The Journal of Studies in Language

The Journal of Studies in Language 34(3), 505-518 (2018)

Production and Perception of F0 by Three-Year-Old Children

Gayeon Son

Kwangwoon University

ABSTRACT

The Journal of Studies in Language **34.3, 505-518**. This study explores how young Korean children produce and perceive stop contrasts in the phonetic dimension of F0. The Korean spoken in Seoul is undergoing tonogenetic sound change in which there has been a loss of VOT differentiation in young adults' speech, while the role of fundamental frequency (F0) in enhancing Korean stop contrasts has been required. To investigate the interrelation between children's production and perception of F0, the experiments were conducted on three-year-old children. Phonetic analysis showed that lenis and aspirated stops have no significant VOT differences, whereas a significant relationship with F0 was found. The results suggest that the acquisition of F0 plays a crucial role in the formation of phonemic categories for lenis and aspirated stops and this process significantly affects articulatory distinction. **(Kwangwoon University)**

Keywords: VOT, F0, child language, production, perception

https://doi.org/10.18627/jslg.34.3.201811.505

pISSN : 1225-4770

Received: October 01, 2018

Revised: October 30, 2018

Accepted: November 14, 2018

1. Introduction

Children acquire stop sounds first cross-linguistically, and to differentiate stop categories, a universal acoustic parameter, Voice Onset Time (VOT) has been used across languages (Lisker and Abramson, 1964). VOT represents the temporal relationship between the oral constriction release and the vocal folds vibration. There are effectively three stop phonation types that differ in VOT, particularly when a stop sound occurs utterance-initially: lead-voice (voicing), short lag (voiceless unaspirated), and long lag (voiceless aspirated) (Klatt, 1975; Ladefoged and Maddieson, 1996; Lisker and Abramson, 1964). Thus, VOT is a key phonetic parameter to investigate children's developmental patterns regarding stop distinctions and has been used to study the process of stop acquisition in relation to the articulatory achievements of young children (e.g., Allen, 1985; Clumeck, Barton, Macken, and Huntington, 1981; Gandour et al., 1986; Kewly-Port and Preston, 1974; Pan, 1994).

In spite of a number of studies on stop distinction in children, little research has reported on the acquisition of Korean stops. Korean has an unusual three-way contrast known as lenis, fortis, and aspirated, which are all pulmonic egressive voiceless stops.

For the Korean stop distinction, VOT as a determinant acoustic cue and fundamental frequency (F0) as a secondary acoustic measure have been used. F0 is usually used a prosodic term, but as in F0 perturbation, a correlation between the feature [voice] of a word-initial consonant and F0 can be found, which makes F0 useful for identifying voicing feature of a consonant (Haudricout, 1954; Hombert, Ohala, and Ewan, 1979). In Korean, F0 is a language-specific phonological component since it makes a phonemic contrast (Jun, 1996). The traditional role of VOT has been changed recently, since Korean stop contrasts are undergoing tonogenesis, in which stop sounds have tonal contrasts in a word-initial position. Regarding on-going tonogenetic changes in Seoul Korean, phonetic trade-off between VOT and F0 has been consistently reported in adult speech (Kang, 2014; Silva, 2006; Wright, 2007). That is, the role of VOT has diminished while F0 enhancement has increased. More specifically, for fortis stops, the shortest VOT can distinguish them from non-fortis stops, while lenis and aspirated stops have much longer VOT values. For the phonemic distinction between lenis and aspirated stops, F0 differentiation has been emphasized, since VOT rarely functions for the phonetic distinction between them. Lenis stops usually have relatively lower F0s and aspirated stops have the highest F0s, whereas fortis stops have intermediate values (e.g., Kang, 2014, among others).

This on-going sound change has been revealed in a few child studies. Kim (1999) examined VOT ranges by four child speakers. His target age was from 5 to 7 years, and his analysis showed that VOT values overlapped especially between lenis and aspirated stops. Kim suggested that VOT differences cannot distinguish between the stop categories, and other phonetic parameters such as stop closure duration or amplitude would be required for the valid differentiation. More recently, Kim and Stole-Gammon (2009) tried to illustrate the developmental pattern in the acquisition of Korean word-initial stops with a large body of data from 40 young children aged 2;6 to 4;0 (years;months). Children showed significantly short VOTs for fortis stops while no significant VOT differences between aspirated and lenis stops were observed except for the three-year-olds, who showed significant VOT differences for all stop categories. Focusing on the order of mastery among the Korean stop contrasts by children, Kong et al. (2011) revealed the relation of multiple phonetic properties of Korean stops to production and perception. Their hypothesis was based on the idea that there is a certain order to mastering the laryngeal features of Korean stop consonants, and this order affects perception by adults, too. The production of the children from 2;0 to 5;11 indicated that VOT cannot be used to differentiate aspirated from lenis stops. Instead, F0 served as a critical cue to distinguish lenis from aspirated stops. The findings of this phonetic analysis are supported by the results of a perception test. Adult listeners were dominantly affected by VOT in correctly identifying fortis stops. Considering the role of VOT in production and perception, the pattern of acquiring and mastering fortis stops earlier than the other two stops should be universal, not exceptional.

