Master's Thesis of Arts

Generational changes in the perception of foreign voiced stops by Korean listeners

세대에 따른 한국어 화자의
외국어 어두 유성 폐쇄음 지각 변화

August 2021

Graduate School of Linguistics
Seoul National University
Phonetics Major

Jang Hyeonjeong

Generational changes in the perception of foreign voiced stops by Korean listeners

Lee Ho-Young

Submitting a Master's thesis of Arts

August 2021

Graduate School of Linguistics
Seoul National University
Phonetics Major

Jang Hyeonjeong

Confirming the master's thesis written by
Jang Hyeonjeong
August 2021

Chair Chung Minhwa

Vice Chair Kim Sunhee

Examiner Lee Ho-Young

Abstract

This study aims to investigate how Korean speakers perceive foreign unaspirated voiceless and voiced stops, and how the mechanisms of stop perception differ across generations. Contemporary Seoul Korean speakers tend to utilize fundamental frequency (f_0) of the following vowel as the primary cue for perceiving lenis stops, while older generations relied more on voice onset time (VOT). In the lenis and aspirated stop perception, f_0 and VOT were also observed to trade off. In this study, a perception and mapping test was conducted with 20 teenage subjects and 20 senior subjects to determine if the cue trade-off and the changing tendency with respect to cue-weighting, observed in Korean speakers' native voiceless stop perception, are also present in their foreign voiced stop perception. The results indicate that younger listeners do perceive foreign voiced stops differently from the older listeners. Consistent with their native stop perception, the teenage subjects utilized f_0 more actively as a cue when perceiving word-initial unaspirated voiceless and voiced stops. The trend towards increased f_0-dependency seems to be consistent across native and foreign stops with or without aspiration or prevoicing. However, the f_0-VOT trade-off observed in previous studies on Korean speakers' voiceless stop perception was not found to be replicated in this experiment with stimuli with near-zero to negative VOT values. For Korean speakers, whose native language lacks voicing contrasts in stop consonants at word-initial positions, negative and positive VOT in word-initial stop consonants may not serve equally well as a cue.

Keywords: generational change, Korean speaker, stop perception, voice onset time, VOT, fundamental frequency, f_0, vowel pitch

Student No.: 2015-20066

Table of Contents

Tables

Figures

1. Introduction

Studies on Korean listeners' perceptual mechanism of stop consonants in word-initial positions, as well as chronological change in perceptual cue weighting, have primarily focused on how the three voiceless stop categories in Korean are distinguished by the native listeners. In word-initial positions, Korean language shows three-way distinction among voiceless stops but lacks voice distinction. Previous studies found that differentiating factors for Korean lenis, unaspirated fortis (hereafter, "fortis"), and aspirated fortis (hereafter, "aspirated") stop categories, in both perception and production, include consonantal and vocalic values, such as linguopalatal contact (Cho & Keating, 2001), glottal aperture (Hirose, Lee, & Ushijima, 1974; Jun, Beckman, & Lee, 1998; Kagaya, 1974; C.-W. Kim, 1970), glottal and laryngeal muscular control (Dart, 1987; Hardcastle, 1973; Hirose et al., 1974; Hirose, Park, & Sawashima, 1983; Jun, 1996; Kagaya, 1974; C.-W. Kim, 1965), transglottal and intraoral airflow (Ahn, 1999; Dart, 1987; Hardcastle, 1973; Kagaya, 1974), closure duration (Park, 2002), voice onset time (Cho, 1996; Cho, Jun, & Ladefoged, 2002; J.-I. Han, 1996; M. S. Han & Weitzman, 1970; Hardcastle, 1973; Hirose et al., 1974; Kagaya, 1974; C.-W. Kim, 1965; M. Kim, 2004; Mi-Ran Kim, 1994; Lisker & Abramson, 1964; Shimizu, 1990;

Silva, 1992), release burst intensity and aspiration intensity (J.-I. Han, 1996; C.-W. Kim, 1965), fundamental frequency of the following vowel (Cho, 1996; Choi, 2002; Dart, 1987; J.-I. Han, 1996; M. S. Han & Weitzman, 1970; Hardcastle, 1973; Kagaya, 1974; C.-W. Kim, 1965; M. Kim, 2004; Mi-Ran Kim, 1994; Mi-Ryoung Kim, 2000; Park, 2002; Shimizu, 1990), the first and second formants of the following vowel (M. S. Han & Weitzman, 1970; Hardcastle, 1973; Park, 2002; Shimizu, 1990), intensity of vowel onset (M. S. Han & Weitzman, 1970), and vowel length (Cho, 1996).

	bilabial	alveolar	velar
lenis	/p/	/t/	/k/
fortis	/p$^{=}$/	/t$^{=}$/	/k$^{=}$/
aspirated	/p^{h}/	/t^{h}/	/k^{h}/

Table 1: Korean stop consonants

The lenis (or lax), romanized as ⟨b⟩, ⟨d⟩, and ⟨g⟩, and phonemically described as /p/, /t/, and /k/, show moderate aspiration in the initial positions whilst voiced in intervocalic positions (Jun, 1993; Silva, 1992). The fortis (or tense), romanized as ⟨pp⟩, ⟨tt⟩, and ⟨kk⟩ and diversely transcribed as they produced with the shortest-lag VOT and properties related to “tenseness” such as greater glottal constriction, high subglottal pressure, tenseness in the walls of the vocal tract, and lowering or other expansion of the larynx, negative H1-H2 (Ladefoged & Maddieson, 1996). Common transcriptions include /p͈/, /t͈/, /k͈/ with the strong articulation diacritics, /p*/, /t*/, /k*/ with asterisks, and /p'/, /t'/, /k'/ with apostrophes. In this study, they are transcribed with unaspirated diacritics, as /p˭/, /t˭/, /k˭/. The aspirated, romanized as ⟨p⟩, ⟨t⟩, and ⟨k⟩ and transcribed as /p^h/, /t^h/, /k^h/, are associated with the long-lag VOT.

Defining the primary cue to the three-way stop distinction in Korean has been a recurring topic in phonetic studies, as Korean speakers' contrast realization and perceptual cue-weighting pattern have shifted over time. Some early studies such as the cross-linguistic voicing analysis by Lisker & Abramson (1964) report the consistent VOT differences among the Korean stop sets as the singly satisfactory differentiating factor, although the unaspirated fortis and lenis show overlapping distribution in VOT values. The average VOT-lag reported by Lisker & Abramson (1964) was 7ms for the fortis, 18ms for the lenis, and 91ms for the aspirated. Other researchers such as Kim (1965) suggested the "tensity" feature as the primary differentiating cue for Korean stop sets. According to Kim (1965), the lenis are distinguished from the aspirated and unaspirated fortis sets using the "tensity" feature, being smaller in bursts and aspiration intensity, amplitude and vibration rate at vowel onset, duration of pressure and amount of airflow, articulatory contact, and lip muscle activity in case of bilabial stops.

Later studies found low f_0 of the following vowel, rather than the aforementioned glottal and laryngeal configurations and related intensity features, to be the primary distinguishing factor for the lenis set (Jun, 1993; Mi-Ryoung Kim, 2000; Silva, 2006). Along with the f_0 differences, what has been observed is seemingly "merging" VOT differences between the aspirated and lenis stops. In the production test conducted by Kim (2017), the average VOT for the aspirated stops were found shortened to 79ms and the average VOT for the lenis lengthened to 69ms, which shows a drastic change in terms of the VOT landscape among the three stop categories–Overlaps now being found in the values for the lenis and aspirated, instead of lenis and unaspirated fortis as in 1960s.

	fortis	lenis	aspirated
Lisker & Abramson (1964)	7 ms	18 ms	91 ms
C.-W. Kim (1965)	12 ms	35 ms	93 ms
M.-R. Kim (2017)	13 ms	69 ms	79 ms

Table 2: Average length of VOT-lag in Korean consonants

Similar changes have also been observed in perception. Kim, Beddor, & Horrocks (2002) demonstrated the perceptual dominance of f_0 over VOT for lenis sets. A comparative study of non-tonal Seoul Korean and tonal Gyeongsang Korean by Lee, Politzer-Ahles, & Jongman (2013) also suggests low f_0 as the primary cue to perceive lenis stops for Seoul Korean speakers, although VOT was found to be the primary cue utilized by tonal Gyeongsang Korean speakers. Kim (2004) demonstrated the trade-off between the f_0 and VOT in the perception and the production of Korean lenis and aspirated stops—Stops with longer VOT can be perceived as lenis stops if the pitch of the following vowel is sufficiently low; Stops with lower pitch can be perceived as aspirated stops if the VOT is sufficiently long. In this study, the stops with zero to near-zero positive VOT are found to be always perceived as tense stops.

With the changing tendency for cue weighting in stop perception, theories on tonal contrast have developed. Some studies suggest that VOT merger between the lenis and aspirated consonants in Seoul Korean leads to the development of tonal contrast, similar to the Tibeto-Burmese tonogenesis (Kang, 2014; Kang & Han, 2013; Mi-Ryoung Kim, 2000, 2016, 2017; Silva, 2006). According to Kang (2014), younger female speakers of Korean have already switched to the tonal contrasts.

Whether this diachronic change in the perceptual mechanism for native voiceless stops can be extended to or have influenced the Korean speakers' cue-weighting in their perception of foreign voiced stops remains unclear. A previous study by Schmidt (1996) reported that Korean speakers mapped English "voiced" stops with both lenis and fortis, but stops with VOT-lead were usually perceived as lenis. The mapping is consistent with what Shinohara, Ji, Ooigawa, & Shinya (2011) reports, in their study on Korean speakers' perception of English, French, Japanese, and Mandarin stops, where most of the tokens with prevoicing were mapped with the Korean lenis. However, the pre-voiced tokens tended to be low-pitched as well, all of them showing f_0 not higher than 150Hz, while voiceless tokens were distributed in high-frequency areas, up to 210Hz and more in the f_0 scale. In Schmidt's study, tenseness-related features other than f_0, such as an abrupt onset of the consonant and the shorter intensity rise time also caused more mapping to fortis as expected.

Chronological or generational changes in Korean listeners' voiced stop perception is a yet little explored topic, although Kang (2008) compared the 1930s Korean speakers' English loanword adaptation practices and the English-Korean perceptual mapping in the 2000s, suggesting that the frequent mapping of the English "voiced" stops to Korean fortis in the 1930s was caused by the tokens' short-lag VOT, which was more important a cue by Korean listeners of that period, while in the increased mapping to lenis in 2000s was induced by the tokens' low f_0, which have become the primary differentiating factor. One of the limitations with this explanation is that loanword adaptation reflects more than perception. It should be noted that in 1930s, many English loanwords were introduced into Korean via Japanese.

Heeding these previous findings, this study attempts to explore Korean listeners' categorization of foreign voiced and voiceless unaspirated stops, focusing on possible generational differences in cue-weighting between f_0 and VOT, which may be related to the changes in their native language perception mechanism.

2. Experiment

A perception test was conducted to identify if younger Korean speakers perceive "tenseness" using vowel pitch more so than older speakers do.

2-1. Stimuli

60 stimuli, consisting of six unaspirated voiceless and voiced consonant sets (/p/-/b/, /t/-/d/, /k/-/g/) followed by five vowels (/a/, /e/, /i/, /o/, /u/) and the second syllable consisting of /ʁ/ and /ɐ/ (following the

syllables with the unrounded vowels) or /u/ (following the syllables with the rounded vowels) were recorded with two pitch patterns (HL and LH). 40 fillers were recorded with /n/, /m/, /l/, and /s/ replacing the word-initial stops. The stimuli were recorded twice consecutively to avoid any dropouts and obtain at least one satisfactory sound for each token.

The stimuli were recorded by a female native European Portuguese speaker in her early 20s, using a Marantz PMD-660 recorder and a Shure SM48S microphone with a sampling rate of 44100 Hz at 16-bit resolution. The distance between the microphone and the subject's mouth was 8cm. The words were given in conventional Portuguese spellings.

	/p/	/b/	/t/	/d/	/k/	/g/
/a/	parra /ˈpa.ʁɐ/	barra /ˈba.ʁɐ/	tarra /ˈta.ʁɐ/	darra /ˈda.ʁɐ/	carra /ˈka.ʁɐ/	garra /ˈga.ʁɐ/
/e/	perra /ˈpe.ʁɐ/	berra /ˈbe.ʁɐ/	terra /ˈte.ʁɐ/	derra /ˈde.ʁɐ/	querra /ˈke.ʁɐ/	guerra /ˈge.ʁɐ/
/i/	pirra /ˈpi.ʁɐ/	birra /ˈbi.ʁɐ/	tirra /ˈti.ʁɐ/	dirra /ˈdi.ʁɐ/	quirra /ˈki.ʁɐ/	guirra /ˈgi.ʁɐ/
/o/	porro /ˈpo.ʁu/	borro /ˈbo.ʁu/	torro /ˈto.ʁu/	dorro /ˈdo.ʁu/	corro /ˈko.ʁu/	gorro /ˈgo.ʁu/
/u/	purro /ˈpu.ʁu/	burro /ˈbu.ʁu/	turro /ˈtu.ʁu/	durro /ˈdu.ʁu/	curro /ˈku.ʁu/	gurro /ˈgu.ʁu/

Table 3: Stimuli recorded by the European Portuguese speaker

	/m/	/n/	/l/	/s/
/a/	marra /ˈma.ʁɐ/	narra /ˈna.ʁɐ/	larra /ˈla.ʁɐ/	sarra /ˈsa.ʁɐ/
/e/	merra /ˈme.ʁɐ/	nerra /ˈne.ʁɐ/	lerra /ˈle.ʁɐ/	serra /ˈse.ʁɐ/
/i/	mirra /ˈmi.ʁɐ/	nirra /ˈni.ʁɐ/	lirra /ˈli.ʁɐ/	sirra /ˈsi.ʁɐ/
/o/	morro /ˈmo.ʁu/	norro /ˈno.ʁu/	lorro /ˈlo.ʁu/	sorro /ˈso.ʁu/
/u/	murro /ˈmu.ʁu/	nurro /ˈnu.ʁu/	lurro /ˈlu.ʁu/	surro /ˈsu.ʁu/

Table 4: Fillers recorded by the Portuguese speaker

As European Portuguese exhibits two-way voicing distinction in stops (Cruz-Ferreira, 1995), it is suitable for testing voicing perception. The [ʁɐ] and [ʁu] in the second syllables are not likely to be readily categorizable to Korean speaking subjects, because uvular fricative /ʁ/ present in European Portuguese is an unfamiliar sound to Korean speakers, whose native language lacks uvular consonants and has only two fricatives, /s/ and /h/. Also, because European Portuguese is a stress-timed language showing paroxytonic stress patterns, often results in devoicing of word-final reduced vowels /ɯ/, /ɐ/ and /u/ (Cruz-Ferreira, 1995). The first syllable of each stimulus, on the other hand, consists of familiar, readily-categorizable segments.

To produce tokens with both high-pitched and low-pitched first syllables, pitch patterns HL and LH in disyllabic stimuli were induced by instructing the speaker to read first with falling intonation as if each word were a sentence, "Barra. Darra. Garra." and then read with rising intonation as if she were reading a list "Barra, darra, garra,".

2-1-1. Characteristics

Each token was analysed with Praat (version 6.0.15), set with 5 ms window length and 50 dB dynamic range (Boersma & Weenink, 2016).

Recorded stimuli with initial voiced stops showed VOT-lead ranging from -51.33ms to -113.18ms, the average being -83.57ms. Stimuli with initial voiceless stops showed VOT-lag ranging from 8.11ms to 19.56ms, the average being 14.84ms. An independent-samples t-test was conducted, using R version 3.0.2, to compare the VOT values of the recorded stops. There was a significant difference in the values for the voiceless stops and the voiced stops ($t = -32.42$, $p <$ 2.2e-16). *P*-value from an *F*-ratio score generated from an ANOVA test showed that VOT did not differ significantly among the consonants within voiced and voiceless sets. ($F = 0.859$, $p = 0.435$ within the voiced group; $F = 0.713$, $p = 0.499$ within the voiceless group).

