UNIVERSITEIT VAN AMSTERDAM

A perception study on the function of creak in Yanbian Korean

MA Thesis

LI JIN (14830000)

Supervised by Silke Hamann
Second Reader: Michaela Watkins

Master of General Linguistics
Faculty of Arts and Humanities
University of Amsterdam
The Netherlands

June 20, 2025

1. Introduction

1.1 The Korean laryngeal contrast

Korean presents a typologically rare ternary laryngeal distinction within its stop consonants at the start of words. In Standard Seoul Korean (SSK), word-initial stops are categorized into three phonation types: fortis (/p*/, /t*/, /k*/), lenis (/p/, /t/, /k/), and aspirated (/p^h/, /t^h/, /k^h/) (see table 1). Aspirated stops are typically characterized by being strongly or heavily aspirated, whereas lenis stops are identified as plain, lax, weak, or lightly aspirated, and fortis stops are typically referred to as tense, reinforced, unaspirated, or constricted (Kang et al., 2022). Furthermore, in the accentual phrase-initial position, these stops are typically realized as voiceless (Cho et al., 2002).

Table 1. Summary of Korean stop consonants

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

Early studies showed that these three stops can be differentiated based on voice onset time (VOT) is the period from the release of stop consonants to the beginning of vocal fold vibration (Kewley-Port & Preston, 1974). Fortis stops have the shortest voice onset time, while aspirated stops show the longest, and the lenis stops fall in between the two but much closer to fortis stops (Lisker & Abramson, 1964). It was also observed that fundamental frequency (F0) of the vowel following these stops can distinguish between the different laryngeal categories. Fundamental frequency (F0) refers to the rate of the vocal folds vibrate, perceived by listeners as pitch (Skuk & Schweinberger, 2013). Kim (1965) reported that vowel onsets after aspirated and fortis stops exhibit a high F0, while those following lenis stops are marked by a lower F0. In most Korean dialects, including the Standard Seoul Korean (SSK), this phenomenon is phonologically integrated into the intonation system such that phrase-initial syllables show a tonal distinction based on the type of the consonants (Cho et al., 2002). Syllables beginning with fortis or aspirated stops are associated with a high tone, whereas those beginning with lenis stops relate to a low tone. However, F0 was not viewed as a primary cue for distinguishing the stop contrast in the past and received less focus in the existing literature. Han & Weitzman (1970) investigated intensity build-up after the voice onset, it takes comparatively longer for glottal intensity to reach its peak after an aspirated or lenis stop than after a fortis stop. Kagaya (1974) measured the formant transition's duration and found that the following vowels of fortis stops show a 'creaky' sound. Later, researchers initiated H1-H2 measurements at voice onset as another acoustic measure related to creakiness (Han, 1998), which refers to the amplitude distinction between the first and second harmonics. In 1990s, researchers started to examine the interaction between VOT and vocalic cues such as F0, spectral tilt, and intensity of following vowel, in both production and perception of the stop contrast (Lee et al., 2020). Lee & Jung (2000) were the first researchers who manipulated vocalic cues and VOT in direct opposition and found that vocalic cues may have greater perceptual influence than VOT. In the meantime, the system of stop consonants has been going through a diachronic development in some Korean dialects, among younger speakers of Seoul Korean, the difference in VOT between aspirated and lenis stops has been reducing (Silva, 2006), along with this change, younger speakers have also been making the F0 difference more noticeable (Kang, 2014). A similar change in VOT has been observed in Korean speakers from Shenyang (a city in the Northeast of China) (Jin, 2008) and in Jeju speakers (Holliday & Kong, 2011), but no diachronic change happened in Yanbian Korean (Oh & Yang, 2013). This is because the Yanbian dialect shows a lexical pitch accent - a feature that existed in Middle Korean but was later disappeared in the Western dialects (Schertz et al., 2019). In the Yanbian dialect, because F0 is already used for word-level distinction, it may not be available as a phonetic cue for distinguishing the three-way contrast, which could in turn hinder the progression of a potential VOT merger (Schertz et al., 2019).

1.2 Creak and voice quality

Voice quality describes the way vocal folds vibrate during speech. It is a significant feature because in some language it can play a contrastive role by distinguishing between different meanings (Garellek, 2019). In Korean, the vowel onset that follows a fortis stop often shows features of creaky voice (Abberton,1972), and vowels that follows a lenis stop often shows breathy voice features (Han, 1998). Creaky voice, also known as 'creak', is a broad term for various types of irregular pattern of voicing. It usually shows low F0, irregular F0, and constricted glottis (Garellek, 2019). When people hear creaky voice, it is often expressed as 'rough', 'tense', 'pressed' or 'harsh' (Garellek, 2019). The tight closure of vocal folds in creaky voice reduces airflow because they stay close together and have a long period of closure during every vibration (Keating et al., 2023). When the vocal folds are tightly closed (high glottal constriction), it usually leads to low H1-H2 values, which are often linked to creaky voice (Huang, 2023). Among the three key properties (low F0, irregular F0, and constricted glottis), low value of F0 and irregular F0 by itself can be perceived as creaky sound by listeners but constricted glottis by itself did not lead listeners to perceive it as creak (Keating et al., 2023). On the other hand, certain combinations of these properties also lead to the perception of creaky voice, such as low F0 + irregular F0 or low F0 + constricted glottis. In SSK, vowels following fortis stops show multiple pulsed (irregular) voice (Watkins et al., 2024) where shifting longer and shorter pulses result in the presence of two simultaneous periodicities (Keating et al., 2015). This appears as two F0s, one low and one roughly an octave higher, continuing through part of the vowel transitioning into modal phonation (Watkins et al., 2024). Watkins et al. (2024) analysed fortis tokens produced by female speakers within the Seoul corpus, and they observed that upward pitch jumps commonly occur about one-third of the vowel, indicating a shift from creaky to modal voice. Regarding the voice quality in Yanbian Korean, Oh & Yang (2013) stressed that H1-H2, which refers to the gap between the first harmonics (H1) and second vocal harmonics (H2), is a useful cue for identifying the voice quality of stop consonants. According to Cho et al. (2002), the three-way contrast in Seoul Korean shows significant differences in their H1-H2 values. The H1-H2 value on the following vowel is highest after lenis stops, moderate after aspirated stops, and lowest after fortis stops.

1.3 Yanbian Korean

Yanbian Korean, the dialect spoken in China's largest Korean Autonomous Prefecture in Northeast China, is similar to the Hamkyung dialect in North Korea (Ito, 2008). Yanbian Korean features a pitch accent system where each word contains a mandatory, single pitch peak, and stems differ based on where the pitch peak occurs (Ito & Kenstowicz, 2009b). For the majority of words, the location of pitch peak is unpredictable, it occurs either on the last syllable or the penultimate syllable. The former is represented as L.H, and the latter as H.L in the transcription. Similar to SSK, Yanbian Korean also show three types of laryngeal stops: aspirated, lenis, and fortis (Oh & Yang, 2013). In contrast to comprehensive research on Seoul Korean stops, there has been relatively limited investigation into Yanbian Korean. Ito and Kenstowicz (2009a) found that in Yanbian Korean, VOT does not function as a primary cue in differentiating lenis from fortis stops. They observed a slight F0 difference between lenis and fortis stops in high-tone syllables (fortis > lenis), but no such difference was found in low-tone syllables. Furthermore, they concluded that voice quality in the following vowel is the primary cue in discriminating lenis stops from fortis stops, with fortis stops showed a lower H1-H2 value (mean value of -7.1dB) and lenis stops showed a higher H1-H2 value (mean value of -2.8dB) in Yanbian Korean. However, the study was solely based on a single female participant who is in her 30s. Oh and Yang (2013) conducted a comparison of the ternary stop contrast in both Yanbian Korean and Seoul Korean. They found that in Yanbian Korean VOT for lenis and fortis stops has merged for both genders, meaning VOT alone is not enough to distinct lenis and fortis stops. F0 only plays a role for differentiating lenis and fortis in male Yanbian speakers, not in female speakers. Instead, H1-H2 differentiates lenis and fortis stops for both male and female speakers in Yanbian Korean, and the value of the vowel come after a lenis stop shows a markedly higher value than a fortis stop. It is because in a tonal language such as Yanbian Korean F0 carries two phonological roles which are lexical-tonal differences in vowels and laryngeal differences in stops. Thus, F0 cannot be the primary cue for differentiating these stops when a VOT merger has happened between these stops. On the other hand, Kang et al. (2022) showed evidence that Yanbian speakers from Hunchun (a city in Yanbian Korean Autonomous Prefecture) still rely on F0 to differentiate lenis and fortis stops but the difference is reduced among younger speakers and words with low pitch. Kang and Han (2012) examined the production of stop consonants by Hamkyung dialect speakers in Qingdao and concluded that speakers from Hamkyung produced lenis stops more aligned with fortis stops than with aspirated stops.