Despite these studies attempted to provide insight into the development of Korean stop contrasts by children, it is still unclear whether F0 significantly functions for the perceptual distinction between aspirated and lenis stops in early stages of development of a native language. In spite of the importance of the interaction between speech production and perception, interrelated studies are limited, but it is essential to investigate the dynamics of speech production and perception especially for child language development. The biggest problem of previously reported studies is that they lack the study of children's perception even though they mainly focus on articulatory distinction of young children. A tight link between perception and production has been suggested as perceptual sensitivity guides production in young children, and their perceptual sensitivity decreases as they get the native language input resulting in the formation of phonetic categories for native contrasts (e.g., Kuhl et al., 1992; Polka and Werker, 1994; Werker and Tees, 1984).

With the recent trend of tonogenesis, the role of F0 has become essential, especially in discrimination between aspirated and lenis stops in adult speech. Since F0 makes a phonological stop contrast in a word-initial position, the relevant perceptual representations for the three different stops in the dimension of F0 should undergo phonetic adjustment until children complete the acquisition of stop contrasts, but developmental patterns in terms of F0 have not been fully dealt with in linguistics studies. Therefore, this study aims to capture the development of F0 with regard to perceptual and articulatory distinction between non-fortis stops, lenis and aspirated stops, and to investigate the interrelation between production and perceptual abilities. To pursue these goals, this study conducts experiments with 12 young children aged three years and provides a phonetic analysis of the data. The experiments consisted of a production test and a perception test. In the production test, to elicit children's natural production of Korean stops, a picture-naming task was used with familiar words. Finding appropriate minimal triplets that toddlers would know and that should be picture-describable for the task would be almost impossible. In the perception test, a point-to-an-object task was used with synthesized-F0 sound stimuli. Through these experiments, the phonetic accuracy of their stop production and the role of their perceptual ability in articulatory accuracy can be analyzed.

2. Production Experiment

2.1 Participants

Twelve toddler speakers who aged between 3;0 and 3;4 (years;months) participated in the production experiment (M = 3;2.7, SD = 0;1.15). They were six male and six female native Korean speakers, and were recruited from a daycare center in Kyungki, where the Seoul dialect is spoken, in South Korea. They were all Korean monolinguals without hearing or speaking difficulties, as reported by their parents/guardians. An additional two children participated but were not able to complete the following perception task, so their responses were not included for the study. All child participants were rewarded with candies and toys during and after the session. Adult speech samples were collected from seven female speakers (M of age = 29.3 yrs) to compare children's productions of Korean stop contrasts. They were all Korean monolinguals without any speaking or hearing disorders.

2.2 Task

A picture-naming task was used for the production test. One set of three words, which include three types of stops (lenis /t/ - fortis /t'/ - aspirated /t^h/), and four fillers were carefully selected and each word was prepared as a describable picture. All pictures were provided in randomized order, and each picture was given three times. The target consonants were positioned word-initially, so /tak/ ('hen'), /t'ʌk/ ('rice cake'), and /t^hajɔ/ ('Tayo', famous cartoon character) were pronounced. Additional four fillers, /k^hamɛra/ ('camera') ,/k^hɔk'iri/ ('elephant'), /p'aŋ/ ('bread'), /p'ɔrɔrɔ/ ('Pororo', famous cartoon character), were also pronounced by the children. Except two words, the speech materials are from the McArthur Communicative Development Inventory-Korean (Pae et al., 2004). The two words are famous cartoon characters (Tayo and Pororo), so they are not technically included in MCDI-K, but all the children confidently pronounced the names of the characters. The participant's responses were elicited to name each given picture with the question "*mwuetici?*" meaning 'what is this?' by the experimenter. If a participant had difficulty, the experimenter assisted them in naming the given picture. An omni-directional YETI microphone was used for the recording and was placed at a distance of 3-15cm from the participant's mouth. The whole session was recorded and digitized on Praat version 5.3.01 (Boersma, 2001) set up on a personal laptop at a 44100 Hz sampling rate. The recordings were conducted in a quiet room in the daycare and saved as WAV files.

2.3 Procedure

Target words were extracted from the recordings, and the word-initial stops were analyzed using Praat version 5.3.01. A total of 171 tokens (3 words × (12 children + 7 female adults) × 3 replicated productions) were obtained from the recordings. Two of these tokens were considered to be inappropriate for phonetic analysis so that they were discarded. Praat calculated VOT of the target stop sound and F0 at the following vowel onset based on the spectrograms and waveforms. Every raw phonetic value was used for analysis.

2.4 Results

Adult speech samples are shown in Figure 1, and general production patterns across the Korean stop contrasts are observed. The representative phonetic definition of fortis stops is seemingly the shortest VOT values compared to lenis and aspirated counterparts. Since VOTs for lenis and aspirated stops are almost similar, it seems that F0 complementarily operates for the distinctive articulations between lenis and aspirated stops. In the dimension of F0, the biggest phonetic difference can be found between lenis and aspirated stops, but in the case of the fortis counterpart, it has intermediate F0 values between them.