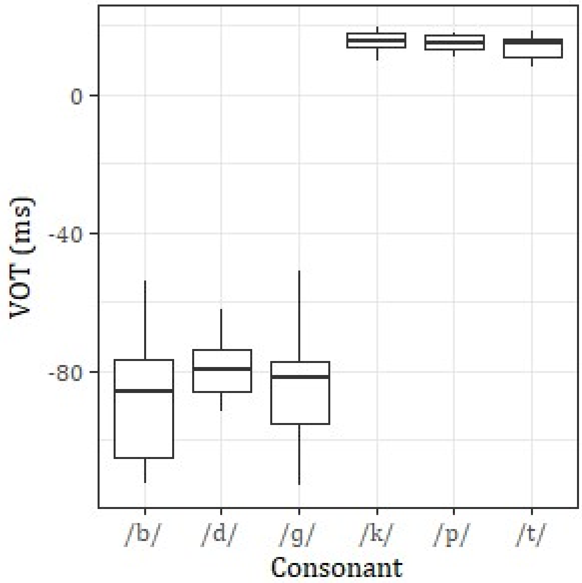


Figure 1: VOT distribution in stimuli by the initial consonant

VOT did not differ in relation to the pitch patterns or the vowel quality of the following vowels. VOT differences were insignificant among the stimuli with different following vowels, both in voiced stimuli group ($F = 2.669$, $p = 0.0557$) and voiceless stimuli group ($F = 0.866$, $p = 0.498$).

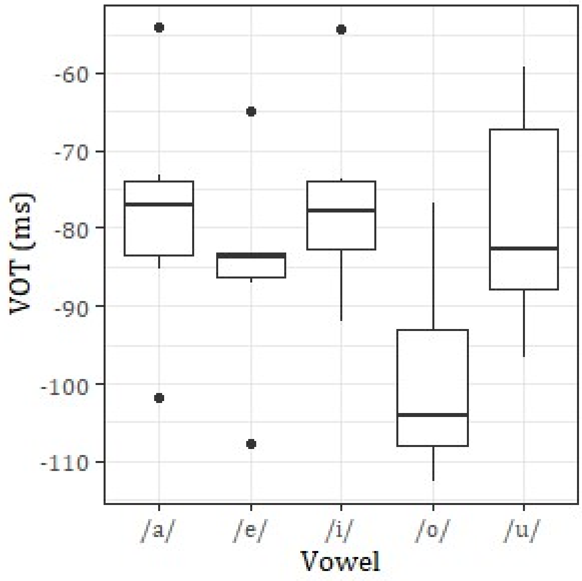


Figure 2: VOT distribution in voiced stimuli by the following vowels

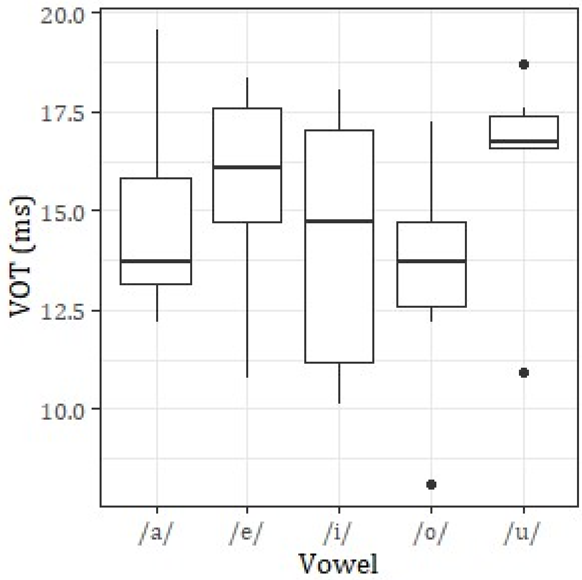


Figure 3: VOT distribution in voiceless stimuli by the following vowels

The differences were also insignificant among the stimuli with different pitch patterns, both in voiced stimuli group ($F = 0.669$, $p = 0.42$) and voiceless stimuli group ($F = 0.141$, $p = 0.71$).

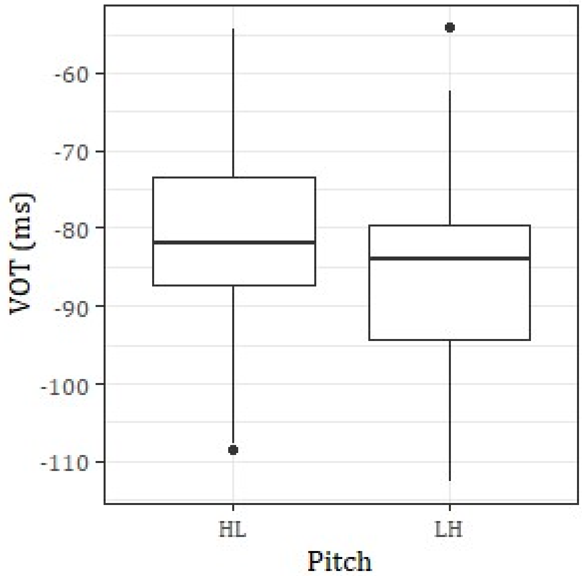


Figure 4: VOT distribution in voiced stimuli by pitch patterns

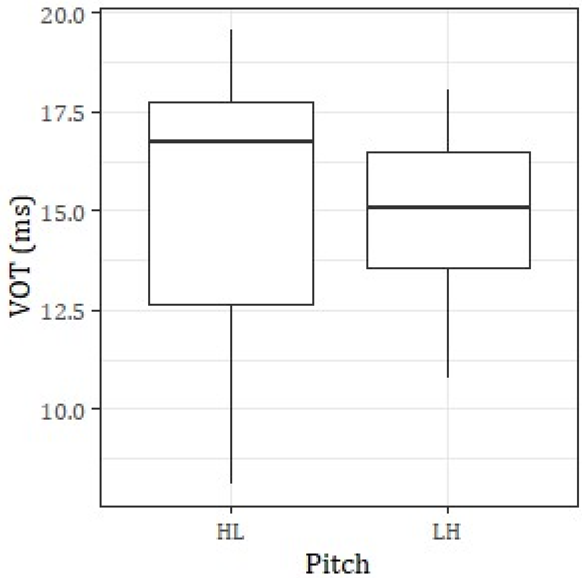


Figure 5: VOT distribution in voiceless stimuli by pitch patterns

f_0 at the onset of the following vowel was lower in the voiced group than the voiceless group, with the average being 247.92 Hz and 259.70 Hz respectively (t = 57.955, p = 8.43e-05). f_0 did not differ significantly among the stimuli with different consonants, within voiced stimuli group ($F=0.765$, $p=0.475$) as well as within voiceless stimuli group ($F=2.751$, $p=0.0818$).

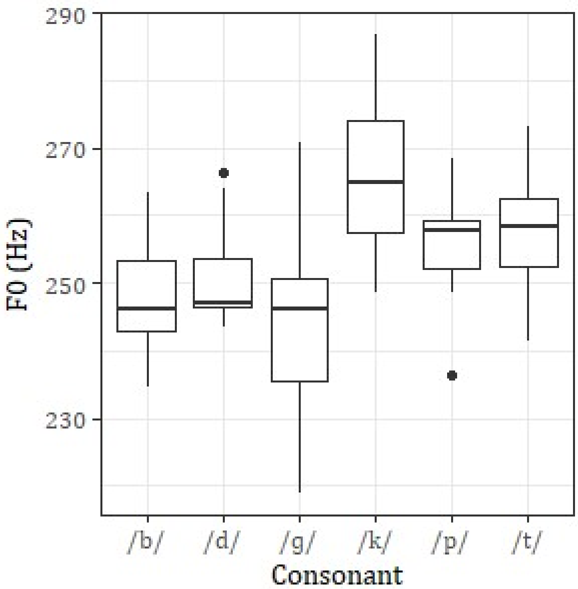


Figure 6: f_0 distribution in stimuli by initial consonant

Vowel quality was not relevant to f_0 values, in both voiced stimuli group ($F=0.986$, $p=0.433$) and voiceless stimuli group ($F=0.615$, $p=0.656$).

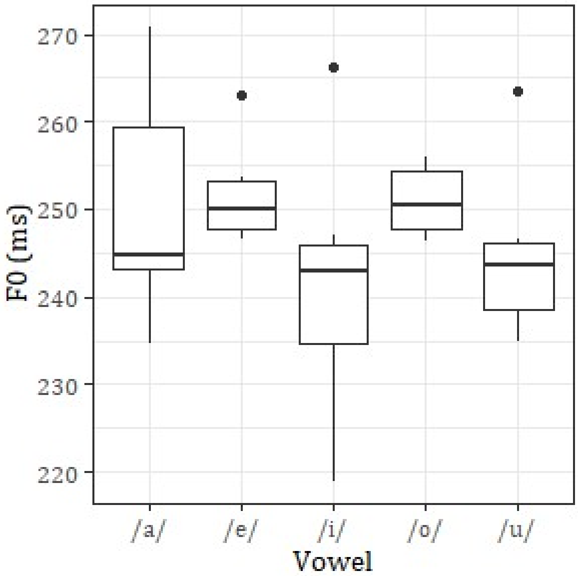


Figure 7: f_0 distribution in voiced stimuli by following vowel

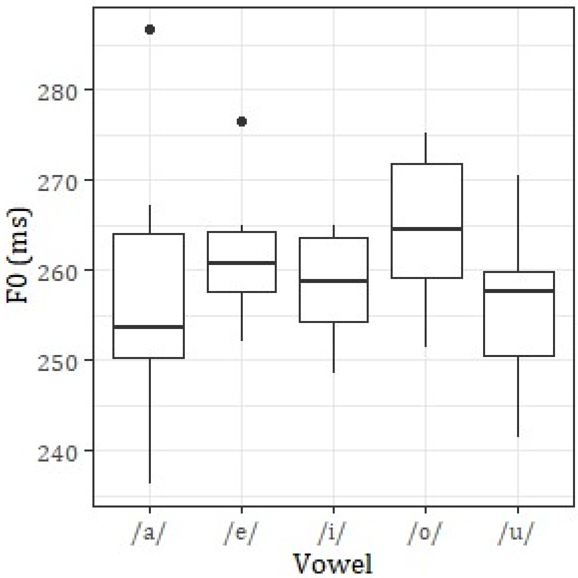


Figure 8: f_0 distribution in voiceless stimuli by following vowel

f_0 of the first syllable varied with pitch pattern in the voiced stimuli group (t = 2.7612, p = 0.01056), with the average being 252.90 Hz when HL and 242.95 Hz when LH. The difference was also significant in the voiceless stimuli group (t=4.7927, p=5.13e-05), with the average being 266.73 Hz when HL and 252.68 Hz when LH.

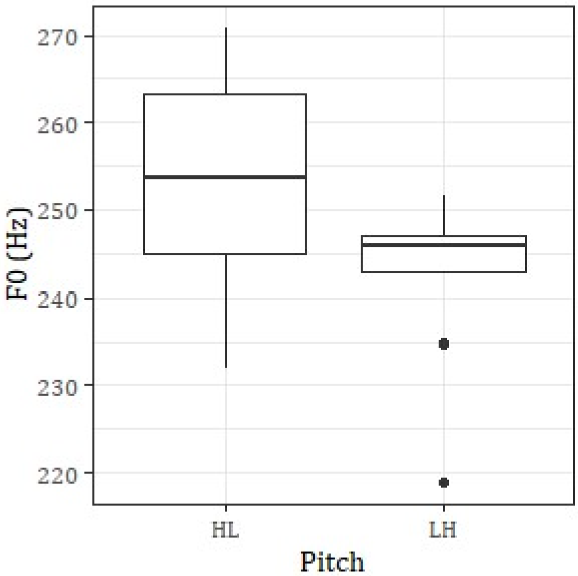


Figure 9: f_0 distribution in voiced stimuli by following vowel

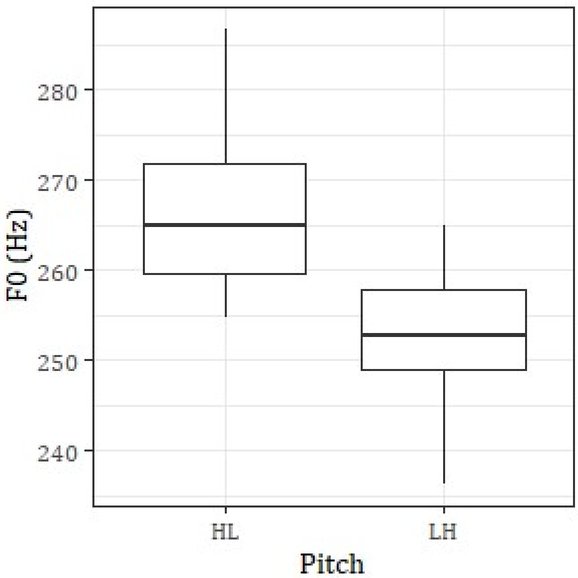


Figure 10: f_0 distribution in voiceless stimuli by following vowel

2-1-2. Manipulation

Each stimulus was manipulated and made into ten tokens, with two VOT steps (voiced, voiceless) and five f_0 steps (-20 Hz, -10 Hz, original, $+10$ Hz, $+20$ Hz). Prevoicing was removed from the recordings of the voiced stops to produce devoiced tokens. Recordings of voiceless stops were made into prevoiced stops by removing the aspiration and pasting the prevoicing copied from the homorganic voiced stimuli.

The stimuli were manipulated using Praat (version 6.0.15) (Boersma & Weenink, 2016). The left edge of prevoicing was manually marked at the first positive zero-crossing of the first periodic waveform, and the right edge was marked at the closest zero-crossing left of the burst. Aspiration longer than 10ms was removed, with left and right edges of the removed portion being zero-crossings, avoiding going near the burst offset and the vowel onset.