1.4 Previous perception studies

Hrabanek (2024) and Hrabanek et al. (to appear) conducted perception studies to investigate whether creak is a perceptual cue for discriminating fortis and lenis word-initial stops in SSK. Both studies used a forced-chosen ABX task. In the former study, A is a natural fortis sound with and without creak, B is a natural lenis sound, X is a lenis with artificial creak. In the latter study, the stimuli were similar, however, sound X included both artificially creaky lenis sounds and naturally produced lenis sounds. The stimuli were delivered to listeners in ABX and BAX orders. Moreover, both studies used two different words, each beginning with coronal stops. The first syllable ended in the vowel /a/, while the last (the second) syllable was either /sja/ or /dyʑut/, resulting in the words 다샤, 따샤 (/tasja/, /t*asja/) and 다즛, 따즛 (/tadyʑut/, /t*adyʑut/). To minimize the influence of VOT as a perceptual cue, VOT was set to 40 ms for all tokens using a Praat script (Boersma & Weenink, 2025) in Hrabanek's (2024) study, and in Hrabanek et al.'s (to appear) study, the VOT was changed directly in Praat. In Hrabanek's (2024) study, 33 native Korean speakers participated. Participants were asked whether they had prior experience learning Mandarin, and this variable (Mandarin experience: yes or no) was included as a predictor in their statistical model. The result of statistical model indicated that Mandarin learning experience had no significant effect on participants' responses in the perception task. Overall, both studies found that adding artificial creak sound to the lenis did not significantly lead the word-initial stops to be perceived as fortis, only a few speakers categorised the lenis with artificial creak as fortis stops. Furthermore, for answering whether the presence of creaky sound affects the perception of the lenis-fortis contrast, Hrabanek's (2024) study showed that the findings indicate that the presence of creaky voice clearly affects listener's decision between fortis and lenis stops, with many participants selecting the 'fortis' option over 10% of the time, some participants even over 20% of the time. However, this study only introduced artificially manipulated creak into sound X, without including a control ABX sets which sound X was naturally produced lenis sound. As a result, participants' decision was only based on one ABX condition. The absence of control ABX sets led the author to conclude that the presence of artificial creak influenced the perception of lenis and fortis stops, even though only 10% of the time participants chose fortis over lenis. In contrast, the study by Hrabanek et al. (to appear) included both naturally produced lenis sounds and stimuli with artificial creak for sound X. Their findings indicated that the presence of creak did not significantly influence participants' categorization of lenis and fortis stops, participants predominantly chose lenis regardless of the presence of creak. Moreover, for both studies, participants showed a significantly higher tendency to select the fortis category in the BAX order than in the ABX order because there is a time gap between the 'A' and 'X' stimuli. In both experiments, each stimulus was separated by a 900 ms interval, meaning that the time between 'A' and 'B' as well as between 'B' and 'X' was 900 ms, which made in a total of 1800 ms between 'A' and 'X'. The extended delay increases the possibility that, when participants hear sound 'X', the memory of sound 'A' might have faded away.

1.5 Acoustic study of the intended stimuli

Prior to conducting the perception study, we analysed several acoustic measurements such as VOT, F0, vowel duration, pitch in the first quarter of the vowel (F0_1/4) and the last quarter of the vowel (F0_4/4), H1-H2 in the first quarter (H1-H2_1/4), the last quarter (H1-H2_4/4) of the vowel, and H1-H2_1/4-H1-H2_4/4, based on the production of all stimulus items (three-way contrast). These measurements were taken from the vowel following the relevant plosives. We recorded our stimuli in a soundproof room of the speech lab at University of Amsterdam. A 35-year-old Yanbian female native speaker who was born and raised in Yanbian produced 105 tokens of all experimental items, which include 54 tokens (18 tokens for each of the three stop types) with a high tone and 51 tokens (17 tokens for each of the three stop types) with a low tone. These measurements were calculated in Praat using a Praat script (see appendix 8.1). Table 2 shows the data of all acoustic measurement.

Table 2. Mean values (Mean) and standard deviations (SD) of VOT, F0_average, F0_1/4, F0_4/4, Vowel duration, F0 ratio, H1-H2_1/4, H1-H2 _4/4, and H1-H2_1/4 - H1-H2 _4/4, measured from the vowel in first syllable of the unmanipulated fortis, lenis, and aspirated with High and Low accent

	/kaki/ (H)	/kaki/ (L)	/k*aki/ (H)	/k*aki/ (L)	/kʰaki/ (H)	/kʰaki/ (L)
VOT						
Mean (ms)	25.98	30.70	20.72	24.01	97.76	116.00
SD (ms)	3.46	4.04	4.88	3.50	12.28	12.87
F0_average						
Mean (Hz)	251.90	208.57	252.80	188.09	267.60	198.53
SD (Hz)	3.94	6.47	6.59	7.16	11.67	11.42
F0_1/4						
Mean (Hz)	244.34	183.99	251.96	176.82	261.93	187.72
SD (Hz)	6.40	13.70	6.81	6.00	12.53	8.14
F0_4/4						
Mean (Hz)	252.54	158.95	252.33	157.16	256.51	160.42
SD (Hz)	3.16	10.13	5.98	15.56	11.49	7.84
F0 ratio						
Mean	0.967	1.158	0.998	1.135	1.021	1.171
SD	0.02	0.06	0.01	0.11	0.01	0.44
Vowel duration						
Mean (ms)	350.81	302.30	349.38	288.59	316.67	282.48
SD (ms)	25.32	40.56	30.30	23.98	26.35	28.12
H1-H2_1/4						
Mean (dB)	6.948	6.719	2.202	-0.227	6.991	2.842
SD (dB)	1.32	3.14	0.95	0.55	1.05	1.61
H1-H2_4/4						
Mean (dB)	5.068	3.068	5.016	1.466	4.989	1.921
SD (dB)	0.70	1.18	0.80	1.65	0.68	1.50
H1-H2_1/4 - H1-H2_4/4						
Mean (dB)	1.880	3.651	-2.814	-1.693	2.002	0.921

In our stimuli, VOT for the three-way stop consonants is higher under the low tone condition than in the high tone condition. Among the stop types, the VOT of fortis stops show the shortest, aspirated show the longest, and lenis stops falls in between, but the VOT values of lenis stops align more closely with fortis, with almost no gap between them. These results support previous research (Lisker & Abramson, 1964), which show that VOT values of fortis and lenis stops are similar. Regarding F0, the high tone condition shows higher values than the low tone condition across all stop types. Previous studies (Ito & Kenstowicz, 2009a) have suggested that in Yanbian Korean, F0 values do not differ between fortis and lenis stops in the low tone condition, while in the high tone condition, fortis stops tend to show higher F0 values than lenis stops. However, in our stimuli, mean F0 values in the high tone condition were similar between fortis and lenis stop, whereas in the low tone condition, there was a notable difference, lenis stops had F0 values approximately 20 Hz higher than fortis stops. Meanwhile, F0 in the first quarter of the vowel (F0_1/4) and F0 in the last quarter of the vowel (F0_4/4) were also measured. F0_1/4 was always lower than the average F0 (F0_average), regardless of tone conditions and stop types. However, the comparison between F0_1/4 and F0_4/4 did not reveal a consistent pattern. In low-tone stimuli, F0 exhibit a clearly falling pitch trajectory, with F0_1/4 always higher than F0_4/4. In contrast, in high-tone stimuli, aspirated stops show a decreasing F0, while fortis and lenis stops show an increasing F0 patterns. Specifically, although the mean F0_1/4 value for fortis stops is lower than the mean F0_4/4 value, but approximately half of the tokens show a higher F0_1/4 than F0_4/4. In early studies, Ohde (1984) analysed nonsense syllables produced by American English speakers, both in isolation and within carrier phrases. He observed that the F0 consistently decreased from a high point after voiceless tops, regardless of the stops were aspirated or unaspirated. However, in our study, this pattern was not consistently observed. Regarding vowel duration, the data also revealed that vowels following stops are consistently much longer in high tone contexts than in low tone.

