DEVELOPMENTAL PATHWAY OF PHONETIC FINE-TUNING OF PHONOLOGICAL CONTRAST IN REFERENCE TO INFORMATION STRUCTURE: A PRELIMINARY STUDY ON THE THREE-WAY CONTRASTIVE STOPS IN KOREAN

Sahyang Kim[1], Anqi Yang[2], Aoju Chen[3], Jiyoung Lee[4], and Taehong Cho[4]

Hongik University[1], Tianjin University[2], Utrecht University[3],
Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University[4]
sahyang@hongik.ac.kr, anqi.yang@tju.edu.cn, aoju.chen@uu.nl, vnalvores@naver.com, tcho@hanyang.ac.kr

ABSTRACT

This study investigates the acquisition of three-way stop contrast (lenis, aspirated, fortis) in Korean and phonetic modulation of the phonological contrast driven by information structure. VOT and F0 of the stops produced by children (4-5, 7-8, 10-11 year-olds) and adults in broad, narrow, and contrastive focus conditions were measured. Results indicated that only the 7-8 year-olds showed the three-way distinction using VOT under (phonemic) contrastive focus, while the other two children groups and adults did not. As for F0, adults made a three-way distinction using F0 in all focus conditions, while the 7-8 and the 10-11 year-olds did so only under limited focus conditions. The 4-5 year-olds did not show the three-way distinction in any focus condition. The results suggest that children build up their phonological awareness and fine-tune phonetic realization in their developmental pathway in conjunction with different functions of information structure.

Keywords: Korean three-way stops, L1 acquisition, phonetics-prosody interface, information structure

1. INTRODUCTION

In spoken language, the information structure of utterances is phonetically manifested in various dimensions [3]. Focus information, for example, is well-known to be phonetically realized at the suprasegmental level. The focused elements tend to be produced with an expanded pitch range and longer duration [e.g., 4,8], although the details of phonetic focus realization differ across languages, depending on the prosodic system of a given language [e.g., 11, 20]. Focus is also known to be phonetically encoded at the segmental level. In general, the focused elements are 'strengthened', such that, for example, voiceless aspirated stops are produced with longer VOT and vowels are produced more peripherally [6]. Such fine-phonetic modulation triggered by the presence of focus in the information structure also differs across languages, as the segmental strengthening contributes to the enhancement of phonological contrasts in a given language [9,10].

These findings suggest that the phonetic manifestation of information structure must be acquired in a language-specific way, interacting with the prosodic and the phonological system of a language. Recent studies have indeed demonstrated that, at the suprasegmental level, children's use of pitch and duration cues in focus-marking differs depending on the prosodic characteristics of their native language [2,17,18,19]. However, our knowledge is limited as to how children acquire focus-induced phonetic modulation for language-specific phonological contrast. This question is of importance as it would also allow us to observe how children build up their native phonological contrast via the phonetic fine-tuning of segments in conjunction with focus, given that focus would yield phonological contrast maximization.

In an attempt to explore this issue, the present study examines how children acquire phonetic modulation of Korean three-way stop contrast (lenis, aspirated, fortis) in different focus contexts. The Korean stops are of interest due to their unique contrast – i.e., they are all voiceless in the word-initial position. While they are distinguished in several phonetic dimensions, VOT and F0 have been considered as major acoustic cues for the distinction [e.g., 1,5]. The fortis stop is produced with short lag VOT, and the other two with long lag VOT; and the lenis stop is produced with lower F0 compared to the other two stops. Another interesting fact about Korean stops is that they are currently undergoing a diachronic sound change. The VOT distinction between the lenis and aspirated stops, with shorter VOT for the former [see 5], was once robust but is now being blurred in the word-initial position; and the F0 of the following vowel plays a more important role in distinguishing the two stops, with a higher F0 for the aspirated than the lenis stops [1,12,16].

Recently, Choi at al. [7] found that Korean speakers in their 20's raise F0 more for the prosodic phrase-initial (therefore word-initial) aspirated than the lenis stop when the word is focused than when not, but no such focus-induced effect was found for VOT in the same position. Given that the presence of focus enhances phonological contrasts [9,10], results in [7] indicates that the F0 is now being employed as a

major cue to the aspirated-lenis contrast at least in the phrase-initial position. ([7], however, showed that VOT plays a role in distinguishing the two stops in the unfocused, phrase-medial (but still word-initial) condition where the lenis stop is produced as voiced.)