Production patterns of the child participants are illustrated in Figure 2. Comparing to the adult productions, the three-year-old participants show similar VOT variations across the stop categories; Fortis stop /t'/ has the shortest VOT, and VOTs for aspirated and lenis stops are almost overlapping. The differences in VOT between stop contrasts are relatively bigger in the case of adult productions than in the child productions. Still, children's productions show that in the case of fortis /t'/, VOT values are clearly shorter than those for lenis /t/ or aspirated /t^h/ stops. Most of the tokens of

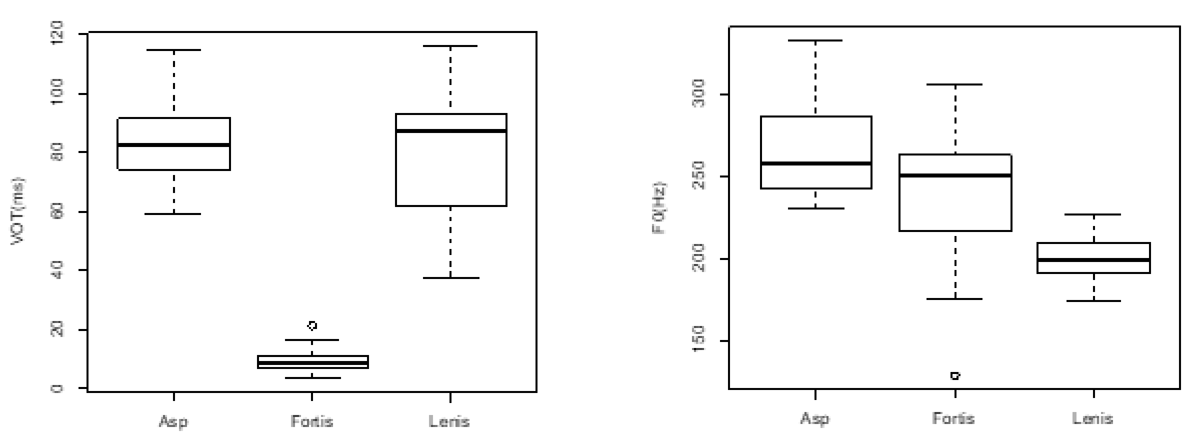

Fig. 1. Adult productions of aspirated /t^h/, fortis /t'/, and lenis /t/ in the dimension of VOT (left panel) and F0 (right panel).

/t'/ were pronounced within 50ms. Lenis and aspirated stops share almost identical ranges of VOT, but in the case of aspirated stop /tʰ/, a slightly wider range of VOT values is found. Aspirated and lenis stops have overlapping VOT values, which are relatively longer than those for a fortis counterpart. The figure indicates that VOT functions enough for distinguishing fortis stops from lenis and aspirated stops.

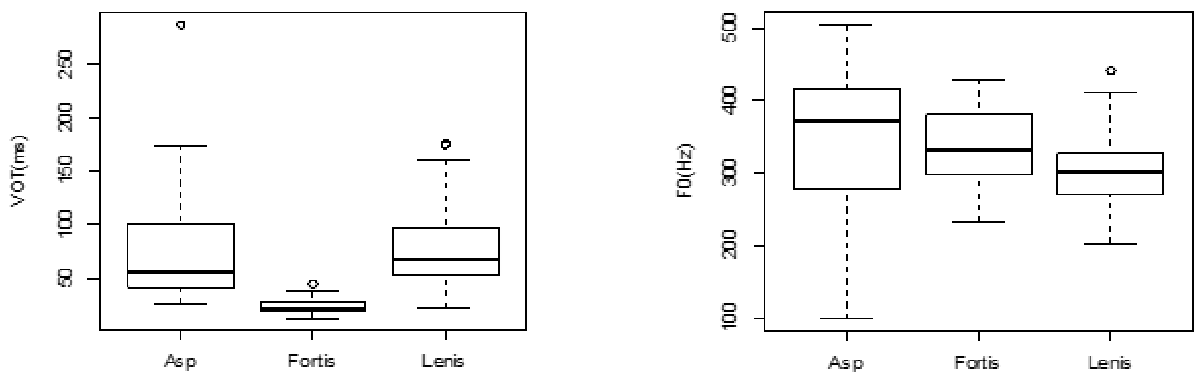

Fig. 2. Child productions of aspirated /tʰ/, fortis /t'/, and lenis /t/ in the dimension of VOT (left panel) and F0 (right panel).

In the dimension of F0, the three stop categories are not apparently distinguishable, since the F0 values estimated are somewhat overlapping across the three stop types. In the case of aspirated stop /tʰ/, F0 values are widely dispersed from 100 Hz to 500 Hz. Relatively, fortis and lenis stops were produced within small ranges of F0. Median F0 values of three different stop categories show some F0 differences: aspirated stop, /tʰ/, had the highest F0, while lenis stop, /t/, had the lowest F0.

To identify the correlation between the estimated phonetic values and stop categories in children's production, a mixed-effects linear regression model (Raundenbush and Bryk, 2002; Snijder and Bosker, 1999) was designed and conducted. The *lme4* package (Bates, Maechler, and Bolker, 2011) was implemented in R (R Development Core Team, 2011). VOT (ms) was the dependent variable and by-speaker adjustment to intercept is a random effect. Aspirated stops for stop category were the reference category in this model.

Table 1. Mixed-effect linear regression for the effect of stop category on VOT (ms), with aspirated as the reference category.