Watkins et al. (2024) proposed the measure of 'F0 ratio' to illustrate how often pitch jumps occur in the tokens. It is calculated by dividing the median F0 in the last third of vowel by the median F0 in the first third of vowel (see formula (1)). In our acoustic analysis, we adjusted the vowel segments from thirds to quarters, as shown in (2). Watkins et al. (2024) emphasized that if the F0 ratio falls between 0.8-1.25, it indicates no pitch jump, and an F0 ratio greater than 1.7 suggests pitch doubling, while a ratio below 0.6 indicates pitch halving (Watkins et al., 2024). In our production data, the F0 ratios for all three laryngeal contrast types fall within the 0.8 to 1.25 range, regardless of tonal condition.

$$\text{(1) F0 ratio} = \frac{\text{median F0 in last 1/3 of vowel}}{\text{median F0 in first 1/3 of vowel}}$$

$$\text{(2) F0 ratio} = \frac{\text{median F0 in last 1/4 of vowel}}{\text{median F0 in first 1/4 of vowel}}$$

Additionally, measurements of H1-H2 in the first quarter (H1-H2_1/4) and the last quarter (H1-H2_4/4) were also taken. H1-H2 represents the amplitude difference between the first and second harmonics (Chai & Garellek, 2022). It is one of the most widely used measure of voice quality in linguistic phonetics and is closely associated with variations in phonation type (breathy vs modal vowels) across many languages. Usually, lower H1-H2 values are typically related to a smaller glottal open quotient (OQ), tighter vocal fold constriction, and greater thickness in the middle part of the vocal folds (Chai & Garellek, 2022). Interestingly, H1-H2 value does not always show the expected pattern. For example, there was no difference in H1-H2 values across phonation types. Kreiman et al. (2012) discovered that even though H1-H2 is strongly connected with the open quotient (OQ), the connection differs significantly between individual speakers. It may be due to the influence of additional factors other than glottal open quotient. Despite some limitations of H1-H2, scholars continue to use it, and Kreiman et al. (2007) provide evidence in support of its reliability. They examined how various acoustic measures (including H1-H2, H2-H4, and spectral slope) relate to spectral and glottal pulse shape. Through correlation and principal component analyses of 78 acoustic parameters, they identified four core independent dimensions, which are H1-H2, H2-H4, overall spectral slope, and noise in high-frequency range. Crucially, they discovered that H1-H2 reflects meaningful variation in spectral and glottal pulse characteristics and remains stable across various models of glottal source, indicating that H1-H2 is a reliable measure, not easily affected by modelling artifacts. To minimise the instability of H1-H2 values, we proposed a new measure H1-H2_1/4 – H1-H2_4/4. For fortis stops, the H1-H2 value in the last quarter of the vowel is always higher than in the first quarter. In contrast, for lenis stops, the H1-H2 value in the last quarter is always lower than in the first quarter. It means that fortis stops (creaky sound) show a negative H1-H2_1/4 – H1-H2_4/4, and lenis stops (breathy sound) show a positive H1-H2_1/4 – H1-H2_4/4. For aspirated stops, the mean value in the first quarter of H1-H2 is always higher than the last quarter of H1-H2 under high tone conditions. However, under low tone conditions, aspirated stops show instability, 45% of them show a lower H1-H2 value in the first quarter compared to the last quarter. The underlying cause of this pattern is unclear.

1.6 The present perception study

The studies by Hrabanek (2024) and Hrabanek et al. (to appear) have concluded that the artificial creaky sound to the lenis stops did not cause listeners to perceive the word-initial stops as fortis in SSK. However, given that voice quality as the key cue for discriminating lenis from fortis stops in Yanbian Korean, could creak function in this dialect as a perceptual cue for the lenis-fortis distinction? The present study therefore aims to address the following question:

Does creak affect the discrimination of fortis and lenis stop types for Yanbian Korean speakers?

We hypothesize that, as creak is a unique feature for fortis stops in Yanbian Korean, it could function as the main cue in perception for differentiating fortis and lenis stops since VOT and F0 can’t be the perceptual cue in Yanbian Korean.

2. Method

The present study uses a forced ABX selection task to examine whether the presence of creak influences the perception of a stop as fortis or lenis. In this experiment, listeners hear three sequential sounds (A, B, and X) and must determine whether sound X is more like sound A or sound B. An ABX task was chosen because it avoids applying any phonological knowledge and does not rely on orthography. Sound manipulation was used for sound X to create artificial creak from a natural lenis sound.

2.1 Stimuli

To limit the impact of semantic involvement, nonce words were used as stimuli and fillers. The selected nonce words were constructed based on Yanbian Korean phonology and each stimulus is a monosyllabic stem followed by the nominative suffix *-i*, resulting in a disyllabic stimulus. In Yanbian Korean, monosyllabic words have two distinct accent classes: H[L] and L[H] (square brackets [] represents the tone on the following suffix). For our data, we only extracted H[L] accent nonce words because the mean F0 values for fortis and lenis stops under the high tone condition were nearly identical in our stimuli (see table 2 in section 1.5), and also because 80% of monosyllabic words are H[L] accented in Yanbian Korean (Ito, 2008). Each nonce word was produced with either a fortis or lenis stop at the start of the word. To minimize place of articulation effects, the experimental stimuli included only words beginning with a velar stop. The first syllable starts with either /k/ or /k*/, followed by a vowel /a/, and always end with /k/ (see table 3). Each filler begins with a nasal (/n/, /m/), fricative (/s/, /h/), or approximants (/l/, /r/) followed by the vowel /a/, and ending with a plosive (/p/, /t/), nasal (/m/, /ŋ/), fricatives (/s/), and approximants (/l/).

