



## Executive function and Korean children's stop production\*

Eun Jong Kong<sup>1</sup> · Hyunjung Lee<sup>2,\*\*</sup> · Jeffrey J. Holliday<sup>3</sup>

<sup>1</sup>*School of Liberal Arts and Sciences, Korea Aerospace University, Goyang, Korea*

<sup>2</sup>*Department of English Language Education, Incheon National University, Incheon, Korea*

<sup>3</sup>*Department of Linguistics, University of Kansas, Lawrence, KS, USA*

### Abstract

Previous studies have established a role for cognitive differences in explaining variability in speech processing across individuals. In the case of perceptual cue weighting in the context of a sound change, studies have produced conflicting results regarding the relationship between executive function and the use of redundant cues. The current study aimed to explore this relationship in acoustic cue weighting during speech production. Forty-one Korean-speaking children read a list of stop-initial words and completed two tests that assess executive function, i.e., Dimensional Change Card Sorting (DCCS) and digit n-back. Voice onset time (VOT) and fundamental frequency (F0) were measured in each word, and analyses were carried out to determine the extent to which children's executive function predicted their use of both informative and less informative cues to the three pairs comprising the Korean three-way stop laryngeal contrast. No evidence was found for a relationship between cognitive ability and acoustic cue weighting in production, which is at odds with previous, albeit conflicting, results for speech perception. While this result may be due to the lack of task demands in the production task used here, it nevertheless expands the empirical ground upon which future work in this area may proceed.

**Keywords:** acoustic cue weighting, executive function, stop laryngeal contrast, sound change in progress, voice onset time (VOT), fundamental frequency (F0), child speech production

### 1. Introduction

Individual variation is observed at all times during phonetic categorization due to the existence of multiple acoustic cues that can blur the phonetic boundary within a phonological contrast, as each listener (and speaker) can give different weightings to each of these cues. For example, when categorizing the stop voicing contrast in English (e.g., /b/-/p/), listeners utilize both voice onset time (VOT)

and fundamental frequency (F0) at the onset of the following vowel as a primary and a secondary cue, respectively. But male listeners tend to weigh VOT less compared to female listeners, with this gender-related tendency associated with their subjective evaluation of the talker (Yu, 2022).

While previous sociophonetic studies examining such variation at the population level have mostly focused on socio-indexical categories such as dialect, age, gender, or socio-economic status,

\* This work was supported by Incheon National University Research Grant in 2019.

\*\* [hjlee123@inu.ac.kr](mailto:hjlee123@inu.ac.kr), Corresponding author

Received 20 August 2023; Revised 11 September 2023; Accepted 11 September 2023

© Copyright 2023 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

other intrinsic human characteristics have been recently noted as a meaningful variable in understanding systematic individual variation in speech processing (e.g., Chandrasekaran et al., 2010; Clayards, 2018; Francis & Nusbaum 2009; Idemaru et al., 2012; Kong & Edwards, 2016; Ou et al., 2023; Pisoni, 1993; Yu, 2010, 2022; Yu & Lee, 2014; Yu et al., 2011). This line of research has tried to understand how individuals' internal attributes, including cognitive abilities, are related to phonetic variation, and has examined the relationship between individuals' intrinsic traits (e.g., personality, working memory, mental flexibility, attention control) and speech processing styles in an effort to determine the internal source of linguistic variation. Given that linguistic changes or sound changes start off with idiolectal differences (Yu, 2023), understanding the relationship between individuals' cognitive characteristics and linguistic variation would broaden our knowledge regarding linguistic changes within a language community.

