Input distribution of Korean affricates in the context of a diachronic sound change: comparison between IDS and ADS

Minji Nam[a], Youngon Choi[a], Hyun Kyung Hwang[b], Mazuka Reiko[c,d,e]

[a]Chung-Ang University, Korea, [b]University of Tsukuba, Japan, [c]RIKEN Center for Brain Science, Japan,
[d]Waseda University, Japan, [e]Duke University, USA
minji.namm@gmail.com; yochoi@cau.ac.kr; hwang.kyung.gu@u.tsukuba.ac.jp; reiko.mazuka@riken.jp

ABSTRACT

Compared to stops, little is known about the distributions of affricates along the acoustic spectrums provided by caregivers for their infants. Korean affricates have three-way contrasts (fortis, lenis, and aspirated), all of which used to be distinguishable by frication and aspiration duration. However, due to a diachronic change, lenis no longer differs from aspirated in the durational cue but differs in fundamental frequency ($f0$) from aspirated.

We examined the distributional pattern of maternal input along these acoustic parameters in their speech to infants and adults. Like stop categories, mothers' productions reflected the diachronic sound change regardless of speech register, showing the difference in $f0$ not in duration for lenis-aspirated contrast. However, the cue differences were selectively enhanced for relevant contrasts in infant-directed speech: larger $f0$ differences for the lenis-aspirated contrast and greater durational differences between fortis and other categories. This might provide additional phonetic support for young Korean learners.

Keywords: affricates, Korean, phonetic input, diachronic change, speech register

1. INTRODUCTION

An affricate is a type of consonant that is produced by combining the articulation of a plosive with the fricative [1]. Therefore, both the duration of frication noise and voice onset time (VOT) play crucial roles in distinguishing affricate categories. Affricates have posed challenges in specifying their underlying phonetic structure [1], and have received relatively little attention in the field of speech development due to their unique combination of articulatory features.

Korean presents an even more challenging case for learners because it has typologically unusual three-way affricate contrasts, which are entirely voiceless. The three affricate categories in Korean are known as fortis (or tense), lenis (lax or plain), and aspirated (/c*, c, cʰ/, see examples in Table 1). Among the three categories, fortis has the shortest total noise duration (combining durations of frication and aspiration, [1]). The aspirated category used to be produced with the longest noise duration, followed by that of lenis (intermediate duration between fortis and aspirated, [2]).

However, like three-way stop categories in Korean, two of the three affricate categories—lenis and aspirated—have undergone a diachronic phonetic change in the past several decades [3]–[6]. As a result, noise duration has shortened for aspirated, largely overlapping with the range of noise duration for lenis. By contrast, the fundamental frequency ($f0$) difference was enhanced between the lenis and aspirated categories [1], [2].

Adult native Korean speakers relied on durational cues in categorizing and discriminating fortis from other categories [1], [5], whereas $f0$ was the main cue to distinguish lenis from others [1]. These demonstrated that native Korean adults utilized the relevant acoustic cues to distinguish the target affricate categories, showing that their perception pattern reflects the shift in the use of phonetic cues.

These results raise a possibility that the Korean caregiver's phonetic input of affricates to young learners has characteristics reflecting the phonetic changes. However, it is also possible that caregiver input may contain exaggerated acoustic parameters to support early learners [7]–[10] such that the durational and frequency cues are presented with greater differences than the typical input observed among speech directed to adults. Early maternal phonetic quality is associated with infants' consonant discrimination abilities [10]–[12], the child's later language outcomes [13], and children's production [14].

In the context of a diachronic change, it is also possible that individual caregivers may differ in terms of reflecting the change in their input to young infants, as reported by Choi et al. [7]. This may result in mixed distributions along the acoustic dimensions.

To our best knowledge, no studies have examined the input distribution patterns of the Korean affricates along the acoustic parameters of noise duration and $f0$ available to young Korean-learning infants. Therefore, the present study investigated the phonetic input pattern of three affricate categories in the

speech of mothers whose infants are between 4 and 10 months of age.

Korean	Meaning	IPA	Category
짜다	salty	t'sada	Fortis
자다	sleep	tsada	Lenis
차다	kick	tsʰada	Aspirated

Table 1: Examples of words used to elicit word-initial three-way affricate categories in Korean.