Random effects:				
Groups	Name	Variance	Std. Dev.	
childid	(Intercept)	42.31	6.50	
Residual		1525.89	39.06	
Number of obs: 106, groups: childid, 12				
Fixed effects:				
	Estimate	Std. Error	t value	Pr (>|t|)
(Intercept)	78.21	6.33	12.35	4.44e-16***
Fortis	-54.99	8.83	100.08	1.12e-08***
Lenis	2.77	9.33	0.29	0.77

NOTES: * = p < .05, ** = p < .01, *** = p < .001

As shown in Table 1, significant effects of fortis and aspirated stops on VOT are observed using this statistical model (p-values < .05). On the other hand, there is no significant effect of lenis stop /t/ on VOT estimates. This regression model indicates that VOT mainly functions for articulatory distinction for fortis or aspirated stops, which shows that fortis stop /t'/ tends to have much shorter VOT values (*coef.* = -54.99 ms) compared to those for aspirated stops. As suggested in Figure 2, there is no significant difference between lenis and aspirated stops in VOT. This pattern is similar to the female adult speakers' production. The same statistical analysis reported that in adult speech, fortis stops have significantly shorter VOT values (*coef.* = -74.21 ms, p < .001), while lenis stops did not show any significant VOT differences (p > .1).

To determine the effect of stop categories on F0, another mixed-effects linear regression model was used. Results in Table 2 show that there is no significant relationship between fortis stop /t'/ and its F0 values (p > .05). Aspirated and lenis stops significantly affect F0 values in the children's productions. According to the results, lenis stop /t/ has relatively lower F0 values (*coef.* = -38.98) compared to the aspirated counterpart /t^h/, while fortis stop /t'/ shows similar F0 values (*coef.* = -1.96) with aspirated /t^h/, which is not a significant change (p > .05). The same multiple linear regression analysis was conducted for the control group, and revealed that the three different stop categories have significantly different F0 values. Aspirated stops have higher pitch, while the lenis counterpart has significantly lower F0 (*coef.* = -64.52, p < .001). The difference from the children's production is that adult speech showed a significant F0 effect on fortis stop (*coef.* = -25.54, p < .01).

To show the child gender effect on F0, a multiple regression model was conducted, and it revealed that there is no significant speaker gender effect on F0 in the production of Korean stop contrasts (p > .1).

Table 2. Mixed-effect linear regression for the effect of stop category on F0 (Hz), with aspirated as the reference category.

Random effects:				
Groups	Name	Variance	Std. Dev.	
childid	(Intercept)	694.3	26.35	
Residual		4924.9	70.18	
Number of obs: 106, groups: childid, 12				
Fixed effects:				
	Estimate	Std. Error	t value	Pr (>|t|)
(Intercept)	341.67	13.32	25.65	< 2e-16 ***
Fortis	-1.96	15.93	-0.12	0.902
Lenis	-38.98	17	102.75	0.024 *

NOTES: * = p < .05, ** = p < .01, *** = p < .001

To see the effect of children's age (in months) on such phonetic values depending on stop categories, one interaction term (age [in months] × stop category) was added to the multi-level regression models for the analysis. The results show that there is no significant effect of children's age on VOT or F0 as a function of stop categories (p-values > .05).

3. Perception Experiment

The three-year-old children production revealed that significant distinction between lenis and aspirated stops requires F0 dimensional differences, which has been observed in adult speech as well. Both speaker groups showed merged VOT of lenis and aspirated stops, while significant effect of F0 on those stop categories. In this section, considering the consequence of the role of F0 in differentiating between lenis and aspirated stops, the effect of F0 difference on perception of lenis and aspirated stops will be analyzed.

3.1 Participants

The same child participants from the previous production experiment also participated in the perception experiment. As mentioned, additional two children were not able to complete the task, so a total of twelve children provided their responses for the analysis. The same female adult speakers who participated in the production test were the control group of this perception experiment.

3.2 Task

To observe the effect of F0 changes on perceptual distinction between alveolar lenis and aspirated stops (/t/ and /tʰ/), a perception experiment was designed and conducted. A forced-choice selection method was used for the perception experiment. For two unfamiliar objects, two monosyllabic non-words were invented (Figure 3). The thing for /taŋ/ is a blue, pentagonal shaped thin object, while the one for /tʰaŋ/ is a red, cylindrical shaped object. The two words /taŋ/ and /tʰaŋ/ have word-initial alveolar lenis and aspirated stop respectively, and the pronunciations of the words sound natural but need not to remind any other real words. The children played with their caregivers and the experimenter using the objects about 30 minutes. Before the session began, the experimenter confirmed if the children acknowledged the correct name of each object. In the real session, the participants were asked to point to or choose one of the two objects right after hearing "*mwueti* [target word] *ici?*" meaning 'which one is [target word]?' from the audio. The session took approximately five minutes and consisted of 10 trials (6 target words + 4 fillers). If a child seemed distracted, the same question was asked again. Every response was documented by the experimenter during the session.

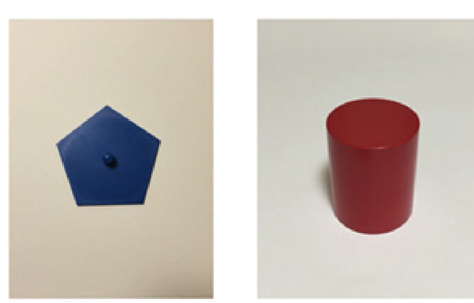

Fig. 3. Objects used for '/taŋ/' (left) and '/tʰaŋ/'(right).