Studies showed that Korean children can produce the three-way phonological contrast by the age of 4 [13,14]. A recent phonetic study, however, have shown that 2 to 5-year-old children exhibit overlap between the fortis and aspirated stops in the F0 dimension and overlap between the lenis and aspirated stops in terms of VOT when the stops were in the initial position of words produced in isolation [15]. It can be therefore hypothesized that the fine phonetic manipulation of VOT and F0 may occur in different focus contexts during the course of the acquisition of the stops. If so, the stop contrasts will be maximized under focus and as a result, there will be no or less overlap among the consonants in their phonetic dimension. In addition, if children are sensitive to the ongoing sound change in their environmental language, they may show similar pattern as that of young adults found in Choi et al. [7], manipulating F0 more under focus for the phonological contrasts in the phrase-initial position. It is, however, still an open question which of the two phonetic dimensions would be more responsive to the demand from the information structure during the course of acquisition and how the diachronic sound change is reflected in children's stop production.

2. DATA ACQUISITION

2.1. Participants

Twelve adults (6F, 6M), eight 10 to 11- year-olds (4F, 4M), eight 7 to 8-year-olds (6F, 2M) and six 4 to 5-year-olds (2F, 4M) participated in the study. All of them were native speakers of Seoul Korean.

2.2. Speech Materials

The data used in this study is a subset of the data used in [17], with 12 target words and 5 focus types in various sentential locations, yielding 60 sentences per speaker. For the current purpose, only three monosyllabic and three multisyllabic target words with a lenis, aspirated, and fortis onset (/pal/ 'foot', /p^{h}al/ 'arm', /p*aŋ/ 'bread'; /kapaŋ/ 'bag', /katɨ/ 'card', /k*amakwi/ 'crow') inserted as an object in target SOV sentences were used in three focus conditions (see below), which yielded 18 sentences per speaker. It should thus be noted that the results of the present study must be taken with caution due to the limited data.

Target sentences were elicited within a series of short dialogues in a picture-matching game, with a question or a statement produced by an experimenter and an answer or a contrasting statement produced by a speaker (see [17] for details). The experimenter had pictures with missing information, and the participant had corresponding pictures with full information of 'who is doing what'. Target words were produced as an answer to a broad-focus question (1), a narrow-focus question on the object (2), and a contrastive-focus question on the object (3). Note that the three sentences were exactly the same, except for their domain of intended focus, as indicated by the curly brackets in the examples. As the domain of the narrow and the contrastive focus was the target word, it was expected that they would yield more phonetic prominence on the target than the broad focus would whose domain was the entire sentence. In addition, as the contrastive focus was always induced by a phonemic contrast within a word, its domain, in principle, was smaller than that of the narrow focus.

(1) Broad-focus in SVO
Experimenter: Look! This picture is very blurry. I cannot see anything clearly. **What happens in the picture**?
Participant: **/kɛ-ka p*aŋ-ul mantʃʌ-jo/**
dog-*nom.* bread-*acc.* touch-*final*
[Dog bread touch]

(2) Narrow-focus on the object in SVO
Experimenter: Look! There's a dog, and it puts out its hand. It looks like the dog touches something. **What does the dog touch**?
Participant: **Dog [bread] touch**.

(3) Contrastive-focus on the object in SVO
Experimenter: Look! There's a dog, and it puts out its hand. It looks like the dog touches something. I will make a guess. **The dog touches an egg**.
Participant: **Dog [bread] touch**.

2.3. Procedures

Participants were tested individually in a quiet room in Hanyang Phonetics and Psycholinguistics Lab.

In order to induce the proper focus context, the target sentences were produced only once without repetition. Tokens with incorrect responses (deviating from the intended sentences in terms of the word choice and the sentence structure), hesitation, and self-repair were excluded for data analyses (19 from adults, 31 from Age 10-11, 49 from Age 7-8, 34 from Age 4-5). The remaining tokens were cross-checked by six trained phoneticians in order to confirm the location of prominence within each sentence and the type of the prosodic boundary before the target object. The tokens with incorrect focus marking (e.g., no phonetic prominence on the target in a focus condition) were also excluded from the analyses (2 from adults, 3 each from Age 10-11 and 7-8, 9 from Age 4-5). Note that more than 90% of the target objects were produced in a phrase-initial position in

all age groups. As focus-induced enhancement patterns of VOT and F0 vary as a function of prosodic position [7], the targets produced in phrase-medial position was also excluded (12 from adults, 10 from Age 10-11, and 4 each from Age 7-8 and 4-5). These exclusion processes further contributed to the limited quantity of data in the present study. In total, 419 tokens were included for data anlyases (176 from adults, 97 from Age 10-11, 85 from Age 7-8, 61 from Age 4-5).