The three-way laryngeal contrast of Korean stops has changed over decades in a way that the relative importance of the multiple acoustic cues to the contrast has reversed, which suggests that the Korean stop contrast can be a good site to examine the relationship between individuals' cognitive resources and variation in phonetic categorization. Specifically, similar to the English stop voicing contrast, Korean stops are also differentiated by multiple acoustic cues, namely VOT and F0. But due to the aforementioned sound change, each acoustic dimension can be primary or secondary depending on the phoneme pair within the contrast, and therefore we can cross-check the relationship between individuals' cognitive resources and the use of an acoustic dimension across the pair types within the three-way contrast. Importantly, under the sound change in progress, the cue primacy between VOT and F0 varies across individuals, particularly for the lenis and aspirated stops. Therefore, exploring the interaction between conservative and innovative cues and individual cognitive abilities may help us better understand the actuation of sound change within a language community. As a follow-up study of Kong & Yoo (2017)'s perception investigation, the present paper particularly focuses on Korean children's cognitive abilities and their use of the multiple acoustic dimensions in the production of the three-way laryngeal stop contrast.

### 1.1. Korean Stop and Executive Functions

Korean has a three-way laryngeal contrast among fortis, lenis and aspirated stops at three places of articulation (i.e., /p', t', k'/, /p, t, k/, /p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/). A great body of work has explored the unusual three-way laryngeal contrast of Korean stops in perception and production over several decades (e.g., Cho et al., 2002; Han & Weitzman, 1970; Kang & Guion, 2008; Kim, 2000; Kong et al., 2011; Lisker & Abramson, 1964; Oh, 2011; Silva, 2006; Wright, 2007; see Lee et al. (2020) for a review). Traditionally, VOT was the sole primary cue for all pairs within the contrast (e.g., Han & Weitzman, 1970; Hardcastle, 1973; Lisker & Abramson, 1964). But under the current sound change in progress, F0 has overtaken VOT as a primary cue, particularly for the lenis-aspirated stop pair (e.g., Kang, 2014; Silva, 2006), and this sound change has spread from standard Seoul Korean into regional dialects, such as the Gyeongsang dialects (Lee & Jongman, 2012; Lee et al., 2013). Specifically, F0 values pattern in the order of lenis<fortis<aspirated stops, and VOT patterns in the order of fortis<lenis=aspirated stops. That is, the importance of the two acoustic dimensions depends on the specific laryngeal contrast: VOT plays an important role in distinguishing

the fortis stop from the other two stops, while F0 is necessary for the lenis-aspirated contrast. Therefore, effective listeners might establish a perceptual strategy by relying on a useful cue and ignoring a less informative cue to identify the stop correctly.

The changing role of VOT and F0 is also reflected in the development of Korean-speaking children's stop production and perception (e.g., Kim, 1999; Kim & Stoel-Gammon, 2009; Kong et al., 2011; Son, 2018). In previous literature, children under age 10 showed significantly shorter VOTs for the fortis stop compared to non-fortis stops. However, there was no difference in VOT between the lenis and aspirated stops, and F0 made up for the lack of a VOT distinction between the lenis-aspirated pair. This stop production pattern in children was confirmed by perception data from adult listeners, who reliably used VOT in identifying children's production of the fortis stop (Kong et al., 2011). In both production and perception studies with 3-year-old children, Son (2018) observed the critical role of F0 for the distinction between lenis and aspirated stops. This previous literature confirmed the reduced role of VOT and enhanced role of F0 based not only on adults' speech, but also on children's perception and production.