3.3 Listening Materials

To minimize the possible effect of word familiarity on children's perception of words, two nonce words were used. The production of /taŋ/ by a thirty-year-old native Korean monolingual female speaker was used for the synthesis. Using the pitch synthesis function in Praat (Boersma, 2001), the onset of the following vowel was manipulated to have a 15 Hz difference between the tokens. The best exemplar of natural /taŋ/ by the speaker was extracted, and VOT and F0 at the vowel onset was calculated; it has 75 ms VOT and 185 Hz. As a result, the F0 values at the vowel onset of /taŋ/ were adjusted to be 185 Hz, 200 Hz, 215 Hz, 230 Hz, 245 Hz, and 260 Hz. The six synthesis stimuli had fixed VOT (75 ms) to get rid of effects of VOT difference. To make every stimulus sound natural, not artificial, following pitch points were also manipulated but the original pitch contour across the word and other phonetic features were maintained.

3.4 Results

The results of the identification of the contrastive pair /taŋ/-/t^{h}aŋ/ with slightly changed F0s at the following vowel onset reveal that a relative F0 difference would exert an effect on perceptual thresholds to the aspirated of the target children. Figure 4 represents the ratio of all the responses for the two alternatives, a lenis /t/ and an aspirated stop /t^h/. Even though not all responses by the children were consistent or categorical, the results still show some perceptual patterns depending on the pitch height of the stimuli; Three-year-old children's perceptual distinction between /taŋ/ and /t^{h}aŋ/ relies on a substantial phonetic difference in F0. When the sound stimulus with 185 Hz was given, all the children responded to it as the lenis counterpart /t/. However, when the stimulus has 75 Hz-higher pitch at the vowel onset (260 Hz), 75% of the responses were the aspirated stop /t^h/.

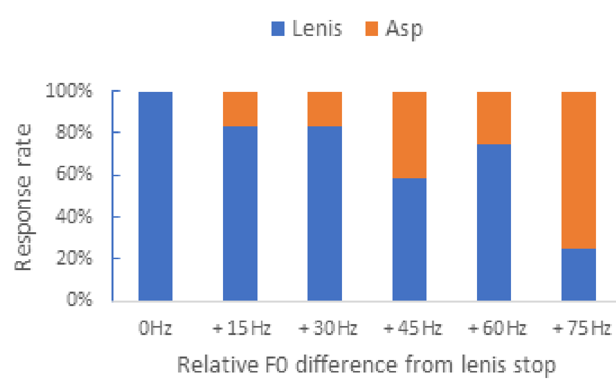

Fig. 4. Children's responses to the synthesized sound stimuli according to the relative F0 difference from the lenis /t/.

To examine the effect of the relative F0 difference on children's perceptual pattern, a mixed-effects logistic regression model in (1) was constructed and applied, as the dependent variable has binary values. This model has a predictor variable *F0 difference*, which means a relative F0 difference from the standard lenis articulation (185 Hz), so six different values in a Hz-scale were used in this continuous variable. The output of this multi-level regression model is presented in Table 3 and the associated figure is provided in Figure 5.

$$\text{(1)} \quad \log\left(\frac{Answer_{ij}}{1-Answer_{ij}}\right) = \beta_0 + \beta_1 \times F0\,difference_{ij} + \epsilon_{ij} + \eta_i$$

($Answer = 1$ if a child's response was /t^h/, otherwise $Answer = 0$)

Table 3. Mixed-effects logistic regression for the effect of F0 differences on children's responses, when aspirated = 1 and lenis = 0.

Random effects:				
Groups	Name	Variance	Std. Dev.	
childid	(Intercept)	0	0	
Number of obs: 72, groups: childid, 12				
Fixed effects:				
	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-2.96	0.72	-4.09	4.37e-05 ***
F0 difference	0.05	0.01	3.51	0.00045 ***

NOTES: * = p < .05, ** = p < .01, *** = p < .001

The results confirm the significant effect of relative F0 differences ($p < .05$) on the children's responses to the stimuli, which also indicates the effect of F0s on the perceptual identification of aspirated stop /t^h/. The fixed-effects coefficient for *F0 difference* is 0.05, which interprets that there is a 5% of increase ($1.05 \approx e^{0.05}$) by F0 difference in the odds of the probability of the stimuli being identified as /t^h/. The same regression model for the adult group revealed that F0 difference significantly affects the perception of aspirated stops (*coef.* = .18, $p < .001$), which implies that in the case of adult listeners, the effect of F0 difference is much greater than in the case of children.

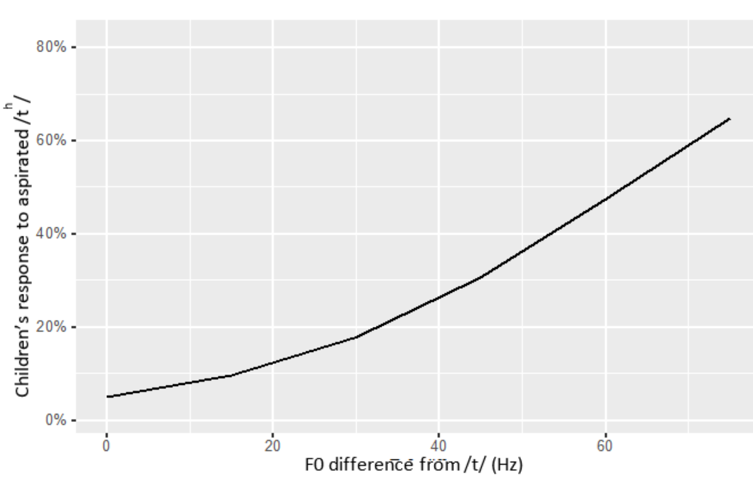

Fig. 5. Relationship between F0 differences and children's perception of /t^h/ with a logistic curve.