As VOT and F0 are two important acoustic parameters that contribute to distinguishing three-way stops in Korean, reflecting the on-going sound change, we measured VOT (ms) of the stops and the F0 of the following vowel at the midpoint of the vowel using Praat. The F0 values were converted in semitone for further analyses.

2.4. Statistical analysis

Since this study aims to explore how VOT and F0 are phonetically manifested in different age groups and different focus types, data analyses were done separately for each age group and for each focus type. VOT and F0 values were centered at zero by subtracting each data point from the grand mean of each parameter before they were submitted to linear mixed effects models. The fixed factor was Consonant Type (Lenis, Aspirated, Fortis), and the speaker was a random factor. An orthogonal coding was used for the fixed factor, and pairwise comparisons were carried out for the consonants by redefining references in the model. The nearly maximal models[i] were fitted for VOT and F0 values.

3. RESULTS

3.1. VOT

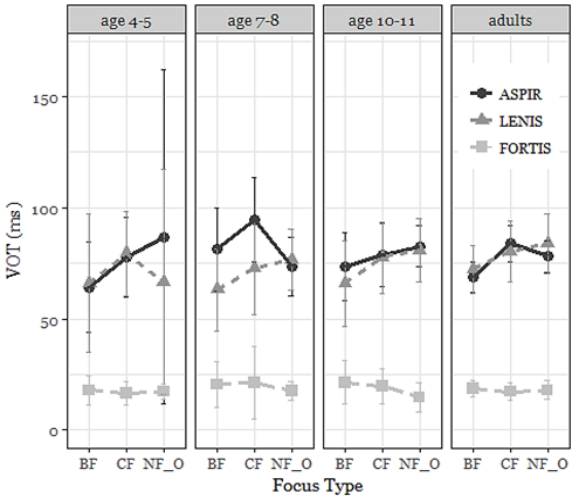
As shown in Table 1 and Fig.1, the fortis stop was clearly distinguished from the other two categories by VOT in all age group and focus types. The lenis-aspirated distinction was not significant regardless of age groups and focus types, with the exception of the 7 to 8-year olds in the contrastive focus condition. Note that the adults' data found in the current study show a different pattern from that found in [7], which showed the adult's use of VOT in the lenis-aspirated distinction under contrastive focus. The difference between the two studies is presumably due to different experimental designs and the limited number of tokens in the current study. Fig. 1 also illustrates substantial variations in children's VOT production, which gradually reduces in the older groups.

In general, as shown in Fig.1, for the lenis and the aspirated stops, VOT increased under the contrastive and narrow focus as compared to the broad focus in all age groups, except for the aspirated stops produced by the 7 to 8-year-olds under narrow focus.

Table 1: The post-hoc comparisons for Consonant Type effect on VOT (in ms) in each focus condition. ('<' refers to p<.05, and '=' to p>.05)

Age	Broad F.	Contrastive F.	Narrow F.
4-5	F < L = A	F < L = A	F < L = A
7-8	F < L = A	**F < L < A**	F < L = A
10-11	F < L = A	F < L = A	F < L = A
Adults	F < L = A	F < L = A	F < L = A

Figure 1: VOT (ms) values for each consonant in each age group. Error bars indicate the standard error.