2. DATA & METHODS

2.1. Participants and Materials

Thirty-six native Korean-speaking mothers participated in the recording session conducted in our lab. The age of the mothers ranged from 23 to 39 years (mean = 31.9 years) and their infants were between 4 and 10 months of age (mean = 6.8 months). All the mothers reported that they grew up and currently reside in Seoul.

A list of three sets of minimal triplet words (짜다/자다/차다, 찌다/지다/치다, 짱구/장구/창구) was provided for the mothers and they were asked to utter each word in the carrier sentence as below in two speech registers: infant-directed speech (IDS) and adult-directed speech (ADS). In the present paper, we specifically report the analyses of the three words listed in Table 1 from the full list.

- ADS: 이것은 ＿＿＿라고 합니다.
  (ikʌsɨn ＿＿＿ lako hapnita,
  'This is ＿＿＿'.)
- IDS: 이것은 ＿＿＿래.
  (ikʌsɨn ＿＿＿ le, 'This is ＿＿＿'.)

In the IDS register, mothers were provided with a set of picture cards depicting target words. They were asked to show the infant the picture cards while recording. In the ADS register, mothers were asked to read the target words from the list as if casually speaking to another adult. The target words were recorded five times for each speech register.

Some mothers missed reading some of the target words. Some repeated the word less than five times. We included partial sets of obtained tokens in the analyses. However, we excluded tokens with noise. As a result, a total of 883 tokens were included in the final analyses, 463 for ADS and 420 for IDS, respectively.

2.2. Procedure

The mothers were taken into a small, sound-treated room with their infants. The utterances were recorded using a portable Marantz digital recorder (PMD 661MKII) and a head-mounted microphone (Shure SM10A).

2.3. Measurements

The noise duration, including frication duration, and aspiration duration, and $f0$ in the following vowel were measured by implementing a Praat script [15]. Frication duration (Fric.) was determined by the presence of high-frequency, high-amplitude noise in the spectrogram. Aspiration duration (Asp.) was measured from the offset of high-frequency noise to the onset of the following vowel and was identified by low amplitude and low-frequency noise. F0 was measured by converting the average wavelength of the first 25 milliseconds of the vowels. Vowel onset was indicated by the first zero-crossing of periodicity after the aperiodic noise.

3. RESULTS

3.1. Acoustic characteristics of the maternal input

The distribution patterns along the noise duration and $f0$ were very similar across the two speech registers. As shown in Figure 1 and Table 2, the noise duration of fortis affricates was much shorter than the other two categories. However, the range of the noise duration for lenis affricates overlapped with those of aspirated affricates in both ADS and IDS.

We estimated the measured durations of Fric., Asp., noise, and $f0$ properties using linear mixed-effects regression models (*lmer* in lme4 package in R version 4.1.0, [16]). Model comparison was done using the likelihood ratio test. The full model included the fixed effects of category (fortis, lenis, & aspirated), speech register (IDS, ADS), and their interactions. Speaker was included as random effects. Model estimates with the absolute value of the t-statistic greater than 2 were considered significant at a 5% significance level [17].

For all the acoustic parameters, significant main effects were found for both the category and register. The interaction effects between the two were also significant. The estimated $f0$ was significantly lower for lenis ($\beta = -95.93$, $t = -20.21$) than aspirated ($\beta = 279.88$, reference category) and this difference was enhanced ($\beta = -42.05$, $t = -6.05$) in IDS. The estimated $f0$ was also significantly lower for fortis ($\beta = -41.18$, $t = -8.80$) than aspirated but the difference