As shown in Figure 5, the pitch height at the vowel onset is positively related to the perception of aspirated stop /t^h/. This relationship between F0 values and children's perception of aspirated stops has been statistically supported, and it also indicates that F0 plays a key role in the phonetic distinction between /t/ and /t^h/.

4. Discussion and Conclusion

The production experiment confirmed that alveolar fortis stop /t'/ can be phonetically defined by its shortest VOT even though lenis and aspirated counterparts need to be distinguished in the dimension of F0. The three-year-old children were able to show a significant relationship between VOT and fortis/aspirated stops but lenis stop /t/ did not show a significant VOT effect. In spite of the fact that they did not show substantial VOT difference between fortis and non-fortis stops as adult speakers did, we propose that three-year-old children have developed VOT differentiation at least for fortis stop. Alongside the clear relationship between fortis and shortest VOTs, the children's distinctive articulations for lenis and aspirated stops need to develop further regarding F0-dimensional differentiation; The actual F0 values were quite overlapping across the stop categories, and the lenis-aspirated F0 difference was not as much as in the productions of adult speakers. Still, it was able to observe some statistical significance of F0 and non-fortis stops: there is a significant correlation between F0 values and lenis/aspirated stops.

Therefore, findings from the production experiment indicate that VOT does not necessarily function as a key acoustic parameter to distinguish between lenis and aspirated stops, and that F0 plays a crucial role in the distinction between word-initially occurring lenis and aspirated stops, as reported in many previous studies on Korean tonogenesis (e.g., Kang, 2014 among others). Three-year-olds were able to have some significant phonetic difference between stop contrasts in each acoustic dimension, but it also seems that their phonetic distinction needs to develop further.

The perception experiment was conducted to examine the role of F0 in the perceptual distinction between lenis and aspirated stops. Since word familiarity would affect the perception of young children (e.g., Swingley and Aslin, 2007), two monosyllabic minimal pairable words were newly created and used. The results confirmed the significant effect of F0 changes on the perception of aspirated stops. When a relatively high-pitched stimulus was heard, the children tended to perceive it as aspirated. This tendency was statistically tested, which suggests that F0 is one of the crucial perceptual cues for determining between lenis and aspirated stops. Since the new words were used for the perception test, it should be noted that the children had to learn and memorize them in a very short time, and this would make the task harder and complicated. However, using new words was the best way to avoid having the word frequency effect or the word familiarity effect on perception of a certain stop sound. It would make the analysis focus on the perception of phonetic realizations of phonological features, even though children's performance could be poorer than usual.

To investigate the interrelated development of the crucial phonetic parameter, F0, in language performance, this section assesses how three-year-old children's perceptual and production ability would be correlated; The findings from children's production and perception patterns are compared. In the children's production data, the F0 values at vowel onset after /t/ and /t^h/ were estimated, and the average F0 differences between the two stop categories were calculated. Since aspirated stops usually have higher F0 values, the F0 differences from aspirated to lenis stops were calculated (≈ 33.5 Hz). Regarding the perception data, recall that not all responses were consistent and categorical. Each stimulus was tested once, so it was not able to observe average perceptual patterns depending on the sound stimuli. Therefore, through a simple logistic regression model (in which as a response, /t^h/=1 and /t/ = 0), every response of each child was determined if it is consistent enough to analyze. With the standard score of 10.0 of AIC, greater AIC scores were excluded for analysis due to its inconsistency. As a result, eight children's responses were considered reliable, and the responses of four children out of the twelve were discarded. Using the responses of the eight stable responders, the children's perceptual F0 threshold for aspirated /t^h/ was calculated. As shown in Figure 6, the values were calculated at the midpoint between the lowest F0 value of the stimulus that was identified as aspirated /t^h/ and the highest F0 values of the stimulus that was identified as lenis /t/. For example, if a child responded 't, t, t, t, t^h, t^h' to '185, 200, 215, 230, 245, 260 Hz' stimuli respectively, then the minimum F0 difference required for perceptual categorizing /t^h/ is calculated as 52.5 Hz (= [the midpoint of 230 and 245 = 237.5 Hz] - [185 Hz]) in this case.

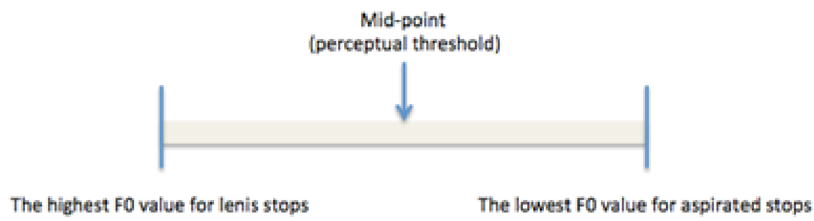

Fig. 6. The calculation of perceptual thresholds for aspirated stops in children's perception.