3.2. F0

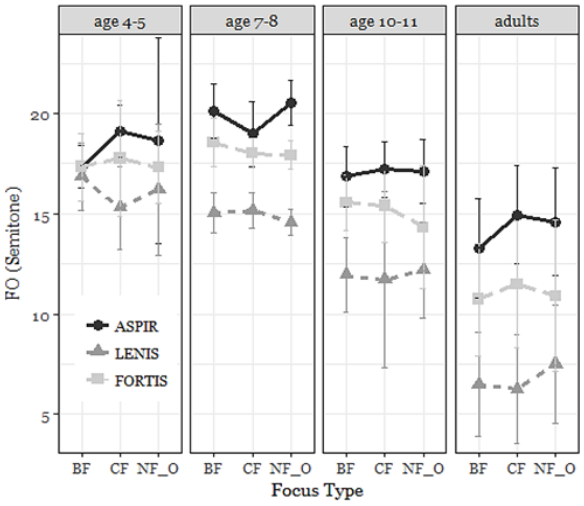
The three-way stop distinction was more clearly observed with F0 than with VOT. In general, the lenis stop showed the lowest F0, the aspirated stop the highest, and the fortis stop showed intermediate values. As summarized in Table 2, the adults showed the three-way contrast with F0 in all focus types. The 10 to 11-year-olds showed the three-way distinction under narrow focus only and marginal distinction between the fortis and the aspirated stops (p=.059) in the contrastive focus condition. In the broad focus condition, however, there was only a two-way distinction between the lenis and the other two stops. The 7 to 8-year-olds showed the three-way contrast in the broad and narrow focus conditions. Only a two-way distinction between the lenis and the other stops was made in the contrastive focus condition. Note that the 4 to 5-year-olds did not show the F0-based stop distinction at all in the broad focus condition. The significant difference between the lenis and the other two stops, however, emerged in the contrastive and narrow focus conditions.

As shown in Fig.2, in the adults' data, the F0 values increased in the contrastive focus condition compared to the broad focus condition in all three stop categories. The children's data, however, did not show such consistency.

Table 2: The post-hoc comparisons for Consonant Type effect on F0 (in semitone) in each focus condition. ('<' refers to p<.05, '≈' to .05<p<.06 and '=' to p>.06)

Age	Broad F.	Contrastive F.	Narrow F.
4-5	L = F = A	L < F = A	L < F = A
7-8	**L < F < A**	L < F = A	**L < F < A**
10-11	L < F = A	L < F ≈ A	**L < F < A**
Adults	**L < F < A**	**L < F < A**	**L < F < A**

Figure 2: F0 (semitone) values for each consonant in each age group. Error bars indicate the standard error.



4. DISCUSSION

A basic finding of the present study is that both adults and children consistently showed a two-way distinction of the fortis and the other two (the lenis and the aspirated) stops in terms of VOT across focus conditions. The only exception was found with the 7 to 8-year-olds who showed a three-way contrast using VOT only in contrastive focus condition. This is in line with the studies which have revealed the declining role of VOT in the three-way stop distinction of Seoul Korean [e.g., 1,12,16]. The three-way contrast, however, was clearly observed with F0 with the adults in all focus conditions, confirming the emerging role of F0 in the three-way stop distinction. Despite the fact that the studies on phonological acquisition suggest the three-way distinction is found before the age 4 [13], our results show that the F0-based consonantal contrast of the older children in the present study has not reached the adult-like level.

More importantly, the present study has observed how Korean children develop fine-phonetic modulation for the three stops under focus. Among the two acoustic measures, the focus-induced contrast enhancement as for VOT was observed only by the 7 to 8-year-olds who showed the lenis-aspirated distinction under contrastive focus. As for the F0 dimension, however, children showed the focus-induced contrast enhancement in their developmental path. The youngest children showed a two-way distinction between lenis and the other stops under contrastive and narrow focus conditions, although no such distinction was observed in the broad focus condition. In a way, the results are similar to Kong et al's [15] findings which showed that children between 2 to 5-year-olds show overlap between the fortis and aspirated stops in terms of F0. In their study, the words were elicited in isolation in a word repetition task. This means that their words were produced in a prosodically strong position, i.e., the utterance-initial position. The results of the present study and of Kong et al. [15] therefore seem to suggest that the Korean children under the age of five are able to show a two-way distinction between the lenis and the other stops in terms of F0, but only in the prosodically strong positions created by focus and prosodic boundary. Interestingly, the F0-based three-way distinction started to emerge with the two older children groups, approximating to the adult-like F0 use. The three-way contrast based on F0 was particularly clear under narrow focus with the older children, suggesting that the focus-induced phonetic modulation is gradually acquired, leading to the language-specific phonological contrast enhancement [9,10]

It is interesting to note that the distinction between the fortis and aspirated stops in the F0 dimension may be considered redundant, given that the two stops are already clearly distinguished as they are in the two extreme ends in terms of VOT. The three-way distinction via F0 therefore verifies its critical role in the phonological contrast among the three stops, and the children's data in the present study suggests that children are in the course of acquiring this critical cue and the presence of focus facilitates the manipulation of this phonetic cue for the phonological contrast.

In conclusion, the current study revealed that Korean children mainly learn to manipulate F0 in order to distinguish the three-way stops, and that they establish their native phonological contrast by fine-tuning the phonetic manifestation in conjunction with the focus structure. Their developmental path for the stop distinction also exhibits how the on-going sound change in Korean is reflected and acquired in children's speech. However, as our data is very much limited in quantity, future studies are required in order to corroborate and generalize the findings.