Table 3. Experiment tokens

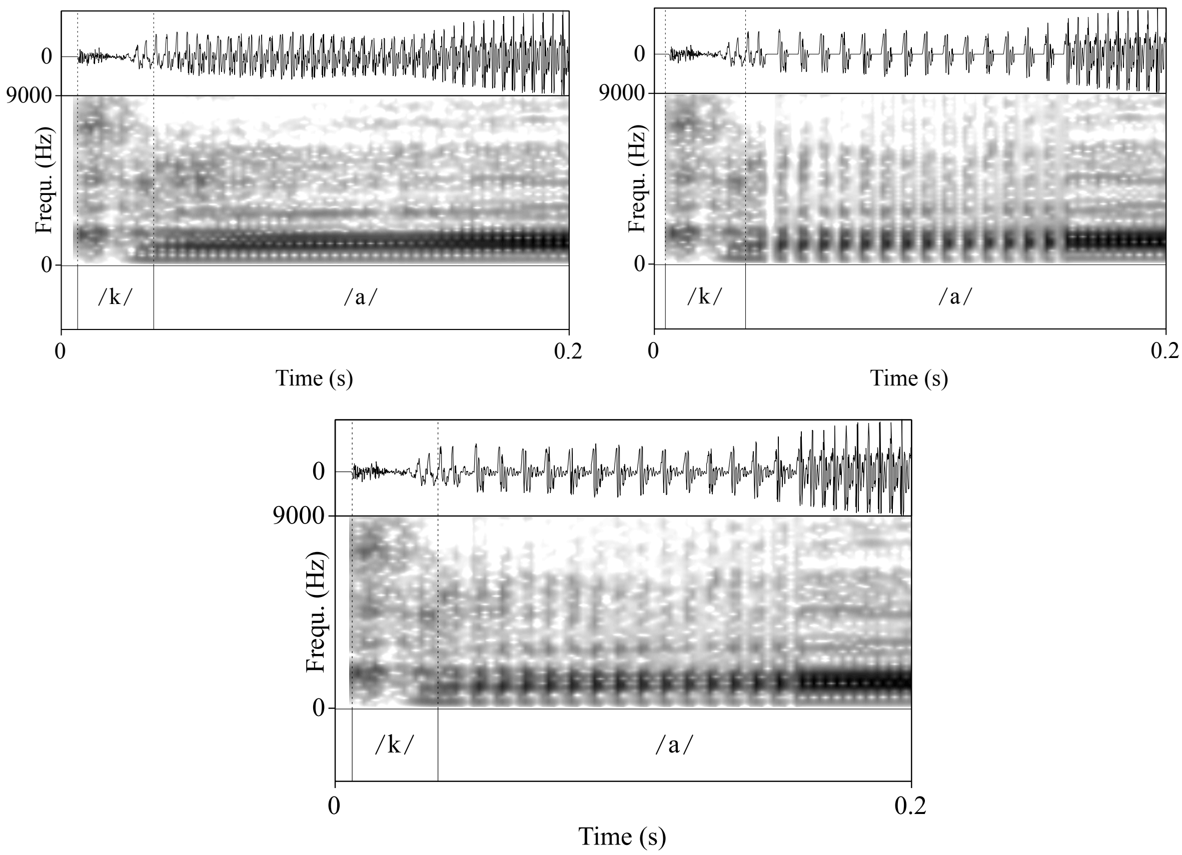
ㄱ /k/	ㄲ/k*/
각이 /kak+i/	깍이 /k*ak+i/

2.2 Stimuli manipulation

This study employed the same methodology used in Hrabanek (2024) and Hrabanek et al. (to appear), with several modifications to the experimental design. First, we did not distinguish whether sound A (the natural fortis tokens) contained creak, as all of sound A in our experimental stimuli showed creaky voice. Second, we only used a single nonce word for the experiment, instead of two. Third, VOT in the X tokens (and neither in A nor B) was not manipulated. Fourth, in generating creaky stimuli, the duplicated sound file contains 15 cycles rather than the 4 used in previous studies. Lastly, in previous studies, when manipulating sound X (the creaky sound), the amplitude of the original sound was set to 60%, while the duplicated sound was set to 40%. In the present study, however, the original sound's amplitude was reduced to 20%, and the duplicated sound increased to 80%, making the manipulated creaky quality more salient.

In the present study, for sound X, artificial creaky sound was created from naturally produced lenis sound. This process was carried out in two steps. Initially, a duplicate of the sound file was made, and within its soundwave, every second cycle in the vowel's periodic waveform, from the onset to the fifteenth even cycle, was set to zero. Unlike the two studies referenced earlier, which created only six cycles, our stimuli require more cycles due to their longer vowel duration. Next, the original file's amplitude was adjusted to 20% of its initial amplitude, the amplitude of the duplicated was adjusted to 80% of the original level, and the two files were merged. As the percentages summed up a full 100%, the merged sound file preserved the original file's properties except where the amplitude was reduced to zero. Within these zero-amplitude parts, only the adjusted to 20% of its original file was shown, leading to an amplitude-modulated sound wave. Figure 1 shows an illustrative example of the stimuli manipulation. The final sound file present audible creak. Additionally, each token's intensity was set to 70 dB. As shown in section1.5, mean VOT and mean F0 for all recorded tokens were extracted using Praat script, the results showed that the VOT and F0 for lenis and fortis stops are very similar for the high tone, consistent with the findings of Ito and Kenstowicz (2009a), see the discussion in section 1.3. Accordingly, no manipulation was applied to VOT or F0 in our experiment.

Figure 1. An example of a lenis stop syllable /ka/ is illustrated with the original sound file (top left) and a modified version with 15 cycles of the vowel portion set to zero (top right). The final audio output (bottom) results from combining the original file at 20% amplitude with the modified version at 80% amplitude.



2.3 Experimental design and participants

The present experiment includes two set of ABX triads (each set contains 5 ABX triads, total 10 ABX triads), the same tokens were used for both sound A and sound B in both sets, with the only distinction being that, for the sound X, the first set used naturally produced lenis tokens, while the second set used manipulated creaky tokens (similar to Hrabanek et al.(to appear)). 5 fortis (/k*aki/) tokens and 10 lenis (/kaki/) tokens (total 15 experimental tokens) were used from the recording. All 5 fortis tokens (sound A), were naturally creaky. For sound B, among 10 tokens of naturally produced lenis stop, half of them (5 tokens) were used as sound B. For sound X, the remaining half (5 tokens) of the naturally produced lenis stops were used for the set 1. For set 2, on the other hand, we manipulated the same half of the naturally produced lenis stops to artificial creak sound. See Appendix 8.2 for the list of experimental items.

Furthermore, 69 tokens were used as fillers and 24 (8 sets *3) tokens of practice trials were extracted from filler tokens. As for the fillers, a total of 36 sets were constructed for the experiment. Of these 20 sets of fillers, sound A and sound B shared same onset and vowel but differed in the coda (e.g. /lapi 1/ - /lami 1/ - /lapi 2/). The remaining 16 sets of fillers, sound A and sound B shared the same vowel and coda but differed in the onset (e.g. /napi 1/ - /mapi 1/ - /napi 2/) (see Appendix 8.3). To ensure that participants fully understand the entire experimental instructions, including the design of the experiment and the use of pressing the M key or Z key for responses, 8 extra ABX sets were used as practice section at the beginning of the experiment (see Appendix 8.4). Each set of experimental items and fillers were played in ABX and BAX order.

20 (15 female and 5 male) native Yanbian Korean speakers (mean age = 28.35 years, SD = 6.56 years, minimum 20 years old, maximum 36 years old) were recruited via family and friend network. All of them were born and raised in Yanbian Korean Autonomous Prefecture. The experiment was conducted online. First, participants were asked about their age, nationality, gender, and the origin of the region (whether they were born and raised in Yanbian or not). After completing the personal information questions, instructions were given as follows: three words (A, B, and X) will be playing. If the final word (sound X) sounds more like the first played word (sound A), press the 'Z' key on the keyboard. If the final word (sound X) sounds more like the second played word (sound B), press the 'M' key on the keyboard. A training section was also included, eight ABX sets were played. Participants could pass the training section if they correctly answered at least three sets. Next, the actual experimental section started. A total of 92 sets (10 sets of experimental items and 36 sets of fillers in ABX and BAX order) were played in random order. All audio stimuli were delivered in stereo format, response instructions (press 'Z' or 'M') appeared only after the three sounds in a set had been played. Each stimulus only played one time. Participants were given a 4-second window to respond before the following set was presented. A break was offered halfway through the session, and participants could resume at their own pace. The entire experiment approximately lasted 15 minutes.

3.Results

3.1 Model design

In our dataset, the majority of participants scored over 90% (mean correctness 93.75%) of correctness on filler items. Only three participants scored below 90%, but all remained above 70% of accuracy. Therefore, the final analysis included all participants. Data analysis was conducted in R studio (Posit team, 2025), mixed-effect models were applied using the lme4 package. Since all participants in our model were from Yanbian Korean Autonomous Prefecture in China, the variable *Region* was excluded from our model. In our model, we used a generalized linear mixed-effect model, with the decision made by participants ('Lenis' or

‘Fortis’) as the outcome variable (coded as ‘Choice’). In terms of predictors, the model included the following fixed effects: *Creak_Presence* (Yes or No), *Order* (ABX or BAX), *Gender* (Male or Female), and *Centered_age* (participants’ real age – mean age), interaction terms involving *Creak_Presence* and all other predictors were incorporated. The model included random intercepts and slopes for participants, letting the influence of *Creak_Presence* and *Order* differ from each participant. Specifically, a random effects structure of the form (Creak_Presence * Order | Participant) was used, indicating that participants may response differently in the Creak_Presence and Order effects. The final model was specified as follows:

*Choice ~ Creak_Presence * Order * Gender * Centered_age + (Creak_Presence * Order | Participant)*