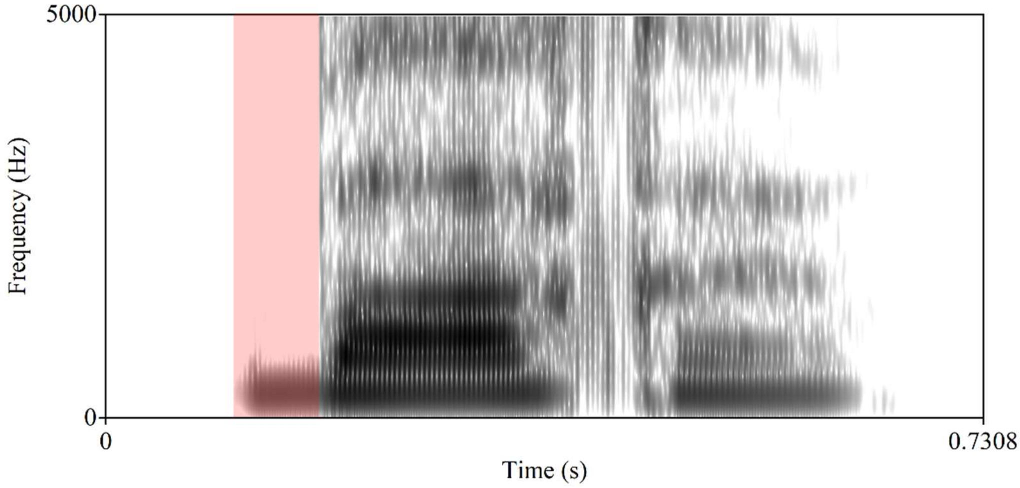


Figure 11: Token /ba/ with the prevoicing portion highlighted

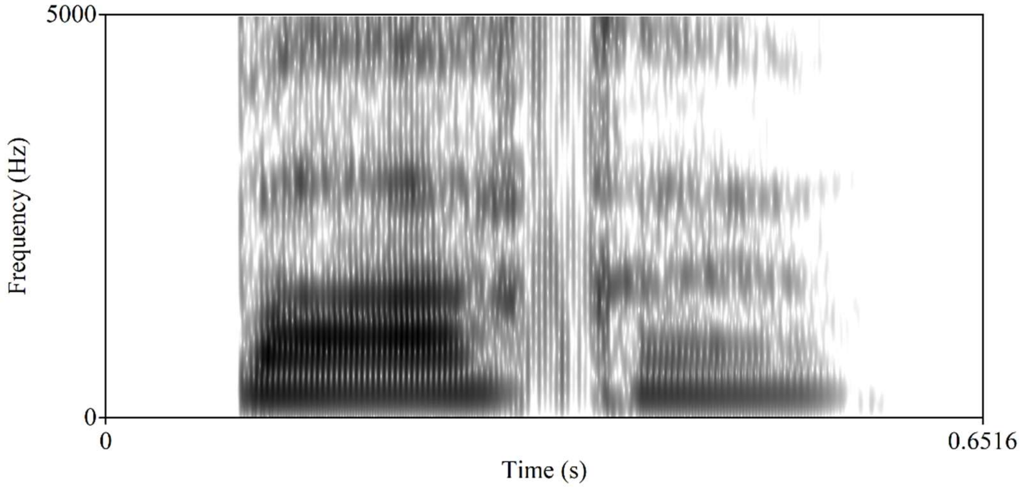


Figure 12: Token /ba/ with the prevoicing removed

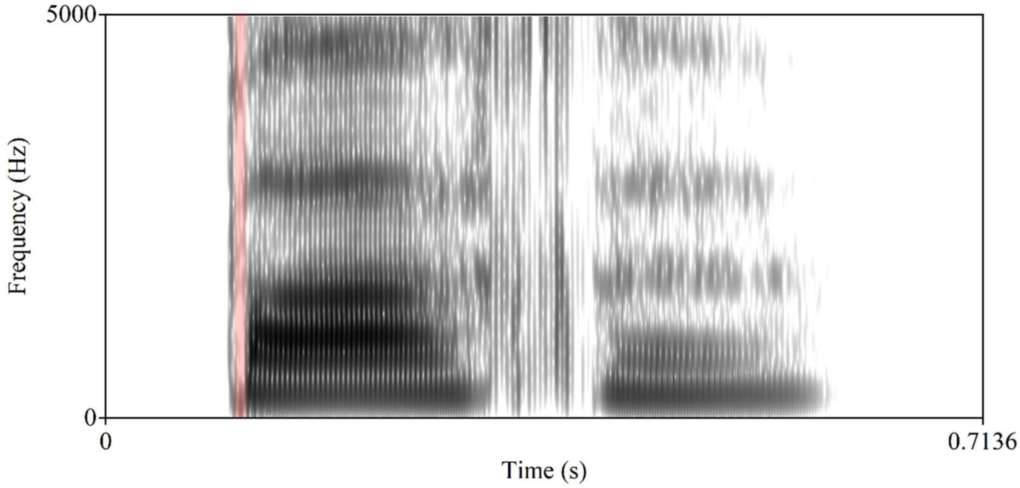


Figure 13: Token /pa/ with the aspiration highlighted

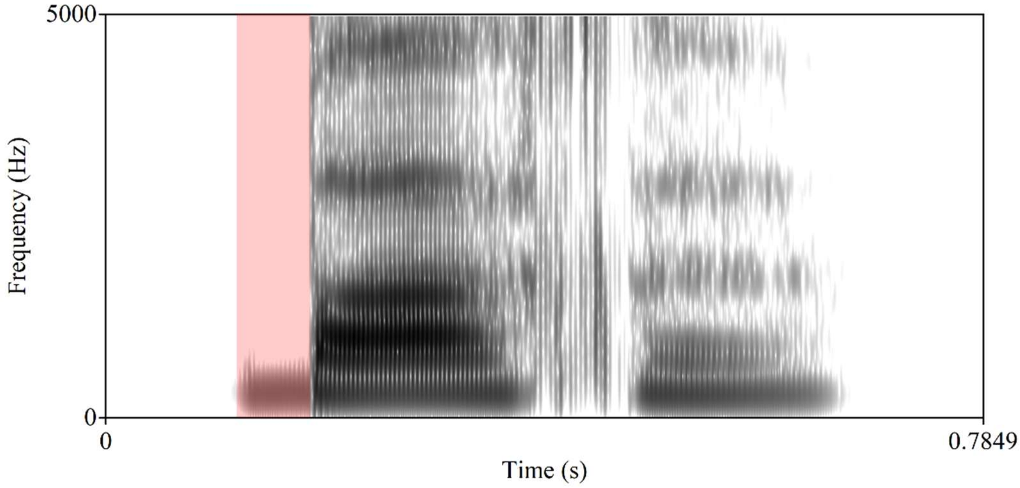


Figure 14: Token /pa/ with the aspiration removed and prevoicing added

f_0 manipulation was done using Praat ManipulationEditor, shifting the pitch frequency of the vowel portion of the first syllable by -20 Hz, -10 Hz, $+10$ Hz, and $+20$ Hz.

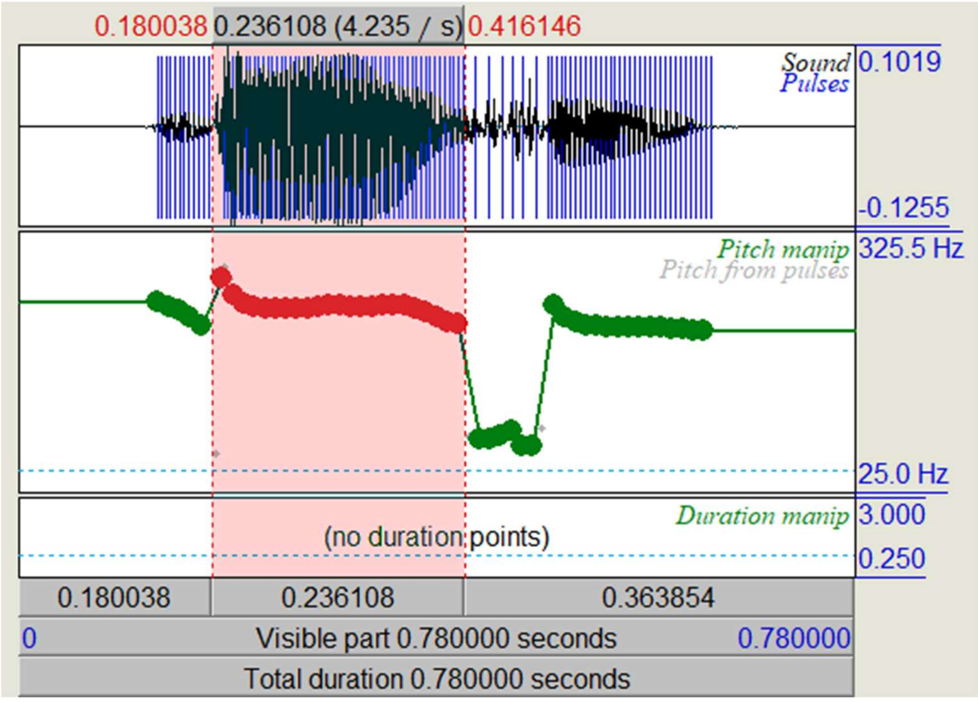


Figure 15: Vowel portion of the stimulus /da/ selected on the Manipulation window

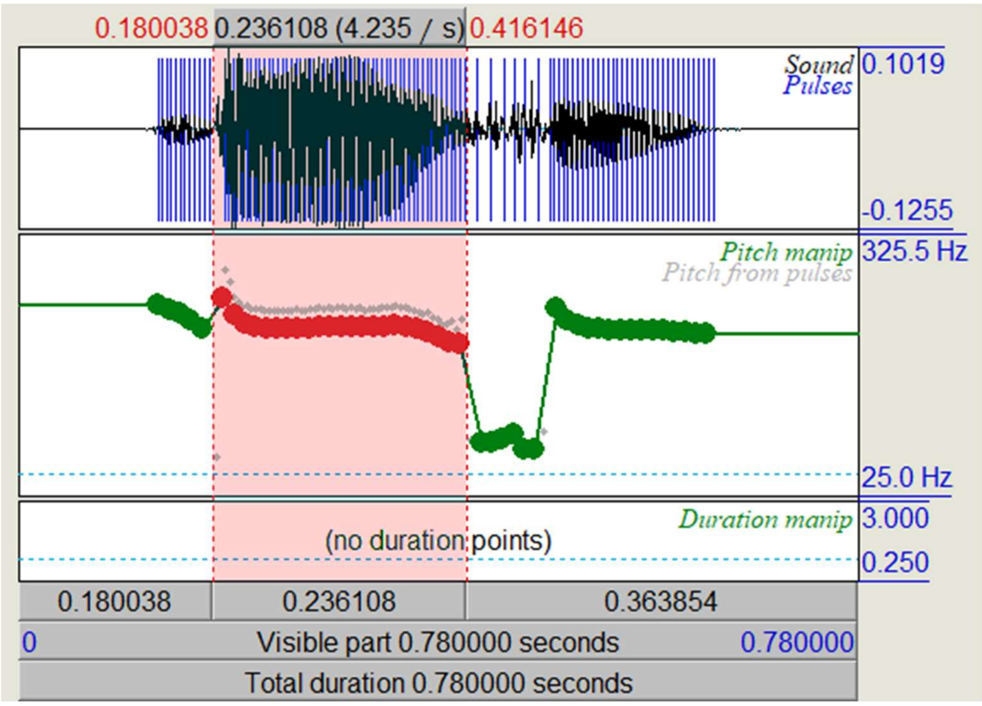


Figure 16: Pitch of the vowel portion of the stimulus /da/ shifted by -20 *Hz*

After manipulation, 560 tokens for the experiment were produced. Each of 60 recorded stimuli, with six stop consonants (/b/, /p/, /d/, /t/, /g/, and /k/) 5 vowels (/a/, /e/, /i/, /o/, and /u/), and two pitch patterns (HL, LH), was manipulated into six tokens, with one original token, four with its f_0 manipulated (-20 Hz, -10 Hz, $+10$ Hz, and $+20$ Hz), and one with its VOT manipulated, yielding 360 tokens in total. Each of 40 recorded fillers was manipulated in the same way, except without VOT manipulation, resulting in 200 tokens.

2-2. Subjects

40 Korean speakers were recruited for the perception test. Among them, 20 subjects were teenagers born in 2000 to 2002. All of them were born and raised in Seoul Metropolitan Area (Seoul, Gyeonggi, or Incheon) where non-tonal Seoul Korean is spoken. None of them has lived outside the area. The subjects have not learned any language other than English, and all of them were self-reportedly not fluent in English (intermediate level or under). Ten subjects were female and the other ten were male. All of them were recruited from Gwangmyeong High School in Gwangmyeong, Gyeonggi Province. 20 senior listeners in their 70s, born before 1947, were recruited from two senior centres in Gwanak District, Seoul, and from various random outdoor rest areas such as pavilions, city parks, and other

public spaces in Gwanak District, Seoul, and in Gwangmyeong, Gyeonggi Province. Twelve subjects were female, and eight were male. They too were born and raised in Seoul, Gyeonggi, or Incheon, and did not live outside the area during their childhood to adolescence. In their adulthood, they have not stayed outside the area for more than two years. All of them were self-reportedly Korean monolinguals.

All subjects reported no speech or hearing problems. They were all literate, able to read words and non-words written in Hangeul (Korean script) without difficulty.

2-3. Methods

2-3-1. Experiment design

An identification test was created using the Praat's multiple forced choice (MFC) listening experiment function. The 560 tokens, among which 360 are stimuli and 200 are fillers, were presented to the subjects in a random order, with three choice alternatives each showing a Hangeul syllable block, consisting of a consonant and a vowel. Although the order was randomized, it was avoided to have two consecutive tokens with homorganic consonants and the same vowel, such as a /di/ token next to a /ti/ token.

The listeners were instructed to choose the Hangeul letter for the first syllable of the word they heard. With tokens whose initial stop is bilabial, /b/ or /p/, choices with bilabial stops, 〈ㅂ〉, 〈ㅃ〉, or 〈ㅍ〉 were provided. With tokens whose first vowel is /a/, provided choices also contained the corresponding 〈ㅏ〉 vowel.

stimuli tokens	/b/, /p/	/d/, /t/	/g/, /k/
/a/	바, 빠, 파 ba, ppa, pa	다, 따, 타 da, tta, ta	가, 까, 카 ga, kka, ka
/e/	베, 뻬, 페 be, ppe, pe	데, 떼, 테 de, tte, te	게, 께, 케 ge, kke, ke
/i/	비, 삐, 피 bi, ppi, pi	디, 띠, 티 di, tti, ti	기, 끼, 키 gi, kki, ki
/o/	보, 뽀, 포 bo, ppo, po	도, 또, 토 do, tto, to	고, 꼬, 코 go, kko, ko
/u/	부, 뿌, 푸 bu, ppu, pu	두, 뚜, 투 du, ttu, tu	구, 꾸, 쿠 gu, kku, ku

Table 5: Choice sets provided to the subjects, according to the stimuli

For the filler tokens, the nasals and the liquid were grouped and provided with the corresponding nasal and liquid consonants, /m/, /n/, and /ɾ~l/, which form a natural category in Korean phonology. Likewise, the fricative tokens were provided with Korean fricatives, /s/, /s͈/, and /h/.

filler tokens	/m/, /n/, /l/	/s/
/a/	마, 나, 라 ma, na, ra	사, 싸, 하 sa, ssa, ha
/e/	메, 네, 레 me, ne, re	세, 쎄, 헤 se, sse, he
/i/	미, 니, 리 mi, ni, ri	시, 씨, 히 si, ssi, hi
/o/	모, 노, 로 mo, no, ro	소, 쏘, 호 so, sso, ho
/u/	무, 누, 루 mu, nu, ru	수, 쑤, 후 su, ssu, hu

Table 6: Choice sets provided to the subjects, for filler tokens

There was no time limit, and the replay was allowed once for each identification task, although the subjects were advised not to replay unless they did not hear the token clearly due to outside factors, such as noise. When the subjects were not sure of what they heard, they were advised to choose the closest answer.

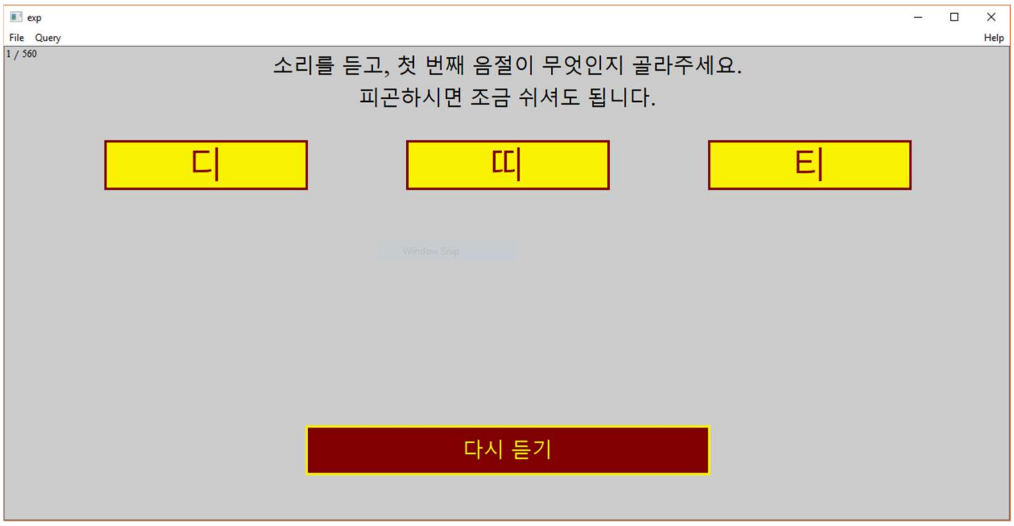


Figure 17: Praat's ExperimentMFC window showing the perception test

2-3-2. Performance

Teenage subjects were provided with a zip folder containing an executable Praat file (.exe), an MFC experiment file (.ExperimentMFC), a hidden folder containing the tokens in the waveform audio file format (.wav), and an instruction document (.PDF) with illustrative screenshots on how to open the MFS experiment file on the Praat window, and extract and save the results as a text file (.txt). They were advised to participate in this experiment in a quiet room using headphones or earphones. The result files were collected via e-mail.