Figure 7 shows the estimated F0 differences in children's and the control group's production and perceptual discrimination. In addition to the child data, the adults in the control group also provided the data for comparison. In children's actual productions, only 33.5 Hz difference in F0 was observed between aspirated /t^h/ and lenis /t/. However, for perceptual identification of /t^h/, the same children needed 44 Hz difference from the lenis /t/. This phonetic mismatch between production and perception can be understood in the framework of the Theory of Adaptive Dispersion, in which phonetic parameters are built at the compromising point of the two different needs in production and perception (Liljencrantz and Lindblom, 1972; Lindblom, 1990; Lindblom and Engstrand, 1989). Perceptually sufficient phonetic differences and articulatory economy should compromise at a certain point in the F0 dimension. The main factor of this model suggests that phonemic categorization is motivated by perception of phonological features in the phonetic input. The relative F0 differences that were estimated in adult speakers' production and perception were almost identical, and this would be in support of the framework of the Theory of Adaptive Dispersion, in that phonemic boundaries in the acoustic dimension work to consistently map the phonetic input onto the appropriate phonemic categories. In perception of adult listeners in the control group, an F0 difference of as much as 61 Hz was required for the distinction between lenis and aspirated stops, while a similar amount of F0 difference was estimated in their articulatory distinction.

On the other hand, the substantial difference in F0 between children's production and perception indicates that the children's phonemic categories in the F0 dimension need to develop further at the age of 3 years. The minimal F0 difference to perceptually distinguish between lenis and aspirated stops is around 44 Hz, which is relatively smaller than the estimated F0 difference in adults' perception. This smaller phonetic difference indicates that three-year-olds would be perceptually more sensitive compared to adult listeners, and their phonemic categories in the dimension of F0 have to develop for consistent mapping. In addition, the children's production did not reflect the same amount of phonetic difference observed in perception, which also means that children's articulatory distinction is still immature in the F0 dimension. Overall, the findings suggest that the language development in three-year-old children is still in progress, decreasing their perceptual sensitivity and making phonetic adjustments for their native phonological contrasts.

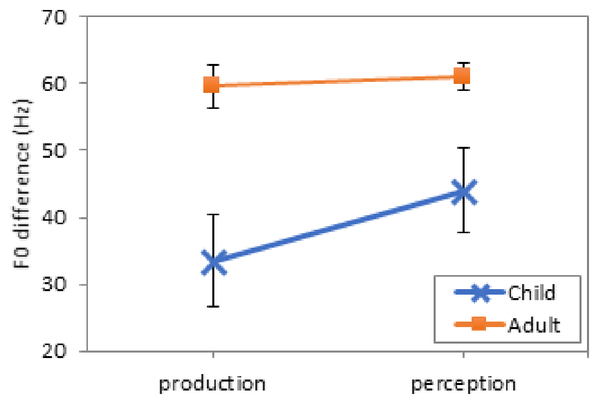

Fig. 7. F0 differences between /t/ and /t^h/in children's production and perception. Error bars represent standard errors.

Under the assumption that 3-year-old children are still acquiring Korean stop contrasts in the acoustic dimension of F0, it is still a puzzle what other phonetic details can differentiate the contrasts, namely lenis and aspirated stops. As shown in the production experiment, VOT plays a crucial role in distinguishing fortis stops, and there was a significant correlation between VOT and aspirated stops. As fortis and aspirated stops can be distinguished in the dimension of VOT as the shortest and the longest respectively, VOT could complementarily function to differentiate the contrasts until children fully develop another dimensional distinction (F0). However, the test stimuli were strictly restricted to have only a small amount of F0 changes, so it was not able to observe any possible effect of other phonetic features such as VOT or durational differences. As a tonogenetic sound change is undergoing in Seoul Korean, it is important to investigate the changing roles of VOT and F0 in production and perception of Korean stop contrasts. Despite that the findings in this study indicate the importance of F0 as a determinant phonetic cue to make a contrast, it is difficult to suggest a clear relationship between VOT and F0 in the acquisition of stop contrasts, mainly because this study lacks children's perceptual patterns depending on varying VOTs. It would be extremely challenging to test young children with a repetitive task because of their short concentration span. For future research, the effect of VOT should be considered to more precisely capture the developmental pattern of stop contrasts in the phonetic dimensions.

Before tonogenesis occurs in Seoul Korean, VOT has been a determinant cue to differentiate three stop series and F0 has been a secondary cue. In this language-specific context, sound input by adult speakers can vary, which also can affect the acquisition pattern of Korean young children. To give an explanation about the mystery in the acquisition of multiple acoustic cues, a more sophisticated design of experiments with young children should be required.

This study attempted to provide empirical evidence of the role of F0 in the acquisition of Korean stop contrast by three-year-old children. However, it is hard to unify developmental patterns in the case of language-learning children. This study only dealt with the small number of three-year-olds, so the findings from this study would not represent the linguistic competence of the same aged children. A large body of data should be collected in future research, since the child listeners were not able to provide stable and consistent responses throughout the experiment. It was difficult to obtain a large amount of stable responses from the children, so the results were inevitably generalized to some extent. Increasing the total number of child participants would increase the probability of getting valid responses, which would produce more reliable and conclusive results.

Another challenge in the analysis was the limited F0 values at vowel onset in the perception experiment. The given relative F0 difference was not that large, which presumably caused not-categorical perceptual patterns. The number of test stimuli was carefully decided due to the children's short concentration span, but the given phonetic difference could not cover the phonetic ranges for categorical perception between lenis and aspirated stops; The test stimuli had 75 Hz difference at most, which was not enough for the categorical perception. For more reliable results, future works should expand the target phonemes to include labial or velar stop contrasts, and various VOT/F0 ranges should be tested for children. Given children's short attention span, the length of the experiment and the number of stimuli should be also considered.