3.2 Model results

Table 4. Results of running mixed-effect model in R studio

	Estimate	Standard Error	z value	Pr (>|z|)
Intercept	1.09536	0.35738	3.065	0.00218 **
Creak_presence-No+Yes	-1.63435	0.63506	-2.574	0.01007*
Order-ABX+BAX	-0.86033	0.52402	-1.642	0.10063
Gender-F+M	-0.02919	0.69957	-0.042	0.96672
Centered_age	0.06906	0.06779	1.016	0.30978
Creak_Presence-No+Yes:Order-ABX+BAX	0.62110	1.03911	0.598	0.55002
Creak_Presence-No+Yes:Gender-F+M	-0.71760	1.23364	-0.582	0.56007
Order-ABX+BAX:Gender-F+M	-0.75333	1.00916	-0.746	0.45537
Creak_Presence-No+Yes:Centered_age	0.01700	0.12044	0.141	0.88773
Order-ABX+BAX:centered_age	0.06256	0.09963	0.628	0.53001
Gender-F+M:centered_age	0.07679	0.13573	0.566	0.57157
Creak_Presence-No+Yes:Order-ABX+BAX:Gender-F+M	0.91703	2.00074	0.458	0.64670
Creak_Presence-No+Yes:Order-ABX+BAX:centered_age	0.07867	0.19766	0.398	0.69063
Creak_Presence-No+Yes:Gender-F+M:centered_age	0.07765	0.24004	0.323	0.74633
Order-ABX+BAX:Gender-F+M:Centered_age	0.06857	0.19876	0.345	0.73010
Creak_Presence-No+Yes:Order-ABX+BAX:Gender-F+M:Centered_age	0.22802	0.39473	0.578	0.56349

Table 4 shows the results of the statistical mixed-effect model in R studio. The intercept, converted from log-odds to odds, indicated that a participant on average was 3 times more likely to choose ‘lenis’ over ‘fortis’ (95% c.i. = 0.38 . . . 1.74; $p < 0.005$). Specifically, this means that the chance of selecting a lenis were 3 times greater than those of selecting fortis. Figure 2 illustrates the decision made by each participant and Table 5 shows each participants’ age, gender, and performance in fillers choice, as well as the number of responses for ‘Lenis’ and ‘Fortis’ in the presence and absence of creak. It also includes the total responses recorded in each condition (‘Lenis’ or ‘Fortis’ without creak, ‘Lenis’ or ‘Fortis’ with creak). Each participant was expected to make 20 choices (10 choices with creak and 10 choices without creak), however, some participants did not complete all choices, and there were 13 ‘none’ responses in total. Notably, 11 out of these 13 ‘none’ responses occurred when creak was present.

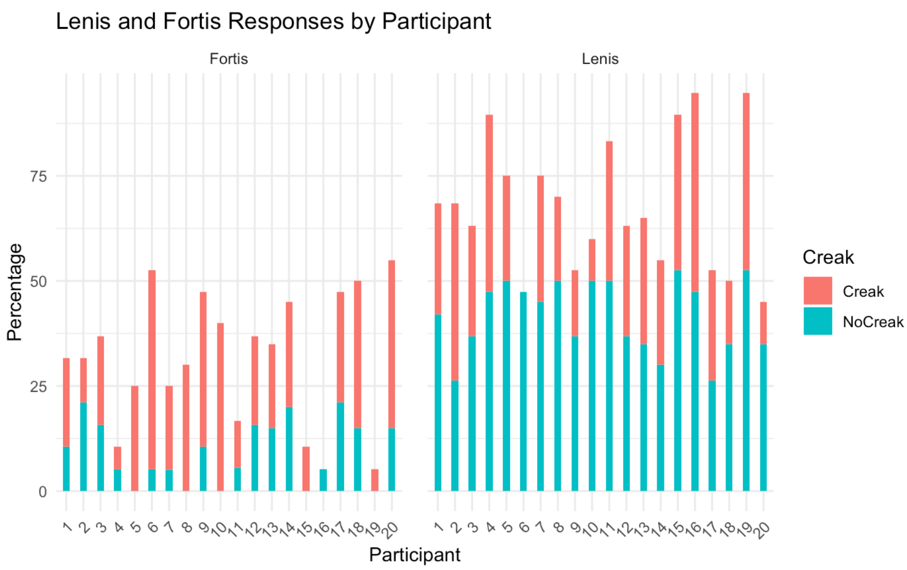
It is clear that most of the participants predominantly chose the lenis option (independent of whether creak was present or not), however, two participants (P6 and P20) chose more fortis choice than lenis choice regardless of the presence of creak (see Table 5 and Figure 2). Specifically, P6 chose 10 times of fortis out of 19 times, and P20 chose 10 times of fortis out of 20 times. A statistically significant effect of *Creak_Presence* was observed. When creak was present, participants chose fortis response more frequently than lenis, with the odds of selecting fortis being 5.36 times more than selecting lenis (95% c.i. = 0.33 . . . 2.8; $p < 0.05$) This means that participants’ choices were influenced by the absence or presence of creak.

The predictor *Order* (ABX or BAX) showed that participants being less likely to make a fortis choice in the BAX order compared to the ABX order, however, this effect was not a significant effect. Similarly, *Gender* (female or male) showed that male participants being very slightly less (1.03 times) likely to choose fortis than female, but this difference was also statistically non-significant. Age, included as a centered continuous variable, showed a minor positive with lenis responses, suggesting a mild tendency for older participants to select the lenis option more often. Nevertheless, this effect was not statistically significant.

Table 5. Age, gender, and the correctness (percentage) on the fillers for each participant, and the number of responses for 'Lenis' and 'Fortis' in the presence and absence of creak

Participants	Age	Gender	Correctness of Fillers	Answer Lenis for X without creak	Answer Fortis for X without creak	Answer Lenis for X with creak	Answer Fortis for X with creak
P1	21	F	98.61%	8	2	5	4
P2	22	F	93.06%	5	4	8	2
P3	20	F	91.67%	7	3	5	4
P4	28	F	98.61%	9	1	8	1
P5	30	M	100%	10	0	5	5
P6	36	F	93.06%	9	1	0	9
P7	36	F	94.44%	9	1	6	4
P8	22	F	97.22%	10	0	4	6
P9	20	F	98.61%	7	2	3	7
P10	20	M	95.83%	10	0	2	8
P11	27	F	95.83%	9	1	6	2
P12	36	F	97.22%	7	3	5	4
P13	26	M	80.56%	7	3	6	4
P14	22	M	93.06%	6	4	5	5
P15	35	F	98.61%	10	0	7	2
P16	36	F	100%	9	0	10	0
P17	35	F	98.61%	5	4	5	5
P18	25	M	97.17%	7	3	3	7
P19	36	F	100%	10	0	8	1
P20	34	F	70.83%	7	3	3	7
Total				**161**	**35**	**104**	**87**

Figure 2. Percentage of 'Fortis' and 'Lenis' responses in all participants, categorized by the presence or absence of creak ('Creak' vs. 'NoCreak')



4. Discussion

Our findings indicate that the presence of creak significantly influenced listeners to perceive word-initial stops as ‘fortis’ more often than ‘lenis’. Participants predominantly chose ‘fortis’ when creak is present, indicating that creaky voice served as a salient cue associated with fortis stops in Yanbian Korean as we hypothesized. In contrast, when creak was absent, participants tend to choose ‘lenis’ more frequently than ‘fortis’ because in this context, sound X and sound A (or B) were all naturally produced lenis sound. Six participants (P6, P8, P9, P10, P18, and P20) identified more ‘fortis’ than ‘lenis’ stops when creak was present, and three participants (P5, P14, P17) chose ‘fortis’ and ‘lenis’ 50% equally (50% each). P6 consistently selected ‘fortis’ 100% of time when creak was present. Since P6 chose 100% of ‘fortis’ all the time when creak was present, given that this participant scored 93.06% on the filler task, the result can be considered reliable. It indicates that for some participants, creak is interpreted as a robust cue for fortis stops in Yanbian Korean. Interestingly, 15 participants (P1, P2, P3, P4, P6, P7, P9, P11, P12, P13, P14, P16, P17, P18, P20) selected ‘fortis’ when the creak was absent (see Table 5), contrary to expectations that they would choose ‘Lenis’ when creak was not present. It illustrates that creak influences the perception of stop types, it is not the only cue, other acoustic or individual perceptual factors may also contribute to the categorization of stops in Yanbian Korean.