Comparison pair	Acoustic Properties	Speech Register	Mean (difference)	SD
Aspirated – Fortis	Fric. duration differnce (ms)	ADS	9.43	7.88
		IDS	14.17	13.44
	Asp. Duration difference (ms)	ADS	46.89	14.63
		IDS	55.75	21.70
	Noise duration difference (fric. + asp., ms)	ADS	53.11	15.72
		IDS	69.50	21.62
	$f0$ differnece (Hz)	ADS	39.53	27.17
		IDS	54.98	38.45
Lenis – Fortis	Fric. duration differnce (ms)	ADS	9.45	7.92
		IDS	13.41	11.49
	Asp. Duration difference (ms)	ADS	56.36	18.41
		IDS	66.29	35.78
	Noise duration difference (fric. + asp., ms)	ADS	62.69	17.06
		IDS	78.51	31.72
	$f0$ differnece (Hz)	ADS	66.84	43.14
		IDS	114.93	80.73
Aspirated – Lenis	Fric. duration differnce (ms)	ADS	5.77	4.09
		IDS	4.57	4.81
	Asp. Duration difference (ms)	ADS	13.33	10.25
		IDS	19.30	16.36
	Noise duration difference (fric. + asp., ms)	ADS	12.49	11.33
		IDS	18.99	16.00
	$f0$ differnece (Hz)	ADS	97.25	30.08
		IDS	139.00	54.12

Table 2. Mean differences in acoustic properties among the three sound pairs in ADS and IDS register.

was not greater in IDS ($\beta = -1.12$, $t = -0.87$) Removing the interaction effect significantly reduced model fit ($\chi^2(2) = 45.70$, $p < .001$), suggesting that $f0$ cue was enhanced selectively for lenis-aspirated contrast.

The estimated noise duration (Fric. + Asp.) was shorter for fortis ($\beta = -53.42$, $t = -28.04$) and longer for lenis ($\beta = 11.84$, $t = 6.13$) than aspirated. The durational difference was greater in IDS ($\beta = -15.47$, $t = -5.65$) for fortis-aspirated but no change was observed for lenis-aspirated contrast ($\beta = 1.45$, $t = 0.51$). Removing the interaction effect significantly reduced model fit ($\chi^2(2) = 44.92$, $p < .001$), supporting that the combined durational cue was enhanced selectively for fortis-aspirated contrast in IDS. Similar patterns were observed with both Fric. and Asp. duration measure.

3.2. Individual differences in how mothers speak the three-way affricates: modern way vs. the traditional way

Lastly, we examined the variability in the production of durational cues and $f0$ cue among these mothers.

According to Lee [2], the estimated noise duration differences between the aspirated and lenis categories were about 10ms and 75 Hz (cf. Fig. 1 in Lee [2]) for the estimated $f0$ differences among speakers in the younger age group. Based on the estimated average differences in the two cues produced by each mother, we categorized them into two groups: (1) those who produce the cue "the modern way" and (2) those that
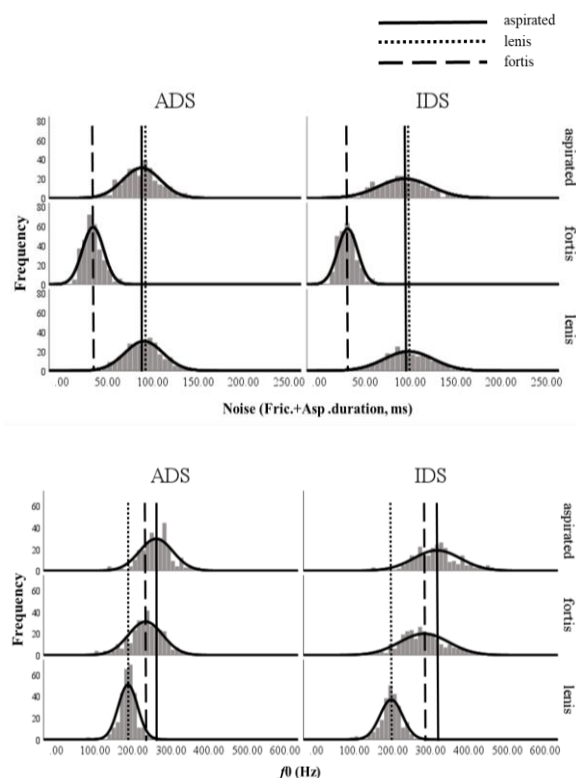

Figure 1: Distribution of noise duration (frication + aspiration) and $f0$ cues in three affricate categories (top panel: aspirated, middle panel: fortis, & bottom panel: lenis) produced by Korean mothers in ADS (left panels in each) and IDS (right panels in each). Vertical lines indicate mean values.

produce the cue "the traditional way." The criterion for the noise duration for this division was 10ms: if the mothers on average produced a noise duration difference shorter than 10ms, they were considered as producing the cue "modern way" but as "traditional way" if the difference was greater than 10ms. A similar approach was applied to the use of the $f0$ cue, so using the 75Hz as the criterion, those mothers that produced the $f0$ difference between aspirated and lenis greater than 75Hz were considered to produce the cue "modern way" and vice versa for other mothers.