Collecting the results via e-mail was not possible for the senior subjects. The author had to instruct each step one by one. Some senior subjects reported eye fatigue and/or arm pain during the test, in which case they were given a break, or the author clicked the button for them when they identified the sound they heard and told the author orally which of the presented choice alternatives they would choose.

3. Results

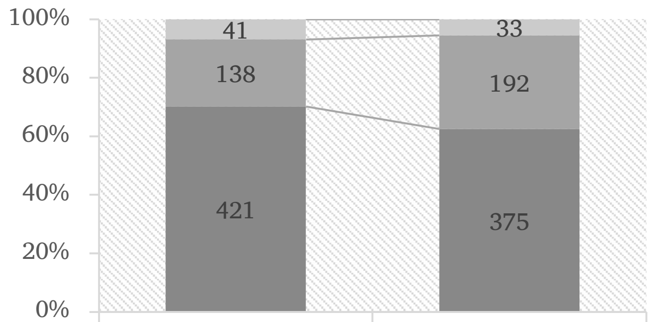
Each of 20 teenage listeners and 20 senior listeners listened to and categorized 360 stimuli. A total of 14,400 answers from were collected and analysed.

3-1. VOT contrasts

To see whether the listeners perceived consonants differently due to the addition or removal of voicing, χ^2 homogeneity test was conducted using R (version 4.1.0), with VOT manipulation as independent variables and heard consonant as the dependant variable. When the p-value is smaller than the significance level of 0.05, it is deemed that the listeners used VOT differences as an acoustic cue for distinguishing and categorizing the consonants.

3-1-1. Young listeners

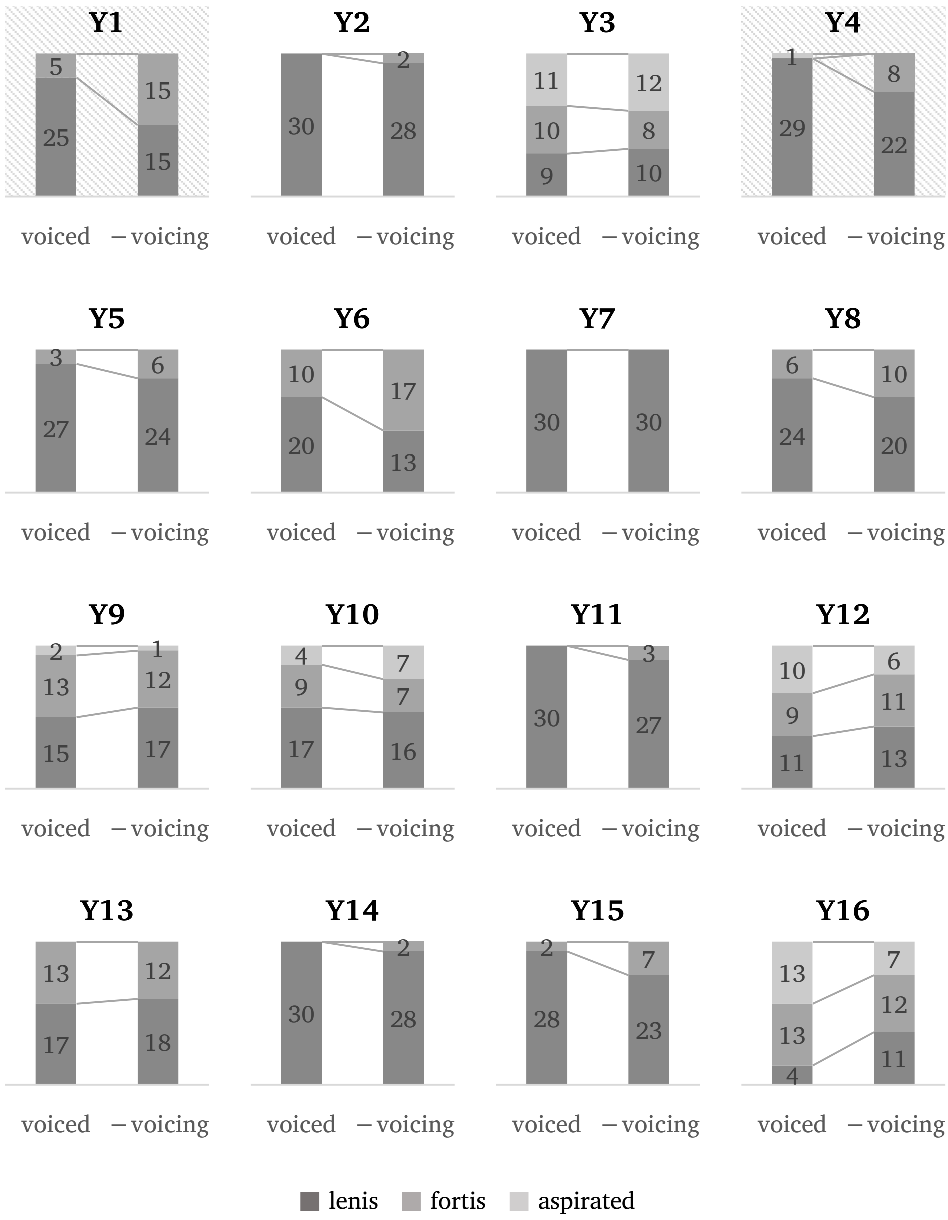
70% (421) of the 600 tokens with voiced, unmanipulated word-initial stops were perceived as lenis by 20 young listeners, while 23% (138) was perceived as fortis. 7% (41) was perceived as aspirated. For the tokens with prevoicing removed, the lenis categorization decreased to 63% (375), while the fortis categorization increased to 32% (192).



	voiced	– voicing
aspirated	41	33
fortis	138	192
lenis	421	375

Figure 18: Percentage of the tokens perceived as fortis, lenis, and aspirated by teenage listeners, by VOT (removed prevoicing)

However, χ^2 homogeneity test showed that only three individual listeners—Y1 ($p = 0.00617$), Y4 ($p = 0.006871$), and Y20 ($p = 0.000532$)—utilized the addition of prevoicing as an auditory cue. For all three of them, most of the fully voiced word-initial stops were perceived as lenis, and the percentage of the tokens categorized as fortis was increased after the removal of prevoicing.



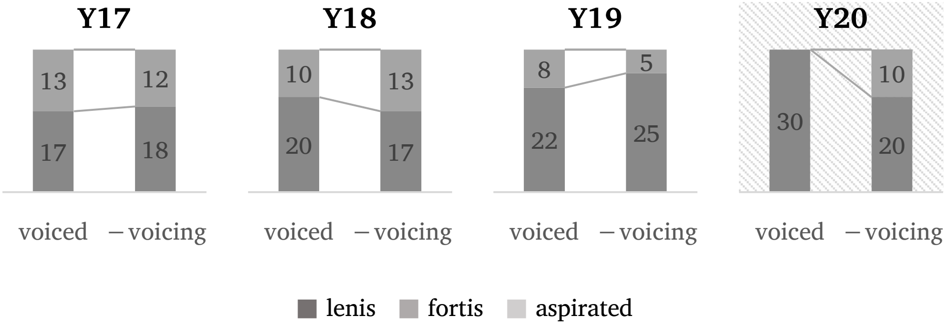
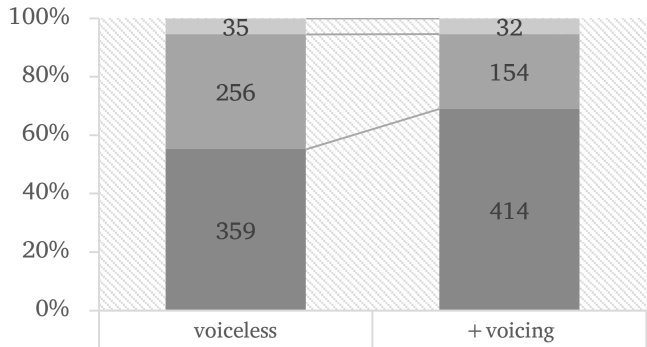


Figure 19: Percentage of the tokens perceived as fortis, lenis, and aspirated by individual teenage listeners, by VOT (removed prevoicing)

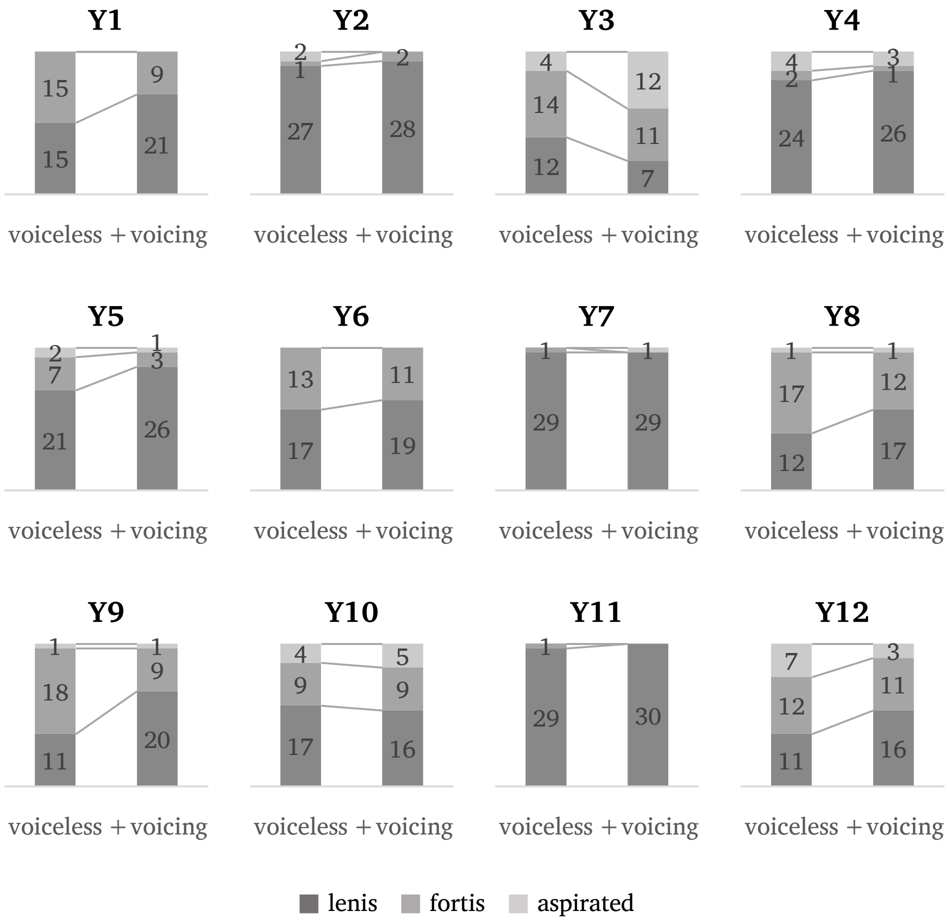
60% (359) of the 600 tokens with voiceless, unmanipulated word-initial stops were perceived as lenis, while 34% (256) was perceived as fortis. 6% (35) was perceived as aspirated. When prevoicing was added, the percentage of the tokens perceived as lenis increased to 69% (414), while percentage of the tokens perceived as fortis decreased to 27% (154).



	voiceless	+ voicing
aspirated	35	32
fortis	256	154
lenis	359	414

Figure 20: Percentage of the tokens perceived as fortis, lenis, and aspirated by teenage listeners, by VOT (added prevoicing)

However, χ^2 homogeneity test showed that only one individual listeners—Y15 ($p = 0.006526$)— utilized the addition of prevoicing as an auditory cue. For that one listener, voiceless word-initial stops were perceived as lenis and fortis, and the percentage of the tokens categorized as lenis was significantly increased after the addition of prevoicing.



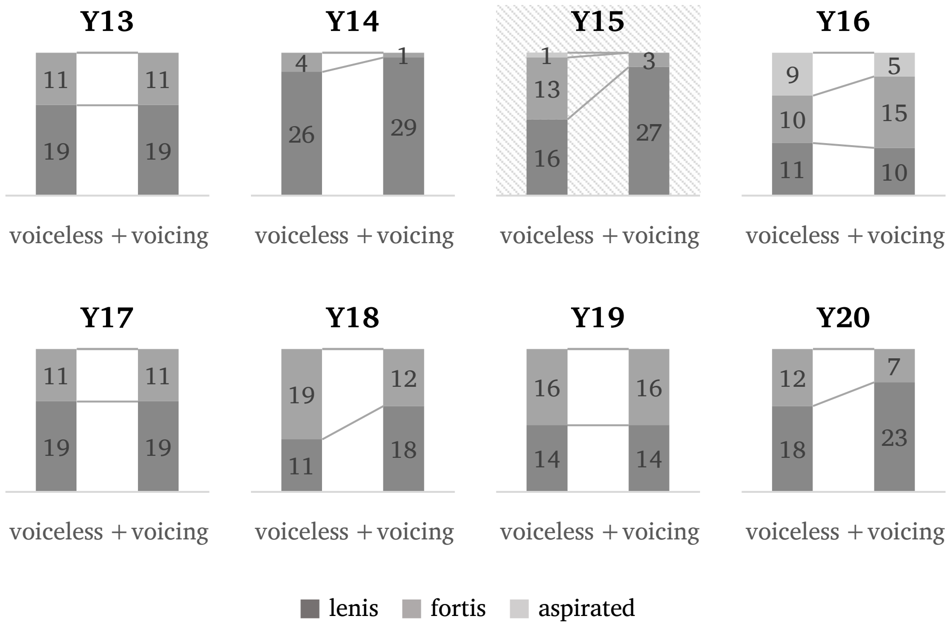
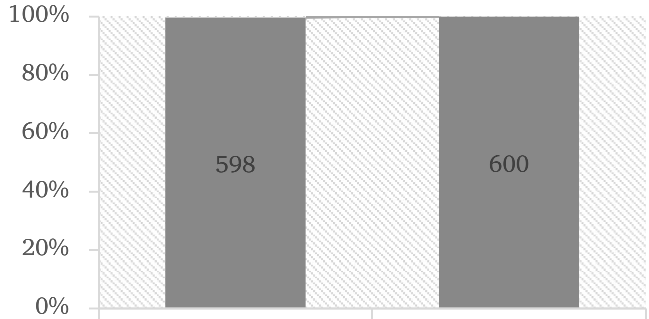


Figure 21: Percentage of the tokens perceived as fortis, lenis, and aspirated by individual teenage listeners, by VOT (added prevoicing)

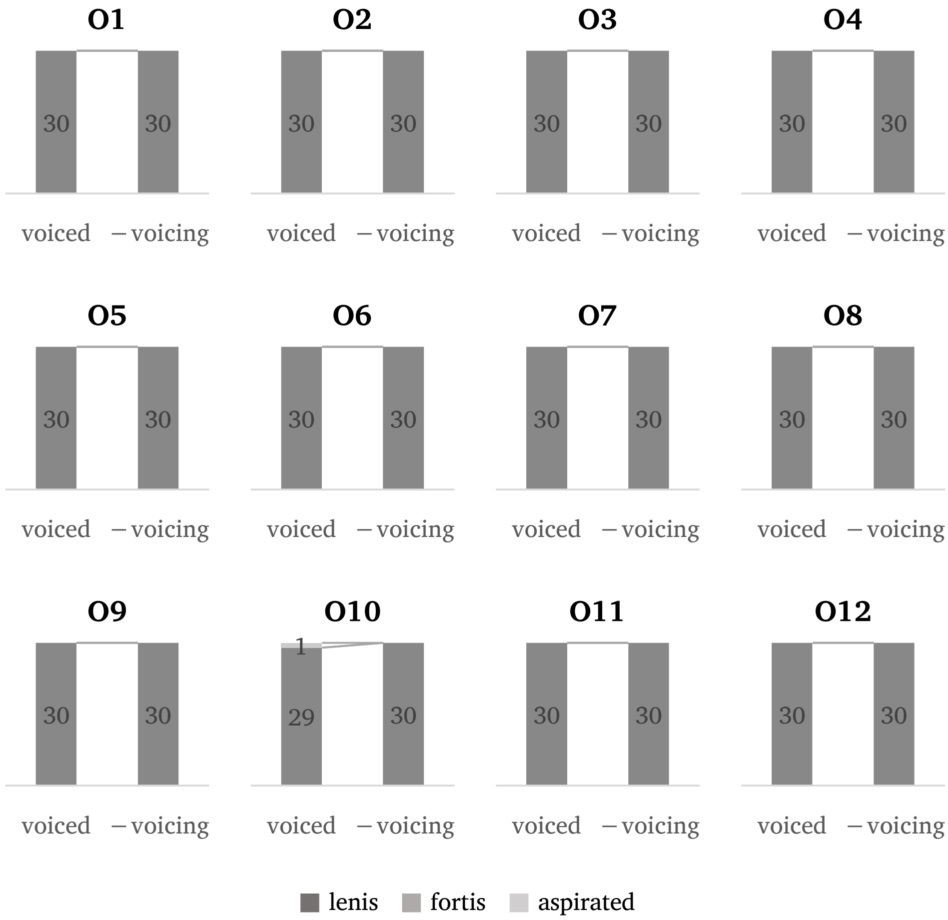
3-1-2. Senior listeners

Almost all (598) the 600 tokens with voiced, unmanipulated word-initial stops were perceived as lenis by senior listeners. Two subjects categorized one token each as aspirated. All 600 tokens with the prevoicing part removed were perceived as lenis. The χ^2 homogeneity test showed that the removal of prevoicing did not affect any of the 20 senior listeners' perception ($p > 0.05$ for all listeners).