5. ACKNOWLEDGEMENT

We are very grateful to the children and their parents for their participation, and to Miroo Lee for her assistance with the data acquisition. We also thank the reviewers for their helpful comments. This work was supported in part by a VIDI grant awarded to the third author by the Netherlands Organisation for Scientific Research (grant number 276-89-001), and by the Ministry of Education of the Republic of Korea and

the National Research Foundation of Korea (NRF-2018S1A5A2A03036736) awarded to the first author.

6. REFERENCES

[1] Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., Yoon, T.-J. 2018. The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *J. Phon.*, 66, 120-144.

[2] Chen, A. 2011a. The developmental path to phonological focus-marking in dutch. In: Frota, S., Elordieta, G., Prieto, P. (eds), *Prosodic Categories: Production, Perception and Comprehension*. Springer, Netherlands, 93-109.

[3] Chen, A., 2011b. Tuning information packaging: intonational realization of topic and focus in child Dutch. *J. Child Lang*, 38, 1055-1083.

[4] Chen, A., 2012. Prosodic investigation on information structure. In: Krifka, M., Musan, R. (3ds), *The expression of information structure*. Berlin: De Gruyter Mouton, 251-286.

[5] Cho, T., Jun, S-A., Ladefoged, P. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *J. Phon*. 30, 193-228.

[6] Cho, T., Lee, Y., Kim, S. 2011. Communicatively driven versus prosodically driven hyper-articulation in Korean. *J. Phon.*, 39, 344-361.

[7] Choi, J., Lee, J., Kim, S., Cho, T. 2018. Prosodically-conditioned phonetic cue use in production of Korean aspirated vs lenis stops. *Proceedings of Hanyang International Symposium on Phonetic and Cognitive Sciences*. 121-123.

[8] Eady, S.J., Cooper, W.E., Gayle, V.K. Mueller, P.R., Lotts, D.W. 1986. Acoustic characteristics of sentential focus: Narrow vs broad and single vs dual focus environments. *Language and Speech*, 29, 233-251.

[9] de Jong, K. J. 2004. Stress, lexical focus,and segmental focus in English: Patterns of variation in vowel duration. *J. Phon.*, 32, 493–516. [10] de Jong, K. J., Zawaydeh, B. A. 2002.Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *J. Phon.*, 30, 53–75.

[11] Jun, S-A., Lee, H-J. 1998. Phonetic and Phonological markers of Contrastive Focus in Korean, *Proc. of the 5th International Conference on Spoken Language Processing*, Vol. 4, p. 1295-1298, Sydney, Australia.

[12] Kang, K-H, Guion, S. G. 2008. Clear speech production of Korean stops: changing phonetic targest and enhancement strategies. *J of Acous. Soc. of Am.* 124, 3909-3917.

[13] Kim, M.J., Pae, S. 2005. The percentage of consonants correct and the ages of consonantal acquisition for 'Korean-test of articulation for children (K-TAC)'. *J. of Korean Association of Speech Science*, 12, 139-149.

[14] Kim, M., Stoel-Gammon, C. 2009. The acquisition of Korean word-initial stops. *J of Acous. Soc. of Am.* 125. 3950-3961.

[15] Kong, E.J., Beckman, M.E., Edwards, J. 2011. Why are Korean tense stops acquired so early?: The role of acoustic properties. *J. Phon.*, 39, 196-211.

[16] Silva, D.J. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23, 287-308.

[17] Yang, A. 2017. *The acquisition of prosodic focus-marking in Mandarin Chinese- and Seoul Korean-speaking children*. Ph.D dissertation. University of Utrecht.

[18] Yang, A., Chen, A. 2014. Prosodic focus marking in Chinese four- and eight-year-olds. *Proc. 7th Speech Prosody*, Dublin.

[19] Yang, A., Cho, T., Kim, S., Chen, A. 2015. Phonetic focus-marking in Korean-speaking 7- to 8-year-olds and adults. *Proc. 18th ICPhS*, Glasgow.

[20] Xu, Y. 1999. Effects of tone and focus on f0 contour formation. *J. Phon.*, 27, 55-105.

[i] (DV ~ 1 + Consonant Type + (1+ Consonant Type | subject)). lme4 package (Bates et al., 2015) in R was used.