4.1 Creak affects the fortis-lenis discrimination

The results address our research question of whether the presence of the creak influences the discrimination between lenis and fortis stops in Yanbian Korean. The presence of creak had a significant effect on listeners’ ability to differentiate fortis and lenis stops. When creak was present, participants were more likely to select the fortis option over the lenis, even though a number of participants still showed a preference for choosing ‘lenis’. This finding aligns with our expectations, given that in Yanbian Korean, VOT and F0 are not considered reliable perceptual cues for distinguishing laryngeal contrasts. Instead, voice quality served as the primary cue for discriminating between lenis and fortis stops in Yanbian Korean. Since voice quality, as measured by H1-H2, is associated with creaky voice, speakers of Yanbian Korean may be more sensitive to creak compared to speakers of Seoul Korean, where VOT and F0 play more dominant role in distinguishing laryngeal contrast in SSK. In contrast, when creak was absent, participants chose ‘lenis’ more often than ‘fortis’, with a total of 161 ‘lenis’ responses compared to only 35 ‘fortis’ responses (see Table 5).

Creak does influence the choice of lenis or fortis stops, but its effect varies across listeners, showing that perception may affected by additional phonetic cues. In our stimuli, VOT and F0 were not manipulated as VOT and F0 cannot be the perceptual cue for discriminating fortis and lenis stops in Yanbian Korean. However, in our experimental tokens, the VOT value of lenis stops was slightly higher than fortis stops (approximately 5.26 ms). Additionally, the F0 of lenis stops was slightly lower than fortis stops (0.9 Hz). This slight differences in VOT and F0 may have a minor effect on listeners' perception, even though they were not intended to serve as primary cues in the discrimination of stop types.

Perception studies conducted by Hrabanek et al. (to appear) showed different results with the present study. They found that the presence of artificial creak in sound X did not affect the differentiate fortis and lenis stops in SSK (see section 1.4) Their results showed a statistically significant effect of Order (ABX or BAX), with participants being more likely to choose the fortis in the BAX order compared to the ABX order. However, the presence or absence of artificial creak in sound X (Condition) did not have a significant effect on the differentiation between lenis and fortis stops. It is because the listeners in their experiment predominantly judge these stops based on VOT and F0. It is also possible that the results were influenced by limitations of their creak manipulation. Since their approach relied primarily on amplitude-based manipulations, the resulting creaky voice may not have been perceptually salient or naturalistic enough. In naturally produced creaky voice, a range of acoustic features typically co-occur, such as inter-cycle variability in amplitude and frequency, jitter, and shimmer. Although the manipulated stimuli were perceived as creaky and natural by all authors, it could be that listeners were less sensitive to the manipulated creak compared to naturally occurring creak, thereby diminishing its influence in the perceptual task. Specifically, the amplitude manipulation in Hrabanek et al.'s (to appear) study, the amplitude level of the original audio was adjusted to 60% of its original amplitude. And the creaky stimuli contained 4 cycles, with every even cycle manually set to zero. In contrast, the present study applied a more extreme manipulation, reducing the amplitude of the original audio to 20%, and included 15 cycles instead or 4 cycles, make the creaky sound more salient and longer. This indicates that the degree and perceptual salience of creak manipulation may play a role in whether creaky voice influences stop consonant differentiation. When the creak is more acoustically salient (as in the present study with a 20% amplitude reduction), listeners may be more likely to consider it into their phonological judgments. Therefore, the mismatch between the two studies suggests that the quality, strength, and length of creaky voice manipulation also may affect perceptual outcomes.

4.2 Yanbian Korean as a tonal language

As 55% (11 participants out of 20 participants) of our participants chose 'lenis' than 'fortis' stops (see Table 5) when the artificial creak was present in sound X, this leads to a theory that creak may influence listener's perception of F0. Some subtypes of creaky voice have been shown to lower perceived pitch, as observed in studies on Mandarin and Black Miao (Kuang, 2013) and Cantonese (Zhang & Kirby, 2020). Due to the connection between creaky voice and low F0, as well as low tones in tonal language, listeners may instinctively perceive creak as low pitch, possibly as a result of an inherent psychoacoustic mechanism (Kuang & Liberman, 2016), particularly among speakers of languages in which creaky voice is commonly use. As a result, the perceived lower F0 may influenced listeners to choose 'lenis' responses during the experiment in SSK. However, since Yanbian Korean is a tonal language and the stimuli carried a high tone, the higher F0 may prompted participants to select more 'fortis' responses when creak was present, compared to previous perception studies in SSK.

Standard Seoul Korean is not a tonal language and F0 serves as an important cue in SSK, while Yanbian Korean is a tonal language and does not use F0 as a primary cue. To what extent a language's pitch accent system influences the potential for F0 variations to function as perceptual cues for distinguishing phonation contrast in stop consonants. Kenstowicz & Park (2006) and Lee & Jongman (2012) suggested that in non-tonal languages, systematic differences in F0 following initial consonants are generally unconstrained, making it likely these preceding consonants can lead to the development of tone (tonogenesis) in such conditions. Kenstowicz & Park (2006) illustrated that in Seoul Korean, which is non-tonal language, when the VOT distinction between aspirated and lenis stops has merged, F0 can become the main cue for distinguishing between the two. On the other hand, in Kyungsang Korean, which is a tonal language, F0 acts two different types of phonological role by conveying both lexical-tone distinctions in vowels and laryngeal contrasts in stop consonants (Kenstowicz & Park, 2006). Unlikely to Seoul Korean, in Kyungsang Korean, lenis and aspirated stops do not undergo VOT merger, and F0 has a minimal function in discriminating laryngeal contrasts. Therefore, Seoul Korean and Yanbian Korean proved the argument which links the phenomenon to how F0 functions in differentiating laryngeal contrasts differently in tonal and non-tonal dialects.

Kang et al. (2022) emphasized that speakers of Yanbian Korean (Hunchun city) continue to use F0 to distinguish between lenis and fortis stops, though this reliance is weaker among younger speakers and in words with low pitch. While our participants were young, the stimuli in our study used high-pitch words, suggesting that F0 may still function as a perceptual cue for younger Yanbian Korean speakers in such contexts. In our experiment, 15 participants chose 'fortis' even the creak was absent. If F0 indeed affects listener's decision of laryngeal contras in Yanbian Korean, this may explain the reason why they chose 'fortis' in cases where 'lenis' would have been expected in the absence of creak. The average F0 calculated in section 1.5 demonstrated that lenis and fortis stops exhibited nearly identical values in high tones. This similarity in pitch may have blurred the acoustic distinction between the two stops, potentially confusing participants' perception of the laryngeal contrast. Additionally, when the native speaker produced the stimuli, she did not use F0 to differentiate between fortis and lenis with a high tone, which may have affected listeners become more indecisive and may have found it more difficult to differentiate between lenis and fortis stops.

Another possible reason for Yanbian Korean speakers may be more sensitive to creaky voice is that all of them are bilingual in Yanbian Korean and Mandarin. In Mandarin, the third tone is often associated with the creak due to its low F0 (Huang, 2020). Since Yanbian Korean speakers are native speakers of both Mandrin and Yanbian Korean, it is possible that their perception in one language influences their perception in another. Specifically, exposure to creaky phonation as a cue for tone in Mandarin may enhance their sensitivity to similar voice quality when processing laryngeal contrasts in Yanbian Korean. This cross-linguistic influence may help explain why creaky voice serves as a more salient cue for Yanbian speakers. In Hrabanek's (2024) study, 6 of the participants (total 33 participants) had approximately one year Mandarin learning experience, and the results indicated that this experience did not have a significant effect on the outcomes. This suggests that, at least in this context, prior exposure to Mandarin may not significantly influence the participants' responses in tasks with laryngeal contrasts. This may also help to explain that why younger participants (in their early 20s) tend to choose lenis less frequently than older participants (in their mid-30s) in the present study. The younger generation has typically received more extensive Mandarin education in school, which may have enhanced their sensitivity to creaky voice as a phonetic cue. As a result, they may rely more heavily on creak when distinguishing between fortis and lenis stops.