	voiced	– voicing
aspirated	2	0
fortis	0	0
lenis	598	600

Figure 22: Percentage of the tokens perceived as fortis, lenis, and aspirated by senior listeners, by VOT (removed prevoicing)



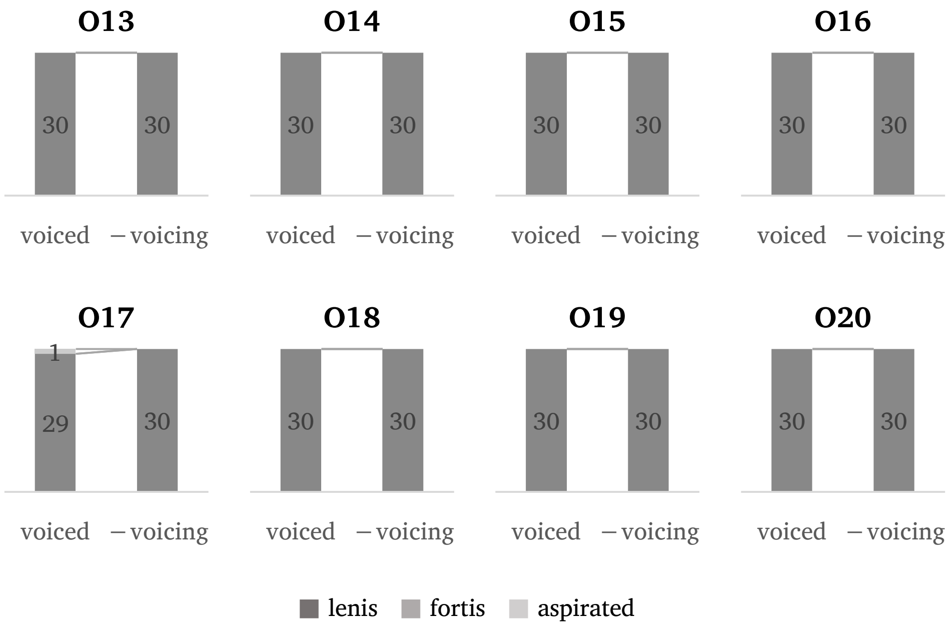
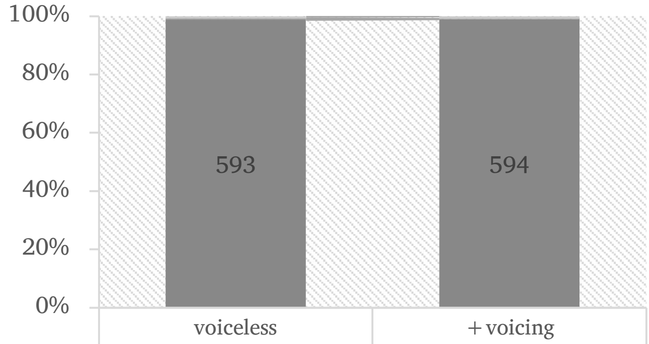


Figure 23: Percentage of the tokens perceived as fortis, lenis, and aspirated by senior listeners, by VOT (removed prevoicing)

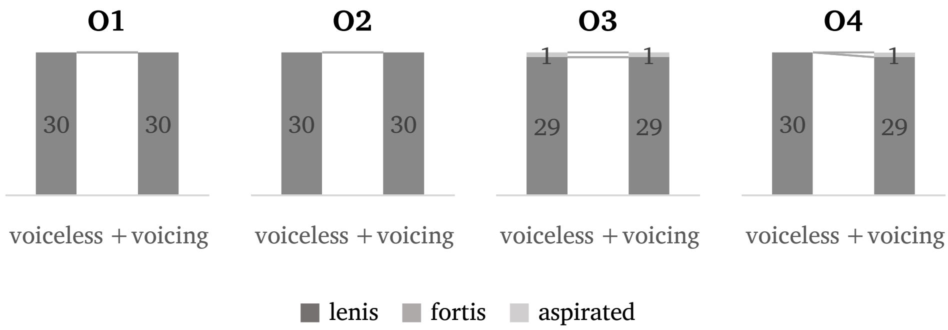
For senior listeners, almost all (593) the 600 tokens with voiceless, unmanipulated word-initial stops sounded like lenis as well. four tokens were categorized as aspirated, and three as fortis.



	voiceless	+voicing
aspirated	4	5
fortis	3	1
lenis	593	594

Figure 24: Percentage of the tokens perceived as fortis, lenis, and aspirated by senior listeners, by VOT (added prevoicing)

All subjects categorized at least 27 of the tokens as lenis. χ^2 homogeneity test showed that none of the 20 senior listeners utilized the addition of prevoicing as an auditory cue in their perception ($p > 0.05$ for all listeners).



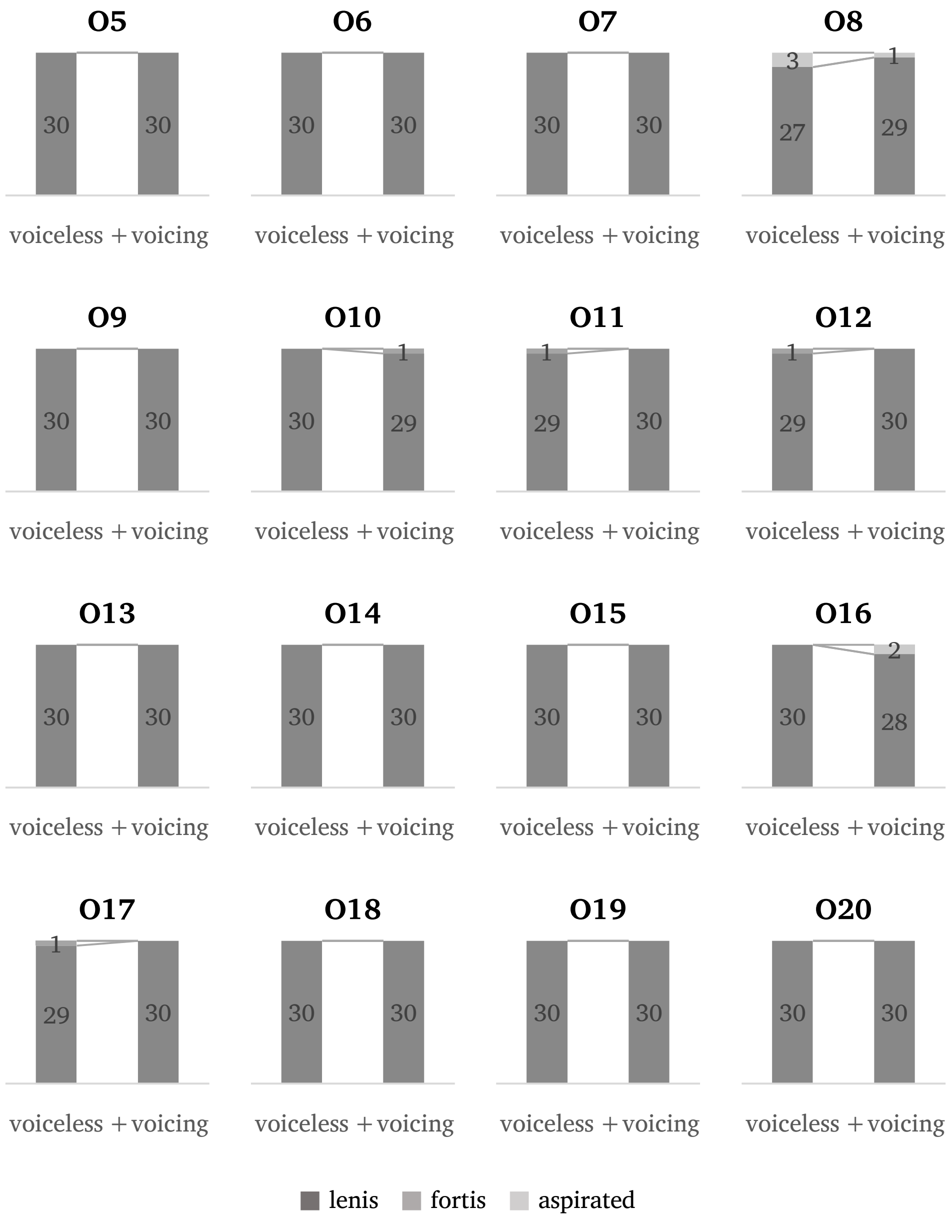


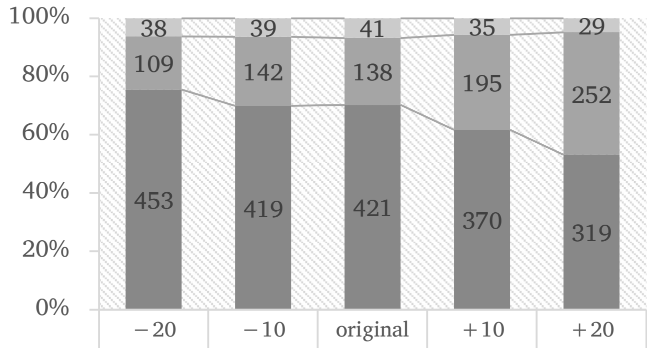
Figure 25: Percentage of the tokens perceived as fortis, lenis, and aspirated by individual senior listeners, by VOT (added prevoicing)

3-2. f_0 contrasts

To see whether the listeners perceived consonants differently due to the raised and lowered pitch of the following vowel, Spearman's correlation test was conducted using R (version 4.1.0), with f_0 manipulation and the number of answers “lenis” divided by the number of answers “fortis”. When the correlation coefficient ρ is positive and the *p*-value is smaller than the significance level of 0.1, it is deemed that the listeners used f_0 differences as an auditory cue for distinguishing and categorizing the consonants.

3-2-1. Young listeners

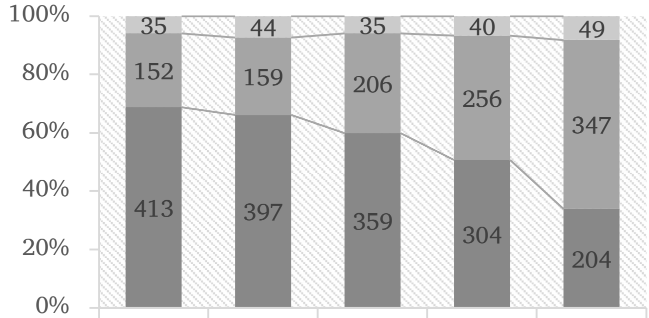
70% (421) of the 600 tokens with voiced tokens with unmanipulated f_0 were perceived as lenis by 20 young listeners, while 23% (138) was perceived as fortis. 7% (41) was perceived as aspirated. As the f_0 of the following vowels increases, more listeners identified the voiced tokens as fortis. When the f_0 of the first-syllable vowel was raised by 10 Hz, the number of tokens perceived as lenis decreased to 370 (−8%) and the number of tokens perceived as fortis increased to 195 (+10%). When the f_0 of the first-syllable vowel was raised by 20 Hz, the number of tokens perceived as lenis decreased to 319 (−17%) and the number of tokens perceived as fortis increased to 252 (+19%). Tokens with their pitch lowered were more likely categorized as lenis. When the f_0 of the first-syllable vowel was lowered by 10 Hz, the number of tokens perceived as fortis increased to 142 (+1%). When the f_0 of the first-syllable vowel was lowered by 20 Hz, the number of tokens perceived as lenis increased to 453 (+6%) and the number of tokens perceived as fortis decreased to 109 (−5%).



	−20	−10	original	+10	+20
aspirated	38	39	41	35	29
fortis	109	142	138	195	252
lenis	453	419	421	370	319

Figure 26: Percentage of the voiced tokens perceived as fortis, lenis, and aspirated by teenage listeners, by added and subtracted f_0 values (Hz)

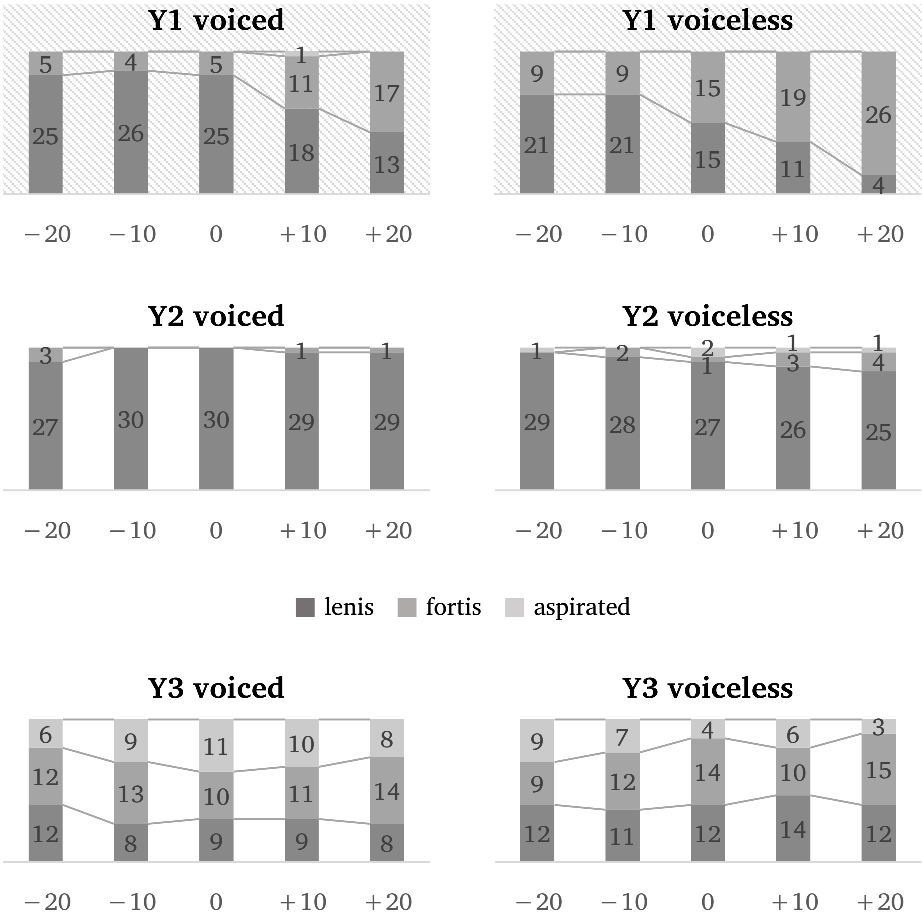
60% (359) of the 600 voiceless tokens with unmanipulated f_0 were perceived as lenis, while 34% (206) were identified as fortis by the young listeners. 6% (35) was perceived as aspirated. As the f_0 of the following vowels increases, more listeners identified the voiced tokens as fortis. When the f_0 of the first-syllable vowel was raised by 10 Hz, the number of tokens perceived as lenis decreased to 304 (−9%) and the number of tokens perceived as fortis increased to 256 (+9%). Tokens with their pitch lowered were more likely categorized as lenis. When the f_0 of the first-syllable vowel was raised by 20 Hz, the number of tokens perceived as lenis decreased to 347 (−2%). When the f_0 of the first-syllable vowel was lowered by 20 Hz, the number of tokens perceived as lenis increased to 413 (+9%) and the number of tokens perceived as fortis decreased to 152 (−10%).