Among 35 mothers in ADS speech (1 mother did not record ADS), 20 mothers (67%) produced VOT in a modern way, whereas 15 of them still produced noise duration difference in the traditional way. For $f0$, 28 mothers (80%) produced the difference in line with the recent change, but 7 of them (20%) still produced an $f0$ difference of less than 75Hz, following the traditional pattern. Mothers who produced both cues, reflecting the complete diachronic change were 46% (see Table 3). The mean and standard deviation of mothers' cues split by this grouping are summarized in Table 3.

In IDS, among 31 mothers (1 mother did not record IDS, 4 mothers did not record either lenis or

	ADS AspLenis Noise duration (ms)		ADS AspLenis *f*0 (Hz)		IDS AspLenis Noise duration (ms)		IDS AspLenis *f*0 (Hz)	
	modern	traditional	modern	traditional	modern	traditional	modern	traditional
Mean	4.55	23.23	107.15	55.33	5.85	25.77	149.88	54.72
SD	3.77	8.68	22.01	15.8	3.32	15.24	44.92	19.34

Table 3. Mean and SD in noise duration and *f*0 of mothers who produced the acoustic cues in modern versus traditional way in ADS and IDS.

aspirated words), 10 mothers (32%) produced noise duration in the modern way, whereas 21 of them (68%) still produced noise duration difference in the traditional way (Table 4). For *f*0, 27 mothers (87%) produced the difference in line with the recent change, but 4 of them (13%) still produced an *f*0 difference of less than 75Hz, following the traditional pattern. Mothers who produced both noise duration and *f*0 reflecting the complete diachronic change were about 29% (*n* = 9). Taken together, we observed that more mothers produced longer noise duration (considered traditional way) in IDS than ADS (21, 15, respectively, fisher's exact test, *p* = .051). However, the group variation on the *f*0 cue did not differ much between the two registers (fisher's exact test, *p* = .659).

Overall, the cue production patterns revealed that not all mothers today are faithful in reflecting the shifted cue pattern, especially for noise duration in IDS.

		Speech Register		
Diachronic Change		ADS	IDS	Total
Traditional	Count	3	3	6
	% of total	5%	5%	9%
Modern	Count	16	9	25
	% of total	24%	14%	38%
Mixed	Count	16	19	35
	% of total	24%	29%	53%
Total	Count	35	31	66
	% of total	53%	47%	100%

Table 4. The distribution of mothers in reflecting the diachronic change in two acoustic cues (noise duration & f0): traditional (no reflection of cue change in both acoustic cues), modern (reflecting change in both cues), & mixed (one cue reflecting the change, the other preserving traditional cue pattern).

4. DISCUSSION

The results of our acoustic analyses showed that native Korean-speaking mothers today provide their young infants with affricate phonetic input that largely reflects the phonological changes that have occurred in the past several decades. The input distribution patterns were overall very similar between mothers' speech to infants as well as to other adults.

These suggest that, unlike in the past, durational cues may no longer play key roles in discriminating between lenis and aspirated as young Korean learners derive these categories, like Korean-speaking adults [1].

However, we observed that some mothers still tried to supply the durational cue difference for the contrast as suggested in our examination of individual differences among mothers (see a similar report on Korean stop input patterns, [7]). This raises an interesting possibility that their infant learners might diverge in terms of relying on different cues, given the strong relationship between children's production and their mothers' quality of production was observed in a previous study [14]. The infants of mothers who provided salient durational and *f*0 cues might also develop sensitivity to this cue, whereas those infants whose mothers consistently relied on a single cue might develop sensitivity to that cue alone. Future studies can illuminate this by testing Korean-learning infants' developing sensitivities to these cues and how they weigh them as they derive these phoneme categories.

Importantly, the specifically relevant cue parameter for category distinction was selectively enhanced in mothers' speech to their infants. In IDS, mothers produced larger differences in *f*0 between lenis and aspirated. By comparison, the durational difference was enlarged for fortis and other categories but not for lenis and aspirated contrast in IDS. These were consistent with many previous reports on the enhanced cue distinctions in IDS than ADS (e.g., [8]). These results suggest that early maternal phonetic input may scaffold the early development of native phonology.

