Online]. Available: https://pubs.aip.org/jasa/article/148/1/EL40/966258/The-role-of-F0-and-phonation-cues-in-Cantonese-low