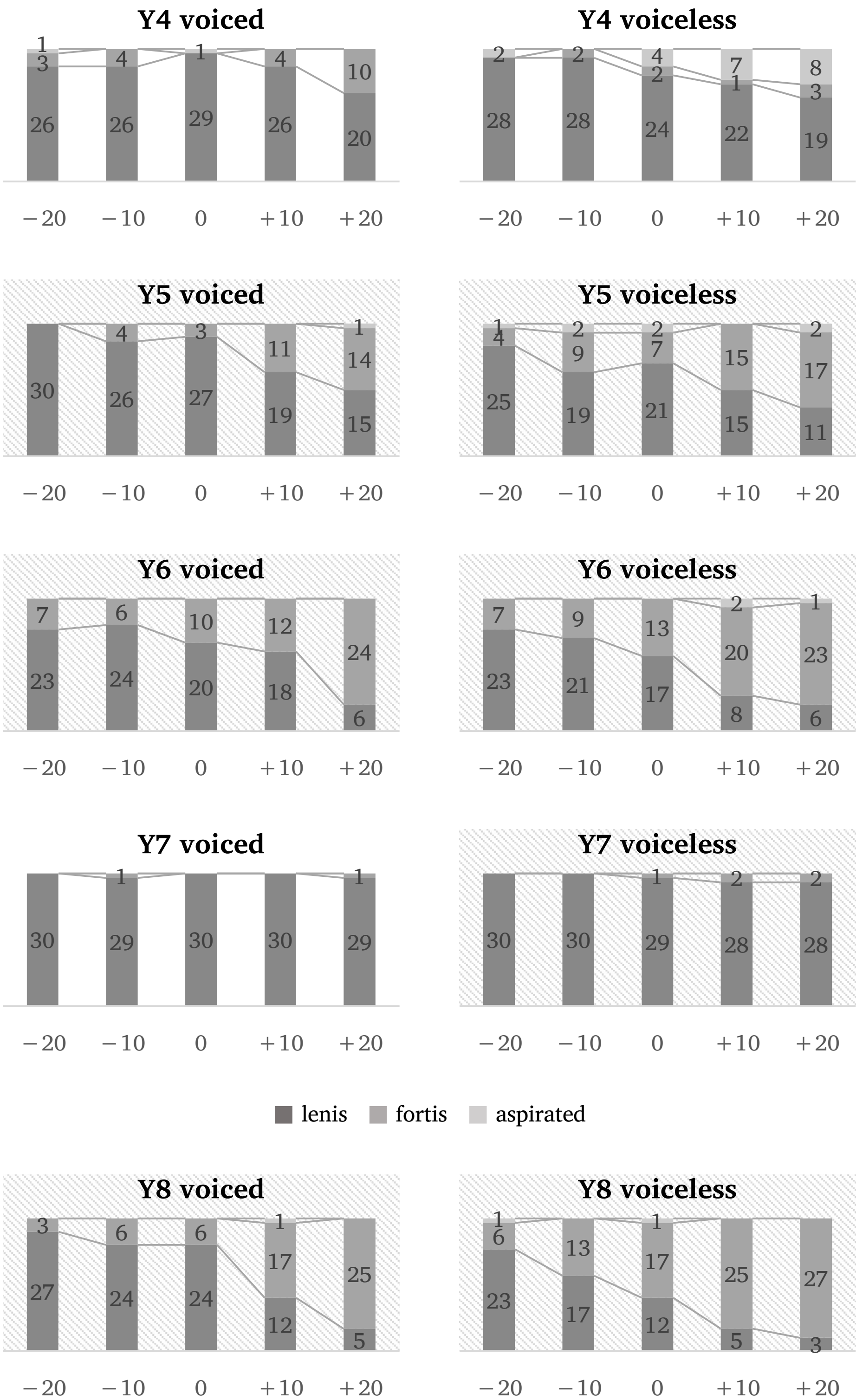


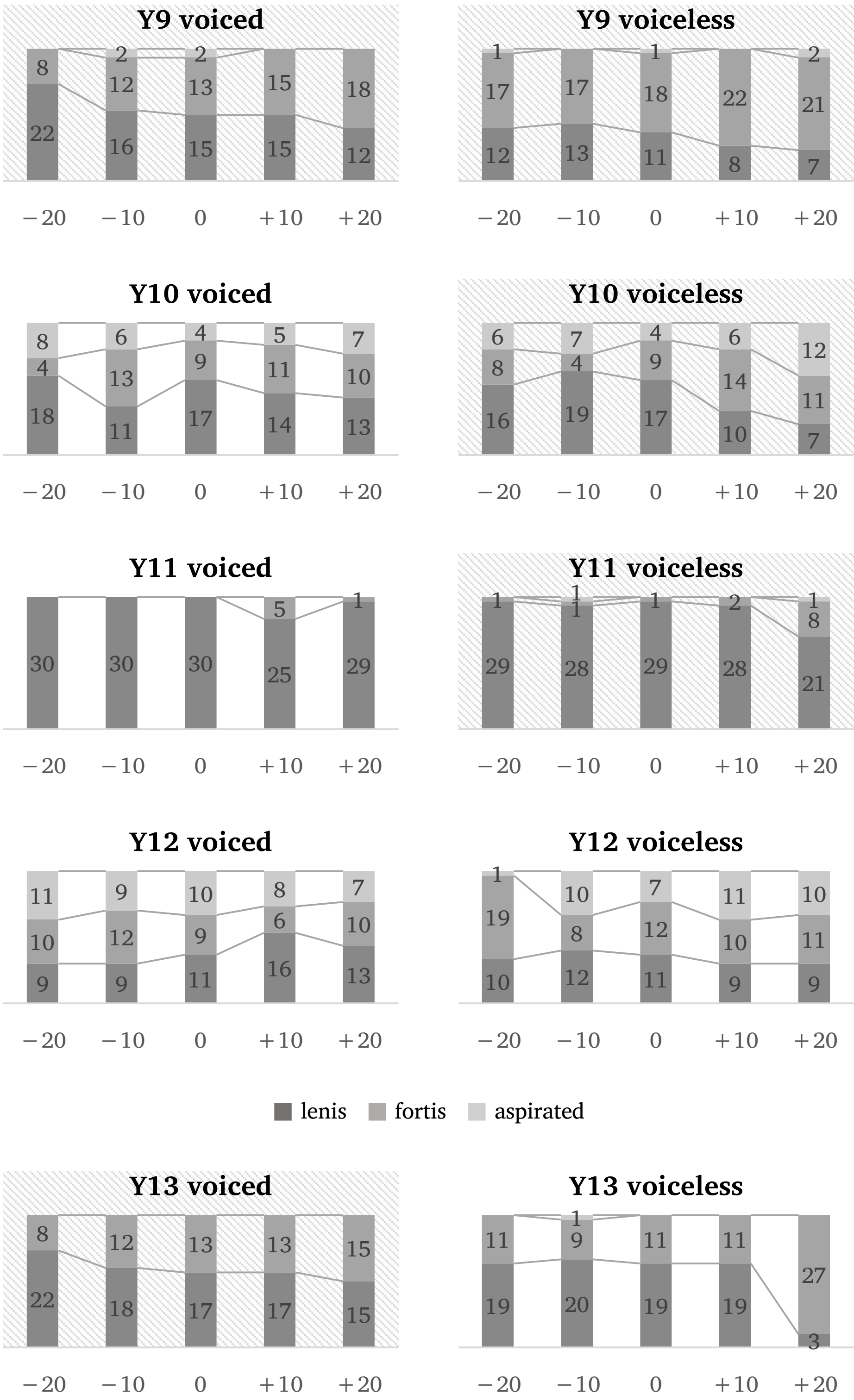
	−20	−10	original	+10	+20
aspirated	35	44	35	40	49
fortis	152	159	206	256	347
lenis	413	397	359	304	204

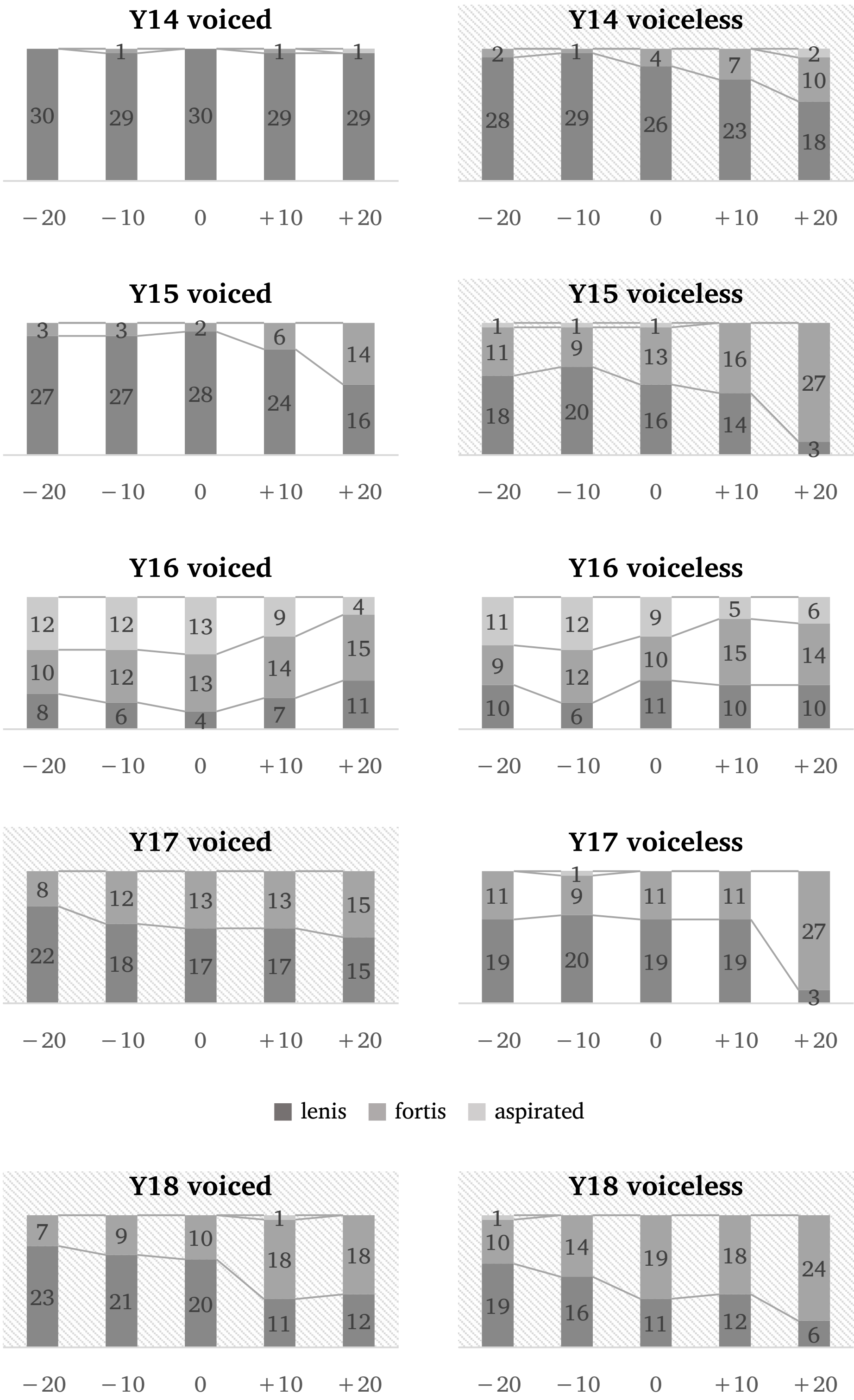
Figure 27: Percentage of the voiceless tokens perceived as fortis, lenis, and aspirated by teenage listeners, by added and subtracted f_0 values (Hz)

Spearman's correlation test showed that 10 individual listeners— Y1 ($\rho = 0.8207827$, $p = 0.08859$), Y5 ($\rho = 0.9$, $p = 0.08333$), Y6 ($\rho = 0.9$, $p = 0.08333$), Y8 ($\rho = 0.50641$, $p = 0.004818$), Y9 ($\rho = 1$, $p = 0.01667$), Y13 ($\rho = 0.9746794$, $p = 0.004818$), Y17 ($\rho = 0.9746794$, $p = 0.004818$), Y18 ($\rho = 0.9$, $p = 0.08333$), Y19 ($\rho = 1$, $p = 0.01667$), Y20 ($\rho = 0.8944272$, $p = 0.04052$)— utilized the f_0 differences in perceiving word-initial voiced stops, and 12 individual listeners—Y1 (ρ = 0.9746794, p = 0.004818), Y5 (ρ = 0.9, p = 0.08333), Y6 (ρ = 1, p = 0.01667), Y7 (ρ = 0.9486833, p = 0.01385), Y8 (ρ = 1, p = 0.01667), Y9 (ρ = 0.9, p = 0.08333), Y10 (ρ = 0.9, p = 0.08333), Y11 (ρ = 0.8207827, p = 0.08859), Y14 (ρ = 0.9, p = 0.08333), Y15 (ρ = 0.9, p = 0.08333), Y18 (ρ = 0.9, p = 0.08333), Y20 (ρ = 1, p = 0.01667)— utilized the f_0 differences as an auditory cue in perceiving word-initial voiceless stops.




Y4 voiced
1 3 26 4 26 1 29 4 26 10 20
−20 −10 0 +10 +20
Y4 voiceless
2 28 2 28 4 2 24 7 1 22 8 3 19
−20 −10 0 +10 +20
Y5 voiced
30 4 26 3 27 11 19 1 14 15
−20 −10 0 +10 +20
Y5 voiceless
1 4 25 2 9 19 2 7 21 15 15 2 17 11
−20 −10 0 +10 +20
Y6 voiced
7 23 6 24 10 20 12 18 24 6
−20 −10 0 +10 +20
Y6 voiceless
7 23 9 21 13 17 2 20 8 1 23 6
−20 −10 0 +10 +20
Y7 voiced
30 1 29 30 30 1 29
−20 −10 0 +10 +20
Y7 voiceless
30 30 1 29 2 28 2 28
−20 −10 0 +10 +20
lenis
fortis
aspirated
Y8 voiced
3 27 6 24 6 24 1 17 12 25 5
−20 −10 0 +10 +20
Y8 voiceless
1 6 23 13 17 1 17 12 25 5 27 3
−20 −10 0 +10 +20


Y9 voiced
Y9 voiceless
Y10 voiced
Y10 voiceless
Y11 voiced
Y11 voiceless
Y12 voiced
Y12 voiceless
lenis
fortis
aspirated
Y13 voiced
Y13 voiceless
−20 −10 0 +10 +20


Y14 voiced
Y14 voiceless
Y15 voiced
Y15 voiceless
Y16 voiced
Y16 voiceless
Y17 voiced
Y17 voiceless
lenis fortis aspirated
Y18 voiced
Y18 voiceless
−20 −10 0 +10 +20