6. REFERENCES

[1] J. Schertz and Y. Kang, "Phonetic cue competition within multiple phonological contrasts: Perception of Seoul Korean sibilants," *Korean Linguist.*, vol. 18, no. 1, pp. 1–17, 2022.

[2] H. Lee, "Consistent sound change between stops and affricates in Seoul Korean within and across individuals: A diachronic investigation," *J. Acoust. Soc. Am.*, vol. 140, no. 6, pp. EL491–EL496, Dec. 2016, doi: 10.1121/1.4971203.

[3] H. Bang, M. Sonderegger, Y. Kang, M. Clayards, and T. Yoon, "The effect of word frequency on the timecourse of tonogenesis in Seoul Korean," in *Proceedings of the 18th International Congress of Phonetic Science (ICPhS2015)*, 2015.

[4] Y. Kang, "Voice onset time merger and development of tonal contrast in seoul korean stops: A corpus study," *J. Phon.*, vol. 45, no. 1, pp. 76–90, 2014, doi: 10.1016/j.wocn.2014.03.005.

[5] M. Kim, "Correlation between VOT and F0 in the Perception of Korean Stops and Affricates," in *INTERSPEECH 2004 - ICSLP (8th International Conference on Spoken Language Processing), Jeju*, 2004, pp. 1–6.

[6] D. J. Silva, "Acoustic evidence for the emergence of tonal contrast in contemporary Korean Acoustic evidence for the emergence of tonal contrast in contemporary Korean *," *Phonology*, vol. 23, pp. 287–308, 2006, doi: 10.1017/S0952675706000911.

[7] Y. Choi, M. Nam, R. Mazuka, H.-K. Hwang, and N. Yamane, "Korean mothers' production of laryngeal stops to their infants as compared with adults in the context of tonogenesis," in *the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 2019, pp. 3260–3264.

[8] H.-M. Liu, P. K. Kuhl, and F.-M. Tsao, "Mothers' speech clarity and infants' discrimination An association between mothers' speech clarity and infants' speech discrimination skills," *Dev. Sci.*, vol. 6, no. 3, pp. F1–F10, 2003.

[9] C. Narayan and T.-J. Yoon, "VOT and f0 in Korean infant-directed speech," *Can. Acoust.*, vol. 39, no. 3, pp. 152–153, 2011.

[10] E. Ko, "Evaluating the clear speech hypothesis of child-directed speech: The case of Korean stop consonants," presented at the The Korean Society of Speech Sciences (KSSS 2016), Seoul, Korea, 2016.

[11] A. Cristià, "Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category," *J. Acoust. Soc. Am.*, vol. 129, no. 5, pp. 3271–3280, May 2011, doi: 10.1121/1.3562562.

[12] M. Nam, Y. Choi, and R. Mazuka, "Caregiver phonetic input supports infants' learning of sound categories: The role of voice onset time distribution in natural phonological acquisition," presented at the in the 46th Boston University Conference on Language Development (BUCLD 2021), Boston, US, 2021.

[13] K. M. Hartman, N. B. Ratner, and R. S. Newman, "Infant-directed speech (IDS) vowel clarity and child language outcomes," *J. Child Lang.*, vol. 44, no. 5, pp. 1140–1162, Sep. 2017, doi: 10.1017/S0305000916000520.

[14] J. H. Sim and B. Post, "Variation in quality of maternal input and development of coda stops in English-speaking children in Singapore," *J. Child Lang.*, vol. 49, no. 6, pp. 1147–1172, Nov. 2022, doi: 10.1017/S0305000921000593.

[15] Xu, Y. 2013. ProsodyPro — *A Tool for Large-scale Systematic Prosody Analysis*. In Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France. 7-10.

[16] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models using lme4." arXiv, Jun. 23, 2014. Accessed: Jan. 07, 2023. [Online]. Available: http://arxiv.org/abs/1406.5823

[17] R. H. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," *J. Mem. Lang.*, vol. 59, no. 4, pp. 390–412, Nov. 2008, doi: 10.1016/j.jml.2007.12.005.