As VOT dominantly functions quite exclusively for distinguishing fortis stops, it is important to acquire F0 as the determinant cue to phonetically differentiate lenis and aspirated stops (Kang, 2014; Kim and Stole-Gammon, 2009; Kong et al., 2011). The children's stop perception and production showed that phonemic category formation is in progress in the acoustic dimension of F0, which directly affects the discriminatory ability for lenis-aspirated contrast. Children at the age of 3 years need to stabilize the sufficient contrastive distinctiveness for lenis and aspirated stops. Children's perceptual development leads to more mature acoustic implementation, so their articulatory distinction will be more adult-like. The findings from the two experiments suggest that children have not fully developed multi-parametric differentiation regarding Korean stop contrasts until the age of 3 years.

References

Allen, G. 1985. How the Young French Child Avoids the Pre-Voicing Problem for Word-Initial Voiced Stops. *Journal of Child Language* 12, 37-46.

Bates, D., Maechler, M., Bolker,B., and Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67.1, 1-48.

Boersma, P. 2001. Praat, a System for Doing Phonetics by Computer. *Glot International* 5(9/10), 341-345.

Clumek, H., Barton, D., Macken, M. A., and Huntington, D. 1981. The Aspiration Contrast in Cantonese Word-Initial Stops: Data from Children and Adults. *Journal of Chinese Linguistics* 9, 210-224.

Gandour, J., Petty, S. H., Dardarananda, R., Dechongkit, S., and Mukongoen, S. 1986. The Acquisition of the Voicing Contrast in Thai: A Study of Voice Onset Time in Word-Initial Stop Consonants. *Journal of Child Language* 13, 561-572.

Haudricourt, A. G. 1954. De l'origine des tons du Vietnamien. *Journal Asiatique* 242, 69-82.

Hombert, J-M., Ohala, J. J., and Ewan, W. G. 1979. Phonetic Explanations for the Development of Tones. *Language* 55, 37-58.

Jun, S. -A. 1996. Influence of Microprosody on Bacroprosody: A Case of Phrase Initial Strengthening. *UCLA Working Papers in Phonetics*, 97-116.

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.

Kewley-Port, D., and Preston, M. S. 1974. Early Apical Stop Production: A Voice Onset Time Analysis. *Journal of Phonetics* 2: 195-210.

Kim, M., and Stoel-Gammon, C. 2009. Tha Acquisition of Korean Word-Initial Stops. *Journal of Acoustical Society of*

America 125.6, 3950-3961.

Kim, S. -H. 1999. Multiple Acoustic Cues of Three-Way Phonemic Contrast in Stop Consonants. *Studies in Phonetics, Phonology and Morphology* 5, 79-103.

Klatt, D. 1975. Voice Onset Time, Frication, and Aspiration in Word-initial Consonant Clusters. *Journal of Speech and Hearing Research* 18, 686-706.

Kong, E. J., Beckman, M. E., and Edwards, J. 2011. Why are Korean Tense Stops Acquired so Early?: The Role of Acoustic Properties. *Journal of Phonetics* 39, 196-211.

Kuhl, P. K., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. 1992. Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age. *Science* 255, 606-608.

Ladefoged. P., and Maddieson, I. 1996. *The Sounds of the World's Language*. Oxford: Blackwell.

Liljencrantz, J., and Lindblom, B. 1972. Numerical Simulations of Vowel Quality Systems: The Role of Perceptual Contrast. *Language* 48, 839-862.

Lindblom, B. 1990. Explaining Phonetic Variation: A Sketch of the H & H Theory. In W. J. Hardcastle & A. Marchal (eds.), *Speech Production and Speech Modeling* (pp. 403-439). Dordrecht:Kluwer.

Lindblom, B. and Engstrand, O. 1989. In What Sense is Speech Quantal? *Journal of Phonetics* 17, 107-121.

Lisker, L. and Abramson, A. S. 1964. Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word* 20, 384-422.

Pae, S., Chang, Y., Kwak, K., Sung, H., and Sim, H. 2004. MCDI-K Referenced Expressive Word Development of Korean Children and Gender Differences. *Korean Journal of Communication Disorder* 9, 45-56.

Pan, H. S. 1994. *The Voicing Contrasts of Taiwanese (Amoy) Initial Stops: Data from Adults and Children*. Ph.D. Dissertation, Ohio State University, Columbus, OH.

Polka, L., and Werker, J. F. 1994. Developmental Changes in Perception of Non-Native Vowel Contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421-35.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. (http://www.r-project.org/)

Raudenbush, S. W., and Bryk, A. S. 2002. *Hierarchical Linear Models (2nd edition)*. Thousand Oaks, California: Sage.

Silva, D. J. 2006. Acoustic Evidence for the Emergence of Tonal Contrast in Contemporary Korean. *Phonology* 23, 287-308.

Snijders, T., and Bosker, R. 1999. *Multilevel Analysis*. London: Sage.

Swingley, D., and Aslin, R. N. 2007. Lexical Competition in Young Children's Word Learning. *Cognitive Psychology* 54, 99-132.

Werker, J., and Tees, R. 1984. Cross-Language Speech Perception: Evidence for Perceptual Reorganization During the First Year of Life. *Infant Behavior and Development* 7, 49-63.

Wright, J. D. 2007. *Laryngeal contrast in Seoul Korean*. Ph.D. Dissertation, University of Pennsylvania, Philadelphia, PA.Gayeon Son, Professor

Son, Gayeon, Professor
20 Kwangwoon-ro, Nowon-gu, 01897, Republic of Korea
Department of English Language and Literature, Kwangwoon University
E-mail: gson@kw.ac.kr