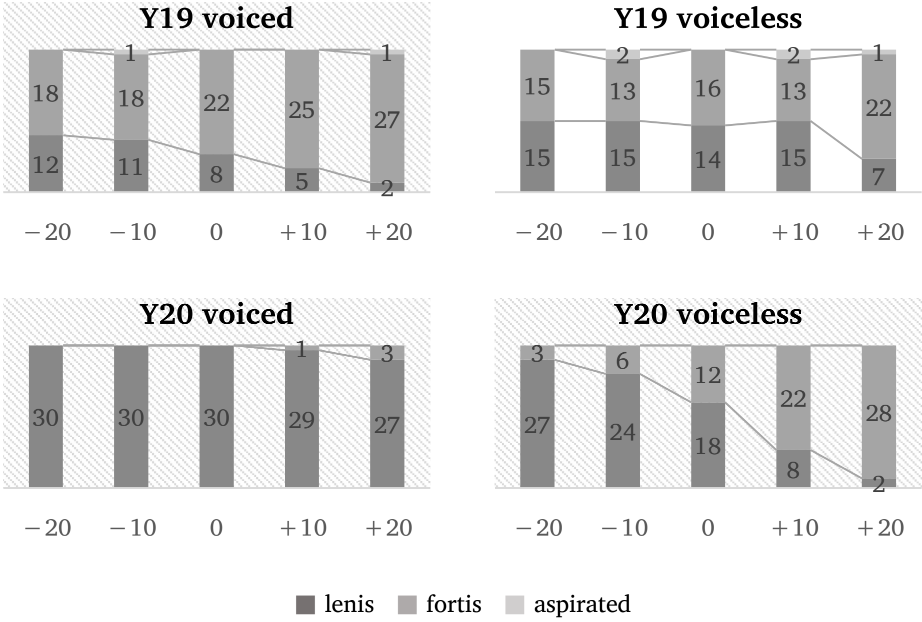
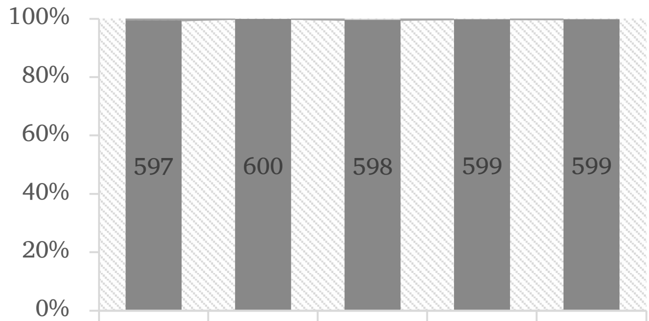


Figure 28: Percentage of the voiced and voiceless tokens perceived as fortis, lenis, and aspirated by individual teenage listeners, by added and subtracted f_0 values (Hz)

3-2-2. Senior listeners

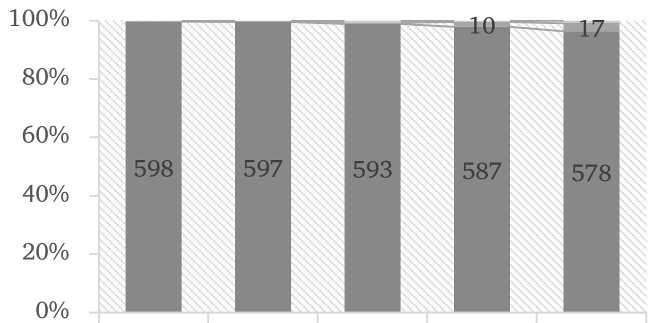
Again, almost all tokens with voiced consonants were perceived as lenis by the senior listeners, regardless of f_0 differences.



	−20	−10	original	+10	+20
aspirated	0	0	2	1	1
fortis	3	0	0	0	0
lenis	597	600	598	599	599

Figure 29: Percentage of the voiced tokens perceived as fortis, lenis, and aspirated by senior listeners, by added and subtracted f_0 values (Hz)

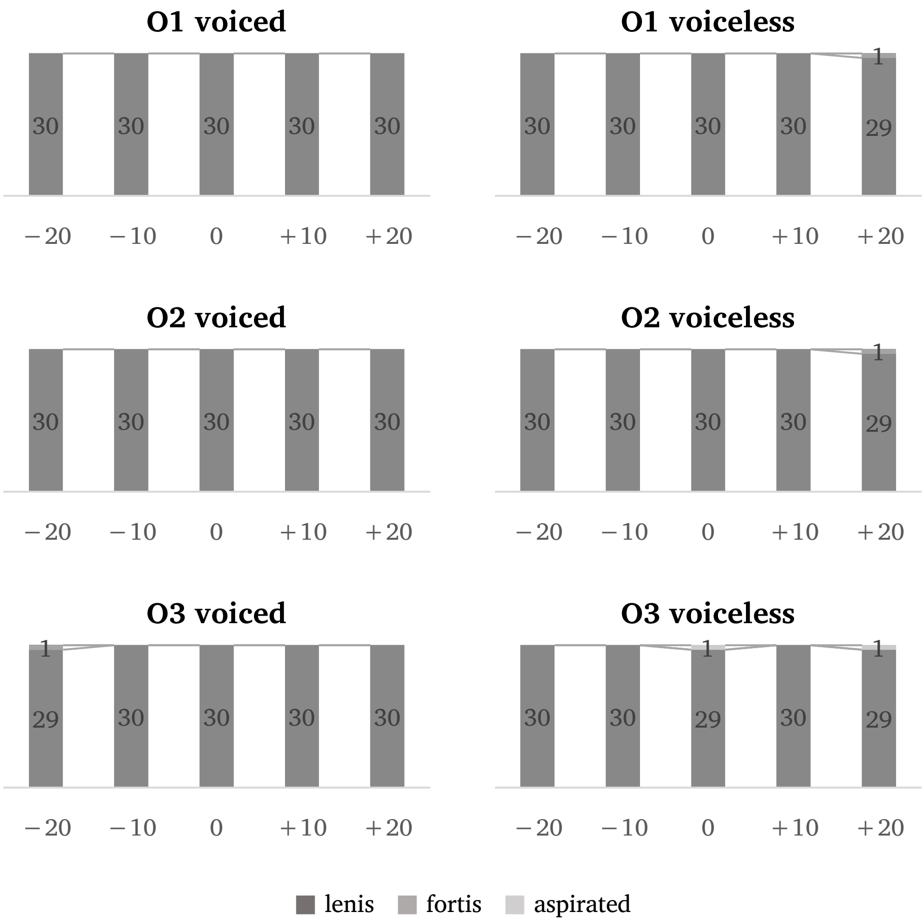
For voiceless stimuli, the number of tokens perceived as lenis decreased slightly with the raise of pitch of the following vowel.

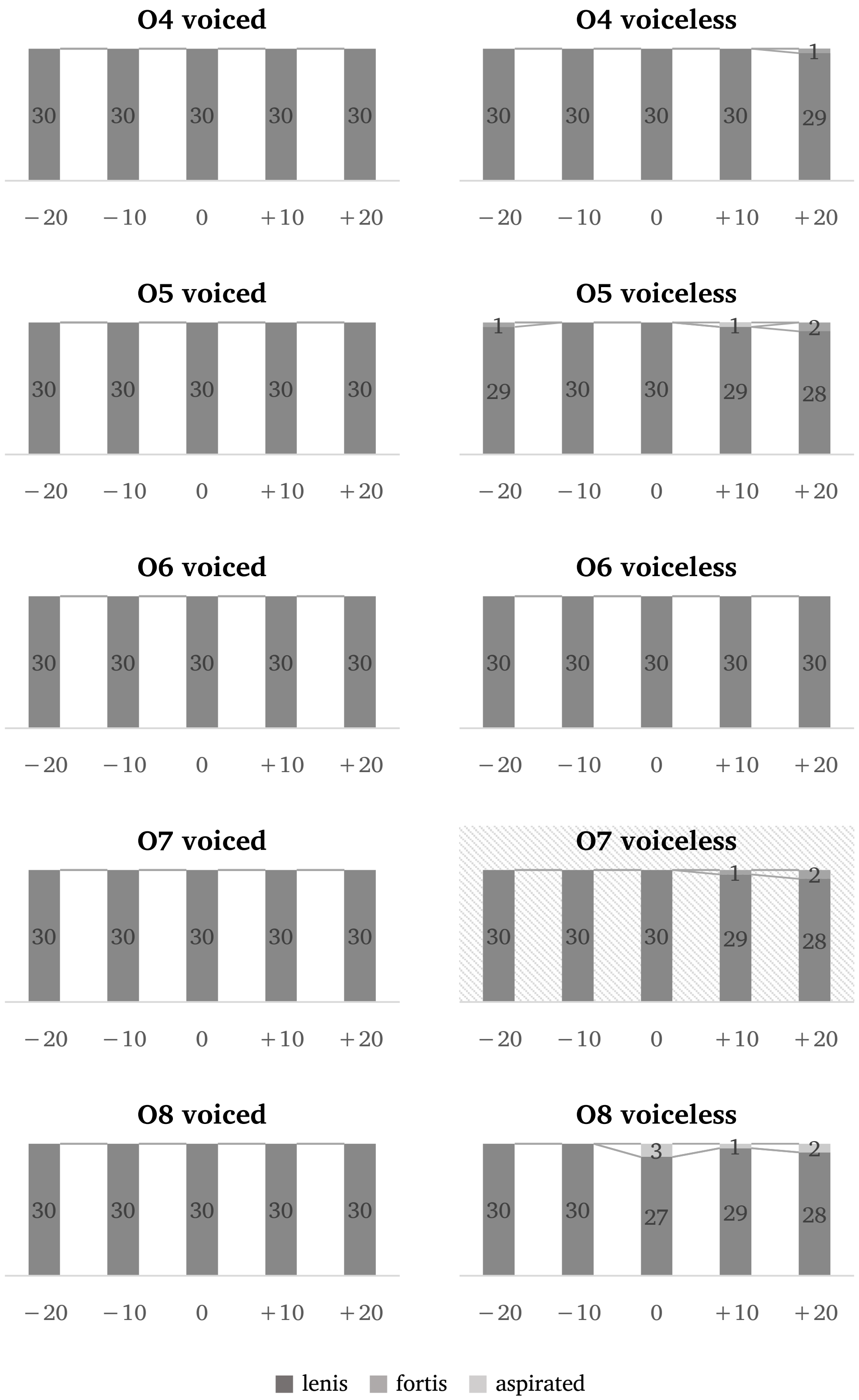


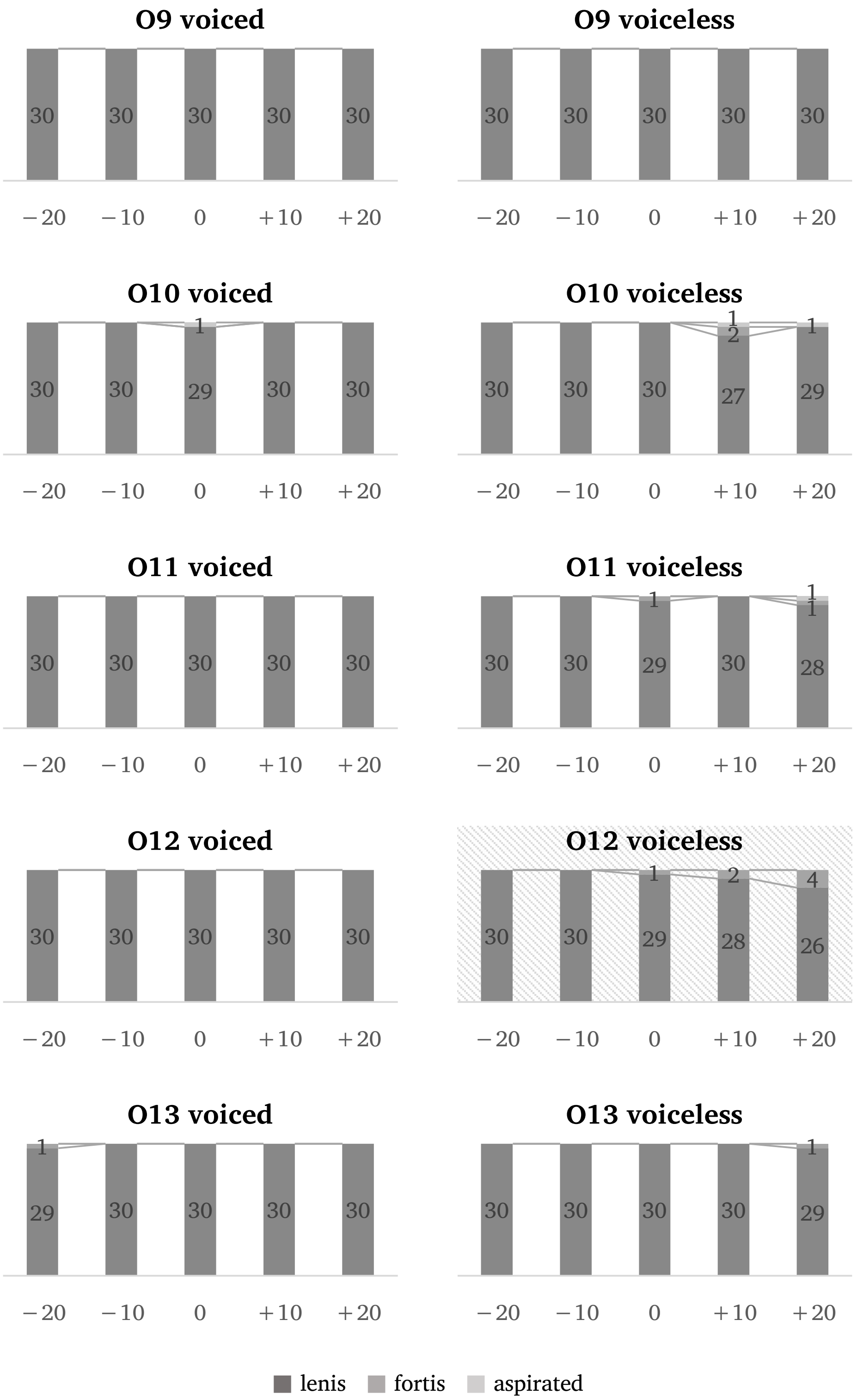
	−20	−10	original	+10	+20
aspirated	1	1	4	3	5
fortis	1	2	3	10	17
lenis	598	597	593	587	578

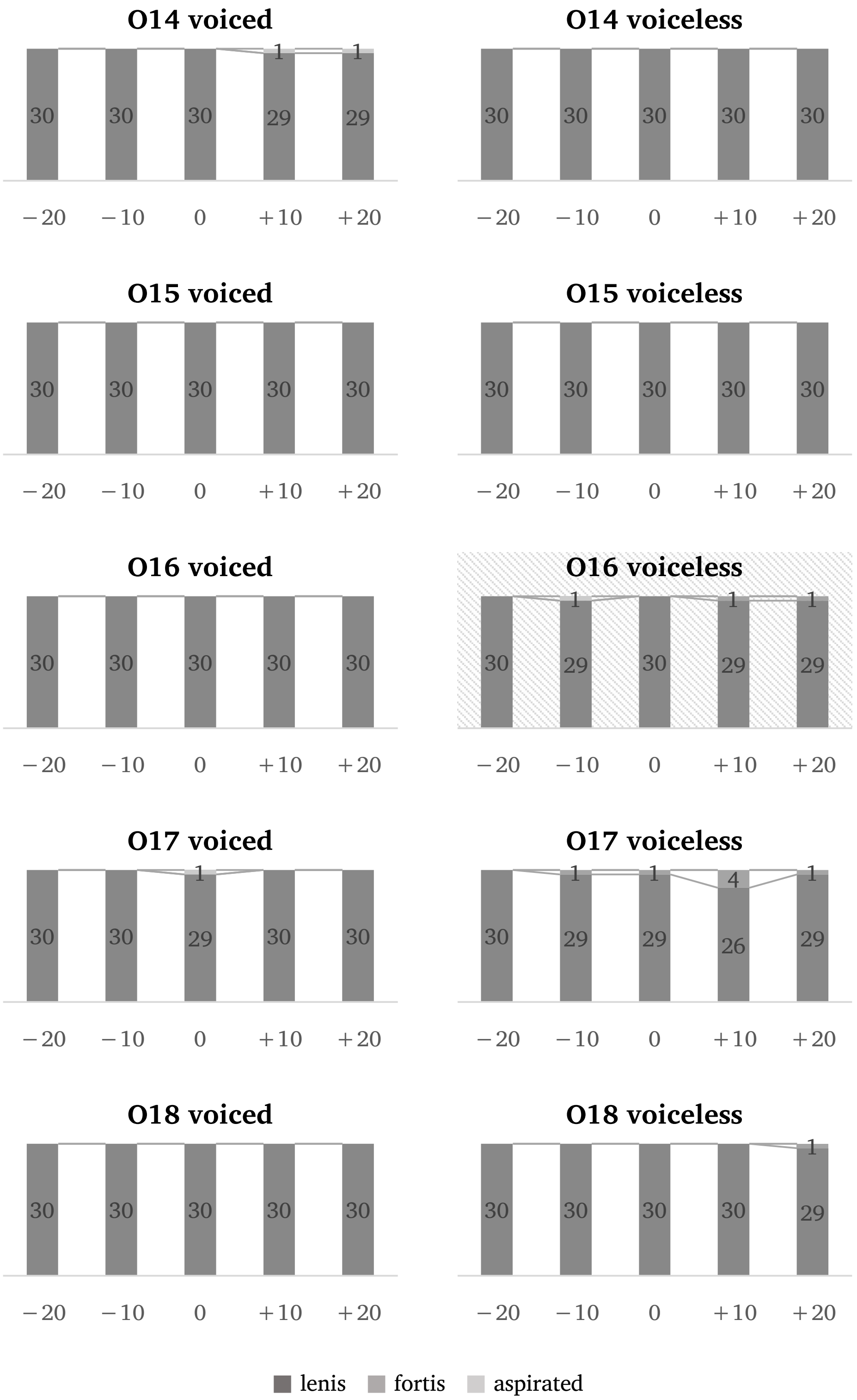
Figure 30: Percentage of the voiceless tokens perceived as fortis, lenis, and aspirated by senior listeners, by added and subtracted f_0 values (Hz)

Spearman's correlation test showed that none of the 20 individual listeners utilized the f_0 differences in perceiving word-initial voiced stops, while three individual listeners—O7 ($\rho = 0.8944272$, $p = 0.04052$), O12 ($\rho = 0.9746794$, $p = 0.004818$), O16 ($\rho = 0.8660254$, $p = 0.05767$)— utilized the f_0 differences as an auditory cue in perceiving word-initial voiceless stops.