4.3 Limitations

The present study presents several limitations. First, the experiment used only a single lexical item, which may limit the generalization of the findings. Relying on one word may not yield fully convincing result due to there would be a potential difference in other words with different syllable constructions. Additionally, only high-tone stimuli were used in the present experiment, which limit a fully understanding of how tonal variation relates to the perception of laryngeal contrast. For example, low-tone stimuli may have different patterns in VOT, F0, and voice quality compared to high-tone stimuli. Without examining both tone types, it is hard to decide whether the findings observed in this study are tone-specific or generalizable across tonal conditions. Another limitation was that VOT was not explicitly manipulated to create ambiguous VOT in the experimental design. Although there was a slight difference in VOT (5.26 ms) between lenis and fortis stops, it may still influenced participants' responses and affected the outcomes of the perception experiment.

4.4 Future studies

The present study examined whether the presence of creak could serve as a perceptual cue for identifying three-way laryngeal contrast in Yanbian Korean. While the findings suggest that creaky voice influence listeners' judgements in relation to the lenis-fortis distinction, several aspects remain open for further investigation. Future studies may consider expanding the lexical variety of stimuli and include words with different syllable structures which would allow more generalizable results. It would also be beneficial to include both high and low tone stimuli to examine how tonal variation interacts with laryngeal perception. Additionally, future experiments could implement more manipulations of acoustic cues such as VOT and F0. The ambiguity of VOT and F0 value could help track how creaky voice contributes to the perception of laryngeal contrasts more accurately. Finally, future studies should also consider comparing younger and older generations of Yanbian Korean. Investigating the older generation may rely less on creaky voice due to limited Mandarin language ability.

5. Conclusion

The current study examined whether creaky voice functions as a perceptual cue in distinguishing the three-way laryngeal contrast in Yanbian Korean. Through a perception experiment using manipulated stimuli, the findings suggest that the presence of creak influenced listeners' judgements in differentiating fortis and lenis stops. The results showed that when creak was present, listeners more tend to choose fortis than lenis, while when the creak was absent, listeners were more likely to choose lenis than fortis. This supports the idea that creaky voice may serve as a cue for identifying laryngeal categories. However, unexpected results were observed, such as participants chose 'fortis' responses even the creak was absent, which highlight the possible role of pitch height (F0), especially the exclusive use of high-tone stimuli. Several limitations including reliance on a single lexical item, absence of low-tone stimuli, and a lack of VOT manipulation which could narrow the generalizability of the findings. Future research should explore a wider range of lexical items, age, tonal conditions and more fine acoustic manipulations to better understand the interaction between pitch, voice quality, and laryngeal perception in bilingual speakers of Yanbian Korean and Mandarin.

6. Acknowledgement

The author would like to express her gratitude to Paul Boersma for proposing an effective statistical model in R, and to Dirk Vet from the University of Amsterdam's Speech Laboratory, for helping with designing and configuring the experiment setup.

7. References

Abberton, E. (1972). Some laryngographic data for Korean stops. *Journal of the International Phonetic Association*, *2*(2), 67–78. https://doi.org/10.1017/s0025100300000517

Boersma, Paul & Weenink, David (2025). Praat: doing phonetics by computer [Computer pro gram]. Version 6.4.35, retrieved 15 June 2025 from http://www.praat.org/

Chai, Y., & Garellek, M. (2022). On H1–H2 as an acoustic measure of linguistic phonation ty pe. *The Journal of the Acoustical Society of America*, *152*(3), 1856–1870. https://doi.o rg/10.1121/10.0014175

Cho, T., Jun, S., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean sto ps and fricatives. *Journal of Phonetics*, *30*(2), 193–228. https://doi.org/10.1006/jpho.2 001.0153

Garellek, M. (2019). The phonetics of voice 1. In *Routledge eBooks* (pp. 75–106). https://doi. org/10.4324/9780429056253-5

Han, J.-I (1998). VOT in the surface distinction of Korean plain and tense stops in initial posi tion : a perception test. *Speech Science*, *3*, 109–117. http://www.dbpia.co.kr/Journal/A rticleDetail/NODE00050783

Han, & Weitzman, R. (1970). Acoustic Features of Korean /P, T, K/, /p, t, k/ and /ph, th, kh /. *Phonetica*, *22*(2), 112–128. https://doi.org/10.1159/000259311

Han, N. (1998). A comparative acoustic study of Korean by Native Korean children and Kore an American Children. *MA Thesis. UCLA*.

Holliday, J. J., & Kong, E. (2011). Dialectal variation in the acoustic correlates of Korean sto ps. *ICPhS*, 878–881. https://dblp.uni-trier.de/db/conf/icphs/icphs2011.html#Holliday K11

Hrabanek, P. (2024). Function of creak in Korean: A perception study. *MA Thesis. University of Amsterdam.*

Hrabanek, P., Watkins, M., & Hamann, S. (to appear). The function of creaky Voice in South Korean: A perception study. *Interspeech*.

Huang, Y. (2020). Different attributes of creaky voice distinctly affect Mandarin tonal perception. *The Journal of the Acoustical Society of America*, *147*(3), 1441–1458. https://doi.org/10.1121/10.0000721

Huang, Y. (2023). Phonetics of period doubling. *University of California, San Diego.* https://escholarship.org/uc/item/33n789f2

Ito, C. (2008). Historical development and analogical change in Yanbian Korean accent. *Harvard Studies in Korean Linguistics XII, 165-178.*

Ito, C., & Kenstowicz, M. (2009a). Mandarin loanwords in Yanbian Korean (1) Laryngeal features. *Phonological Studies*, *12*, 61–72.

Ito, C., & Kenstowicz, M. (2009b). Mandarin Loanwords in Yanbian Korean II: Tones. *Second Language Research*, *45*(1), 85–109. http://web.mit.edu/linguistics/people/faculty/kenstowicz/Yanbian-II.pdf

Jin, W. (2008). Sounds of Chinese Korean: A variationist approach. https://rc.library.uta.edu/uta-ir/handle/10106/970

Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics*, *2*(2), 161–180. https://doi.org/10.1016/s0095-4470(19)31191-x

Kang, Y. (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, *45*, 76–90. https://doi.org/10.1016/j.wocn.2014.03.005

Kang, Y., Schertz, J., & Han, S. (2022). The phonology and phonetics of Korean stop laryngeal contrasts. In *Cambridge University Press eBooks* (pp. 215–247). https://doi.org/10.1017/9781108292351.009

Kang, Y., & Han, S. (2012). Dialectal variation in Korean plosives. *Ms. Korea*.

Keating, P. A., Garellek, M., Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. *ICPhS*. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0821.pdf

Keating, P. A., Garellek, M., Kreiman, J., & Chai, Y. (2023). Acoustic properties of subtypes of creaky voice. *The Journal of the Acoustical Society of America*, *153*(3_supplement), A297. https://doi.org/10.1121/10.0018918

Kenstowicz, M., & Park, C. (2006). Laryngeal Features and Tone in Kyungsang Korean: a Phonetic Study. *Studies in Phonetics Phonology and Morphology*, *12*(2), 247–264. http://digital.kyobobook.co.kr/digital/article/articleDetail.ink?barcode=4010020080069

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, *2*(3), 195–210. https://doi.org/10.1016/s0095-4470(19)31270-7

Kim, C. (1965). On the Autonomy of the Tensity Feature in Stop Classification (with Special Reference to Korean Stops). *WORD*, *21*(3), 339–359. https://doi.org/10.1080/00437956.1965.11435434

Kreiman, J., Gerratt, B. R., & Antoñanzas-Barroso, N. (2007). Measures of the Glottal source spectrum. *Journal of Speech Language and Hearing Research*, *50*(3), 595–610. https://doi.org/10.1044/1092-4388(2007/042

Kreiman, J., Shue, Y., Chen, G., Iseli, M., Gerratt, B. R., Neubauer, J., & Alwan, A. (2012). Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *The Journal of the Acoustical Society of America*, *132*(4), 2625–2632. https://doi.org/10.1121/1.4747007

Kuang, J. (2013). Phonation in tonal contrasts [PhD thesis]. *University of California Los Angeles.*

Kuang, J., & Liberman, M. (2016). The effect of vocal fry on pitch perception. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5260–5264. https://doi.org/10.1109/icassp.2016.7472681

Lee, H., Holliday, J. J., & Kong, E. J. (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, *14*(7). 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 Kyungsang diale cts. *Journal of the International Phonetic Association*, *42*(2), 145–169. https://doi.org/10.1017/s0025100312000035

Lee, K.-H., & Jung, M. S. (2000). Acoustic characteristics and perceptual cues for Korean stops. *Speech Sciences*, 7 (2), 139–155.