Group differences have also been observed for such phonetic trade-off relations between the two acoustic dimensions in a way that the degree of reliance on F0 (or VOT) varies across speaker/listener characteristics such as gender (Kang, 2014) or dialect (e.g., Lee & Jongman, 2012). Focusing on human-internal resources inducing such phonetic variation, a few perception studies have examined correlations between individuals' cognitive abilities and perceptual strategies related to acoustic cue use (e.g., Kong & Yoo, 2017; Lee & Kong, 2016). Lee & Kong (2016) examined the relationship between individual traits and variability in processing Korean stops by testing 28 adult Korean-speaking listeners. They conducted a three-alternative forced-choice stop perception task (3AFC) in two different conditions: in one condition listeners just completed the stop identification task, and in the other condition listeners also had to complete simple arithmetic problems that served as distractors. In addition to the speech perception tasks, a series of executive function (EF) tasks assessed each participants' cognitive abilities through a *digit n-back* task, a *flanker* task, and a *trail-making* task for working memory capacity (WM), inhibition, and cognitive flexibility, respectively. In Lee & Kong (2016), a meaningful correlation between individuals' cognitive traits and the use of multiple cues in perception was observed in that individuals with better executive function were more likely to utilize the less informative VOT in distinguishing between the lenis and aspirated stops. Based on these findings, Lee & Kong (2016) suggested that individual language users' cognitive traits could be one factor explaining systematic individual variation.

Kong & Yoo (2017) examined how 7- to 8-year-old Korean-speaking children utilized VOT and F0 in the perception of Korean stops, and tested whether and how their cognitive abilities are associated with their perceptual strategies in using multiple acoustic dimensions. 15 children's cognitive abilities were assessed by testing their working memory, inhibition, and shifting attention abilities. These cognitive abilities were measured by a series of executive tasks including a *digit n-back* task (WM), a *flanker* task (inhibition), a *Stroop* test (inhibition) and a *trail-making* (attention) test in which accuracy and response times (RT) were obtained. Overall, in Kong & Yoo (2017), the individual observations indicated a meaningful correlation between executive function and the use of acoustic cues

in the perception of Korean stops. Based on a series of correlation tests, Kong & Yoo (2017) found a strong correlation between the use of VOT and F0 cues and children's cognitive abilities in identifying the pairs of stop contrasts. Specifically, child listeners with better WM and inhibition control exhibited more effective perceptual strategies by relying less on the unimportant F0 cue in the identification of the fortis-aspirated stops. Likewise, individuals with better attention reported effective use of an informative cue (i.e., VOT) in identifying the fortis stop from the other two stops. Overall, Kong & Yoo (2017) showed that children with better cognitive abilities showed more effective perceptual strategies by effectively utilizing a more informative cue and ignoring a less informative one. This observation is somewhat opposite from Lee & Kong (2016), in which adult listeners with better cognitive ability tended to utilize an inefficient cue, but may reflect children's more limited cognitive resources compared to adults' in processing sub-phonemic information in speech perception.

## 1.2. Current Study

From the previous studies (Kong & Yoo, 2017; Lee & Kong, 2016), the following questions arise. First, given that the previous observation was limited to speech perception, the present study asks whether and how this relationship between individuals' perceptual strategies and cognitive resources are reflected in the production of Korean stops, at least in the case of children. Hypothesizing consistency between speech perception (Kong & Yoo, 2017) and production, we might predict that children with better cognitive capacity or EF scores would be more likely to use an informative acoustic cue and suppress a redundant cue (e.g., enhancing VOT, but suppressing F0 for the fortis-aspirated contrast) in producing Korean stops. In a similar vein, we also aim to closely look into the presence or absence of significant correlations considering stop pair types and EF tasks. While it is expected to find a meaningful association between better EF and children's accurate pronunciation of the phonemes, as existing production studies showed (e.g., Eaton & Ratner, 2016; Netelenbos et al., 2018), the present study uniquely investigates sub-phonemic acoustic properties of the sounds in children's productions in relation to EF. Finally, the present study will also expand the empirical foundation for exploring individual variation in cognitive factors as they relate to speech science, allowing us to revisit the empirical rift between Kong & Yoo (2017) and Lee & Kong (2016). For these purposes, we examined the production data of forty-one Korean-speaking children, consisting of a subset of the data presented in Holliday et al. (2023).<sup>1</sup> While the statistical analyses presented here consider gender and dialect as factors, the discussion of the present paper mainly focuses on the relationship between individual variation in cue-weighting and cognitive capacities estimated by EF tasks.