O4 voiced
30 30 30 30 30
− 20 − 10 0 + 10 + 20
O4 voiceless
1
30 30 30 30 29
− 20 − 10 0 + 10 + 20
O5 voiced
30 30 30 30 30
− 20 − 10 0 + 10 + 20
O5 voiceless
1 1 2
29 30 30 29 28
− 20 − 10 0 + 10 + 20
O6 voiced
30 30 30 30 30
− 20 − 10 0 + 10 + 20
O6 voiceless
30 30 30 30 30
− 20 − 10 0 + 10 + 20
O7 voiced
30 30 30 30 30
− 20 − 10 0 + 10 + 20
O7 voiceless
1 2
30 30 30 29 28
− 20 − 10 0 + 10 + 20
O8 voiced
30 30 30 30 30
− 20 − 10 0 + 10 + 20
O8 voiceless
3 1 2
30 30 27 29 28
− 20 − 10 0 + 10 + 20
lenis
fortis
aspirated


O9 voiced
30 30 30 30 30
−20 −10 0 +10 +20
O9 voiceless
30 30 30 30 30
−20 −10 0 +10 +20
O10 voiced
1
30 30 29 30 30
−20 −10 0 +10 +20
O10 voiceless
1
2
1
30 30 30 27 29
−20 −10 0 +10 +20
O11 voiced
30 30 30 30 30
−20 −10 0 +10 +20
O11 voiceless
1
1
1
30 30 29 30 28
−20 −10 0 +10 +20
O12 voiced
30 30 30 30 30
−20 −10 0 +10 +20
O12 voiceless
1 2 4
30 30 29 28 26
−20 −10 0 +10 +20
O13 voiced
1
29 30 30 30 30
−20 −10 0 +10 +20
O13 voiceless
1
30 30 30 30 29
−20 −10 0 +10 +20
lenis
fortis
aspirated


O14 voiced
30 30 30 29 29
1 1
−20 −10 0 +10 +20
O14 voiceless
30 30 30 30 30
−20 −10 0 +10 +20
O15 voiced
30 30 30 30 30
−20 −10 0 +10 +20
O15 voiceless
30 30 30 30 30
−20 −10 0 +10 +20
O16 voiced
30 30 30 30 30
−20 −10 0 +10 +20
O16 voiceless
30 29 30 29 29
1 1 1
−20 −10 0 +10 +20
O17 voiced
30 30 29 30 30
1
−20 −10 0 +10 +20
O17 voiceless
30 29 29 26 29
1 1 4 1
−20 −10 0 +10 +20
O18 voiced
30 30 30 30 30
−20 −10 0 +10 +20
O18 voiceless
30 30 30 30 29
1
−20 −10 0 +10 +20
lenis
fortis
aspirated

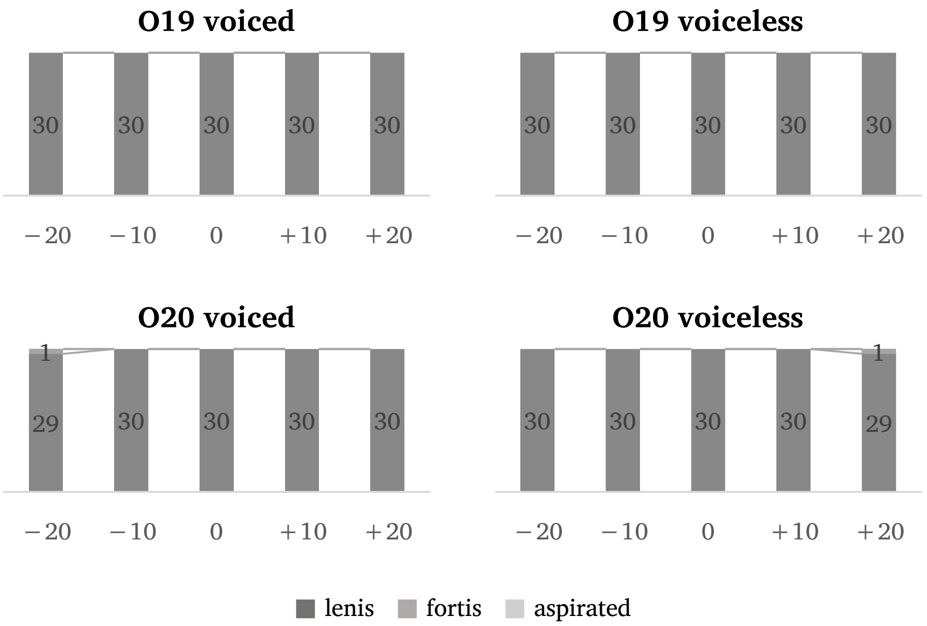


Figure 31: Percentage of the voiced and voiceless tokens perceived as fortis, lenis, and aspirated by individual senior listeners, by added and subtracted f_0 values (Hz)

3-3. Correlations

Seven teenage subjects—Y1, Y5, Y6, Y8, Y9, Y18, Y20— used f_0 as a differentiating cue for both voiced and voiced stops. Three teenage subjects—Y13, Y17, Y19— differentiated only the voiced stops using f_0, while five teenage subjects—Y7, Y10, Y11, Y14, Y15— and three senior subjects—O7, O12, O16— utilised f_0 as an auditory cure for the voiceless stops only.

All four subjects—Y1, Y4, Y15, Y20— who utilized the existence of prevoicing as an auditory cue also utilized f_0 variation as a cue, but not vice versa.

subject	VOT perception		f_0 perception	
	+voicing	−voicing	voiced	voiceless
Y1		✓	✓	✓
Y4		✓		
Y5			✓	✓
Y6			✓	✓
Y7				✓
Y8			✓	✓
Y9			✓	✓
Y10				✓
Y11				✓
Y13			✓	
Y14				✓
Y15	✓			✓
Y17			✓	
Y18			✓	✓
Y19			✓	
Y20		✓	✓	✓
O7				✓
O12				✓
O16				✓

Table 7: Auditory cue utilization by subjects

4. Conclusions

The main goal of the current study was to determine how Korean speakers perceive foreign unaspirated voiceless and voiced stops, and how the mechanisms of stop perception differ across generations.

The perception test was able to confirm that older and younger listeners' cue-weighting in stop perception differ, as the teenage listeners utilized the

f_0 of the following vowel more actively as an auditory cue. The pattern is consistent with their native stop perception mechanism, in which they use low f_0 as the primary cue for the lenis. Prevoicing, which the Korean language lacks, was found less of a differentiating factor in perceiving foreign stops by young Korean listeners, although a few subjects did hear the difference when prevoicing was removed or added to the stimuli.

Older Korean listeners, who were in previous studies reported to depend more on positive VOT than f_0 when identifying native stop consonants, did not use negative VOT more actively than younger listeners when categorizing unaspirated stops. The result implies that negative VOT, unlike positive VOT, does not trade off with f_0 in Korean speakers' perceptual mechanism. Further experimental work is required to verify if the senior speakers of the Korean language do use aspiration more actively as an auditory cue compared to the younger listeners.

Korean listeners' demonstrated tendency of mapping more unaspirated voiceless and voiced stops to lenis may be accounted for by some of the laryngeal configurations and/or other acoustic characteristics of the European Portuguese stops, or the possible speaker-specific factors present in the recordings used in this study. A further study with more focus on factors other than VOT and f_0 that influence Korean listeners' categorization of foreign stop consonants is therefore suggested.

Although this study focuses on foreign stop perception, the findings contribute to our general understanding of Korean listeners' stop perception and its generational changes, by providing experimental evidence in support of the surge of f_0 as the dominant cue in younger Koreans' stop perception. The findings also provide support for the influence of L1 cue-weighting patterns to Korean listeners' perception of foreign-language stops.

References

Ahn, H. (1999). *Post-release phonatory processes in English and Korean: Acoustic correlates and implications for Korean phonology*. The University of Texas at Austin.

Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.15. Retrieved March 14, 2017, from http://www.praat.org/

Cho, T. (1996). *Vowel correlates to consonant phonation: An acoustic-perceptual study of Korean obstruents*. The University of Texas at Arlington.

Cho, T., Jun, S., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30*(2), 193–228.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*(2), 155–190.

Choi, H. (2002). Acoustic Cues for the Korean Stop Contrast - Dialectal Variation. *ZAS Papers in Linguistics, 28*(28), 1–12.

Cruz-Ferreira, M. (1995). European Portuguese. *Journal of the International Phonetic Association, 25*(02), 90.

Dart, S. N. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. *The Journal of the Acoustical Society of America, 81*(1), 138–147.

Han, J.-I. (1996). Perception of Korean Tense and Lax Consonants: Evidence for a Geminate Analysis of Tense Consonants. In N. Akatsuka, S. Iwasaki, & S. Strauss (Eds.), *Japanese/Korean Linguistics* (pp. 407–424). Los Angeles; CA: CSLI Publications.

Han, M. S., & Weitzman, R. S. (1970). Acoustic Features of Korean /P, T, K/, /p, t, k/ and /ph, th, kh/. *Phonetica, 22*(2), 112–128.

Hardcastle, W. J. (1973). Some observations on the tense–lax distinction in initial stops in Korean. *Journal of Phonetics, 1*, 263–271.

Hirose, H., Lee, C. Y., & Ushijima, T. (1974). Laryngeal control in Korean stop production. *Journal of Phonetics, 2*(2), 145–152.

Hirose, H., Park, H. S., & Sawashima, M. (1983). Activity of the thyroarytenoid muscle in the production of Korean stops and fricatives. *Annual Bulletin Research Institute of Logopedics and Phoniatrics, 17*, 73–81.

Jun, S. (1993). *The phonetics and phonology of Korean prosody*. The Ohio State University.

Jun, S. (1996). Influence of microprosody on macroprosody: a case of phrase initial strengthening. *UCLA Working Papers in Phonetics, 92*, 97–116.

Jun, S., Beckman, M. E., & Lee, H.-J. (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics*, 43–68.

Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics, 2*(2), 161–180.

Kang, Y. (2008). English voiced stops in Korean in the 1930s. In *A talk presented at the annual meeting of the Montreal-Ottawa-Toronto (MOT) Phonology workshop*.

Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics, 45*(1), 76–90.

Kang, Y., & Han, S. (2013). Tonogenesis in early Contemporary Seoul Korean: A longitudinal case study. *Lingua, 134*, 62–74.

Kim, C.-W. (1965). On the Autonomy of the Tensity Feature in Stop Classification (with Special Reference to Korean Stops). *WORD, 21*(3), 339–359.

Kim, C.-W. (1970) A Theory of Aspiration, *Phonetica, 21*(2), 107–116

Kim, M.-R. (1994). *Acoustic characteristics of Korean stops and perception of English stop consonants*. University of Wisconsin–Madison.

Kim, M.-R. (2000). *Segmental and tonal interactions in English and Korean: a phonetic and phonological study*.

Kim, M.-R. (2016). Tonogenesis in Korean: some recent speculations on the sound change. *Eoneo. 37*(2), 243–283.

Kim, M.-R. (2017). Inter-speaker and intra-speaker variability on sound change in contemporary Korean. *Phonetics and Speech Sciences, 9*(3), 25–32.

Kim, M.-R., Beddor, P. S., & Horrocks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics, 30*(1), 77–100.

Kim, M. (2004). *Correlation between VOT and f_0 in the perception of Korean stops and affricates*. Seoul National University.

Lee, H., Politzer-Ahles, S., & Jongman, A. (2013). Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics, 41*(2), 117–132.

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD, 20*(3), 384–422.

Park, H. (2002). *Temporal and spectral characteristics of Korean phonation types*. University of Texas at Austin.

Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *The Journal of the Acoustical Society of America, 99*(5), 3201–3211.

Shimizu, K. (1990). *Cross-language study of voicing contrasts of stop consonants in Asian languages*. University of Edinburgh.

Shinohara, S., Ji, S. R., Ooigawa, T., & Shinya, T. (2011). The limited role of perception in Korean loanword adaptation: The Korean three-way laryngeal categorization of Japanese, French, English and Chinese plosives. *Lingua, 121*(9), 1461–1484.

Silva, D. J. (1992). *The Phonetics and Phonology of Stop Lenition in Korean.* Cornell University.

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology, 23*(02), 287–308.

국문 초록

세대에 따른 한국어 화자의 외국어 어두 유성 폐쇄음 지각 변화

서울대학교 대학원

언어학과 음성학전공

장현정

이 연구는 지각 실험을 통해 한국어 화자의 어두 무기 유·무성음 지각 기제와 그 변화 양상을 밝히고자 한다. 서울말 화자가 어두 예사소리를 구분할 때, 과거에는 성대 진동 시작 시간(VOT)이 더 주요한 단서였으나 최근에는 후행 모음의 기본 주파수(f_0)가 더 적극적으로 활용된다. 거센소리와 예사소리 지각에서는 f_0와 VOT 의 트레이드오프(trade-off) 관계가 보고되기도 했다. 이 연구에서는 한국어 화자의 어두 무성 폐쇄음 지각 양상에서 보이는 단서 트레이드오프와 단서 비중 변화가 유성 폐쇄음 지각에서도 나타나는지 살펴보기 위해 10 대 청소년 화자와 70 대 노인 화자를 대상으로 청취 실험을 시행해, 외국어

어두 무기 유·무성음 인지에서 세대 차가 있음을 밝혔다. 젊은 한국어 화자는 외국어 유·무성 폐쇄음을 인지할 때 이전 세대보다 후행 모음의 f_0를 더 적극적으로 활용하며, 이는 이들의 모어 무성 폐쇄음 인지 기제와 일치한다. 한국어 화자의 어두 폐쇄음 지각에서 f_0의 단서 비중 확대는 기식성이나 성대 진동 여부, 모어와 외국어 소리에 상관없이 두루 나타나는 것으로 보인다. 한국어 화자의 모어 무성 유·무기 폐쇄음 지각에서 관찰되었던 VOT 와 f_0 간 트레이드오프는 외국어 무기 유·무성음 폐쇄음 지각에서는 관찰되지 않았다. 모어의 어두 폐쇄음에서 유·무성 대립이 나타나지 않는 한국어 화자의 폐쇄음 지각 기제에서, 음수 VOT 는 양수 VOT 와 다른 성격을 지니는 것으로 보인다.

주요어: 세대 간 변화, 한국어 화자, 폐쇄음 지각, 성대 진동 시작 시간, VOT, 기본 주파수, f_0, 모음 높낮이

학번: 2015-20066