Lisker, L., & Abramson, A. S. (1964). A Cross-Language study of voicing in initial stops: Acoustical measurements. *WORD*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Oh, M., & Yang, H. (2013). The production of Stops by Seoul and Yanbian Korean Speakers. *Phonetics and Speech Sciences*, *5*(4), 185–193. https://doi.org/10.13064/ksss.2013.5.4.185

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, *75*(1), 224–230. https://doi.org/10.1121/1.390399

Posit team (2025). RStudio: Integrated Development Environment for R. *Posit Software*, PBC, Boston, MA. URL http://www.posit.co/.

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 Journal of the Association for Laboratory Phonology*, *10*(1), 13. https://doi.org/10.5334/labphon.67

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, *23*(02), 287–308. https://doi.org/10.1017/s0952675706000911

Skuk, V. G., & Schweinberger, S. R. (2013). Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of Speech Language and Hearing Research*, *57*(1), 285–296. https://doi.org/10.1044/1092-4388(2013/12-0314

Watkins, M., Boersma, P., & Hamann, S. (2024). Revisiting pitch jumps: F0 ratio in Seoul Korean. *Interspeech 2022*, 3135–3139. https://doi.org/10.21437/interspeech.2024-1184

Zhang, Y., & Kirby, J. (2020). The role of F0 and phonation cues in Cantonese low tone perc eption. *The Journal of the Acoustical Society of America*, *148*(1), EL40–EL45. https://doi.org/10.1121/10.0001523

8. Appendix

8.1 The script used in Praat for acoustic measurements in our stimuli

```
if numberOfSelected ("Sound") <> 1 or numberOfSelected ("TextGrid") <> 1
   exit "Please select a Sound and a TextGrid first."
endif
textGrid = selected ("TextGrid")
sound = selected ("Sound")

table = Create Table with column names: "table", 0, { "word", "category", "tone",
"beginVowel", "vowelDuration_ms",
... "Duration_log", "pitchMean_Hz", "pitch_1/4", "pitch_4/4", "F0ratio", "H1-H2mean_dB",
"H1-H2_1/4", "H1-H2_4/4" }

selectObject: sound
pitch1 = To Pitch (raw autocorrelation): 0, 75, 600, 15, "no", 0.03, 0.45, 0.01, 0, 0.14
selectObject: sound
pitch2 = To Pitch (raw autocorrelation): 0, 75, 600, 15, "no", 0.03, 0.45, 0.01, 0.35, 0.14
selectObject: sound
pointProcess = To PointProcess (periodic, cc): 75, 600

selectObject: textGrid
numberOfIntervals = Get number of intervals: 2

for interval from 1 to numberOfIntervals
	label$ = Get label of interval: 2, interval

	if label$ = "Vowel"
		starttime = Get start time of interval: 2, interval
		endtime = Get end time of interval: 2, interval
		vowelDur = (endtime - starttime)
		wordint = Get interval at time: 1, starttime
		word$ = Get label of interval: 1, wordint
		if mid$ (word$, 2, 1) = "a"
			category$ = "lenis"
		elsif mid$ (word$, 2, 1) = "h"
			category$ = "aspirated"
		elsif mid$ (word$, 2, 1) = "*"
			category$ = "fortis"
		endif
		if right$ (word$, 1) = "L"
			tone$ = "low"
		else
			tone$ = "high"
		endif

		selectObject: pitch2
		f0mean = Get mean: starttime, endtime, "Hertz"
		selectObject: pitch1
```

```
			pitch_first = Get mean: starttime, (starttime + (vowelDur/4)), "Hertz"
			pitch_last = Get mean: (endtime - (vowelDur/4)), endtime, "Hertz"
			f0ratio = pitch_first/pitch_last

			selectObject: pitch1, pointProcess, sound
			report$ = Voice report: starttime, endtime, 75, 600, 1.3, 1.6, 0.03, 0.45
			h1minush2 = extractNumber (report$, "Mean of the points: ")
			selectObject: pitch1, pointProcess, sound
			report_first$ = Voice report: starttime, (starttime + (vowelDur/4)), 75, 600,
1.3, 1.6, 0.03, 0.45
			h1minush2_first = extractNumber (report_first$, "Mean of the points: ")
			selectObject: pitch1, pointProcess, sound
			report_last$ = Voice report: (endtime - (vowelDur/4)), endtime, 75, 600, 1.3,
1.6, 0.03, 0.45
			h1minush2_last = extractNumber (report_last$, "Mean of the points: ")

			selectObject: table
			Append row
			rowNumber = Get number of rows
			Set string value: rowNumber, "word", word$
			Set string value: rowNumber, "category", category$
			Set string value: rowNumber, "tone", tone$
			Set string value: rowNumber, "beginVowel", fixed$ (starttime, 3)
			Set string value: rowNumber, "vowelDuration_ms", fixed$ (vowelDur*1000,
3)
			Set string value: rowNumber, "Duration_log", fixed$ (log10(vowelDur*1000),
3)
			Set string value: rowNumber, "pitchMean_Hz", fixed$ (f0mean, 3)
			Set string value: rowNumber, "pitch_1/4", fixed$ (pitch_first, 3)
			Set string value: rowNumber, "pitch_4/4", fixed$ (pitch_last, 3)
			Set string value: rowNumber, "F0ratio", fixed$ (f0ratio, 3)
			Set string value: rowNumber, "H1-H2mean_dB", fixed$ (h1minush2, 3)
			Set string value: rowNumber, "H1-H2_1/4", fixed$ (h1minush2_first, 3)
			Set string value: rowNumber, "H1-H2_4/4", fixed$ (h1minush2_last, 3)
			selectObject: textGrid
		endif
endfor

removeObject: pitch1, pitch2, pointProcess

selectObject: table
Save as tab-separated file: "results.Table"
View & Edit
selectObject: textGrid, sound
```

8.2 The list of experimental items

k*aki1	kaki6	kaki1
k*aki2	kaki7	kaki2
k*aki3	kaki8	kaki3

k*aki4	kaki9	kaki4
k*aki5	kaki10	kaki5

k*aki1	kaki6	kaki1_creak
k*aki2	kaki7	kaki2_creak
k*aki3	kaki8	kaki3_creak
k*aki4	kaki9	kaki4_creak
k*aki5	kaki10	kaki5_creak

8.3 The list of fillers

lapi1	lami1	lapi2
lali1	lapi3	lali2
lami2	lali3	lami3
lapi4	lali4	lapi5
lapi6	lami4	lami5
lali5	lapi7	lapi8
lami6	lali6	lali7
napi1	naŋi1	napi2
napi3	naŋi2	naŋi3
sami1	sasi1	sami2
sati1	sami3	sati2
sasi2	sati3	sati4
sami4	sasi3	sasi4
sasi5	sami5	sasi6
sati5	sami6	sati6
hasi1	hami1	hami2
hami3	hasi2	hasi3
haŋi1	hami4	haŋi2
hasi4	haŋi3	hasi5
hami5	hasi6	hami6
napi1	mapi1	napi2
mapi2	napi3	mapi3
lapi1	napi4	lapi2
lapi3	napi5	napi6
mapi4	lapi4	mapi5
mapi6	lapi5	lapi6
lami1	sami1	lami2
sami2	lami3	sami3
hami1	sami4	hami2
sami5	hami3	sami6
lami4	hami4	hami5
hami6	lami5	lami6
sasi1	hasi1	sasi2
sasi3	hasi2	hasi3
haŋi1	naŋi1	haŋi2
naŋi2	haŋi3	naŋi3

8.4 The list of practice trials

lapi1	sami1	sami2
lami1	sasi1	lami2
hasi1	lapi2	hasi2
mapi1	sami3	sami4
mapi2	hami1	mapi3
lali1	sati1	sati2
hasi3	lali2	lali3
napi1	naŋi1	napi2