## 2. Methods

### 2.1. Participants

Forty-one Seoul (13 F, 10 M) and Gyeongsang Korean (GK, 12 F, 6 M) elementary school children participated in the experiment

and were given monetary compensation. The children were recruited in Seoul and Changwon (a city in South Gyeongsang province) and tested either at their home or at the elementary school they attended [see Holliday et al. (2023) for more methodological details]. All were between age 9 and 10 at the time of testing.

### 2.2. Tasks and Procedures

**Speech production:** The children read a list of stop-initial words (Appendix 1, three repetitions of 18 words in a randomized order) presented on a printed paper. They could take their time, with no time restriction, and on average the recording session took about 15 minutes. While researchers did not provide any feedback on speed or accuracy during the session, a few participants self-corrected their mispronunciations, in which case (19 tokens) the corrected productions were used for acoustic analysis. We used digital recorders and condenser microphones for the task, digitizing the recordings at a 44,100 Hz sampling rate and 16-bit quantization.

**Executive function:** We administered Dimensional Change Card Sorting (DCCS) and digit n-back tasks to assess children's mental flexibility and WM, respectively (Owen et al., 2005; Zelazo et al., 2003). (It is noted that besides the two EF tasks, a Stroop test was also given to a subset of participants as an inhibition control measure. However, we could not administer the task to every participant due to attention and time constraints.) In each DCCS trial, participants were asked to sort cards (e.g., 'blue ball' and 'yellow truck') as fast as possible according to either color (blue or yellow) or shape (ball or truck). Both the counts and RT of correct responses were collected when the two cues *shifted* (e.g., 'color' followed by 'shape') and *stayed* (e.g., 'color' followed by 'color'). RT differences between *shift* and *stay* conditions and accuracy of the *shift* condition were regarded as cognitive cost of mental flexibility. In the digit n-back task, children were asked to decide as fast as possible whether the single digit on the current slide is identical (*same*) to or different (*different*) from that of the previous slide. In consideration of time and the difficulty of the task, we administered only the 1-back session, in which children were to compare the two numbers one slide apart. Similar to the DCCS, RT difference (RT[*same*]–RT[*different*]) and accuracy counts on the *different* trials were used as numeric indices to represent the participants' WM. Both tasks were programmed in E-Prime (ver. 3) to automatically record response details.

### 2.3. Analysis

**Acoustic analysis:** We measured stop VOT and the F0 of the following vowel using Praat (Boersma & Weenink, 2023). For VOT, we manually pinpointed the stop burst and the voicing onset in the following vowel, using a Praat script only to locate each target word. For F0, the script first computed F0 by averaging the pitch values over the initial 15 ms beginning at the voicing onset. If there was no measurable F0 during the initial 15 ms due to creak, the script moved forward in 5 ms steps until a measurable F0 value was obtained.

**Statistical analysis:** Three logistic mixed-effects regression models were built to predict the stop category among the three contrasts

<sup>1</sup> Holliday et al. (2023) examined the three-way laryngeal contrast of stops produced by 197 Korean speakers across dialect, age, and gender.

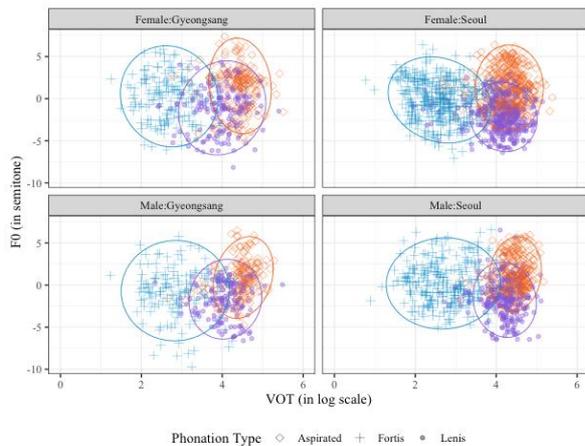
dependent variable (DV pairs: lenis-aspirated, fortis-aspirated, and fortis-lenis stops) based on fixed effect variables of VOT and F0 along with their interactions with gender (male vs. female) and dialect region (GK vs. Seoul). By-subject intercepts and slopes of VOT and F0 were included as random effects, from which individual speakers' coefficients for VOT and F0 were estimated.

As preliminary simple correlation tests revealed that the VOT and F0 coefficients were highly correlated [Fortis-Lenis model:  $r(39)=.84$ ,  $p<.0001$ ], a series of partial correlation tests were conducted to define the relationship between EF capacities and acoustic cue utilizations in producing the stops. The partial correlation tests take each EF score (RTs and accuracies of DCCS and digit 1-back tasks) as one test variable (X), each acoustic variable coefficient (VOT and F0) as the other test variable (Y), and either F0 or VOT coefficients was taken as a control variable (Z). Because the test yields correlation coefficients between X and Y subtracting Z's correlations with X and/or Y, we can avoid misleading estimates of association between X and Y that actually come from the confounding variable Z. We used the *ppcor* package in RStudio (Kim, 2015; Posit Team, 2023).

### 3. Results

#### 3.1. Stop Production: Voice Onset Time (VOT) and Fundamental Frequency (F0) Weighting

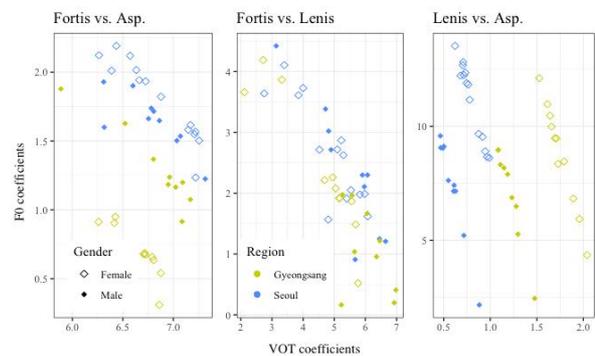
Figure 1 displays F0 values (converted to semitone) as a function of VOT values (log-transformed) separated by gender and dialect region. Overall, all groups of the children exhibited similar distributions of VOT and F0. Children from both dialects differentiated aspirated stops from fortis stops by having longer VOT, although the two dialect groups differed slightly in their realization of lenis stops: GK children's lenis stops were more or less between fortis and aspirated in the VOT and F0 dimensions, but those of Seoul children rarely overlapped with fortis stops.



**Figure 1.** VOT and F0 distributions of the Korean stops (Fortis vs. Lenis vs. Aspirated) separated by dialect region and gender.

These rough observations were confirmed in the statistical analyses in which the three pairs of the stops were predicted by VOT and F0. Table 1 summarizes the mixed-effects regression models for the three pairs of contrastive stops (i.e., Lenis-Asp., Fortis-Asp., Fortis-Lenis models). For the fortis-aspirated stop model, the fixed effect coefficient of VOT but not that of F0 was

statistically meaningful, while both VOT and F0 were significant predictors for the lenis-aspirated and the fortis-lenis stop models. There were no dialect or gender group interactions in the two models, while the lenis-aspirated model yielded a marginally significant interaction of VOT with Region in which a negative interaction coefficient ( $\beta_{\text{VOT:Region}}=-.97$ ,  $SE=.56$ ,  $p=.08$ ) indicates GK children's VOT coefficients were greater than those of Seoul children. Figure 2 presents individuals' VOT and F0 coefficients from the models (random coefficients added to group-averaged fixed effect coefficients), showing a trend that GK children's VOT coefficients were slightly greater than those of Seoul children. The current result finding no dialect group difference supports recent studies examining GK speakers' stop productions in which children realized the innovative variants of lenis and aspirated stops similarly to Seoul speakers (Lee, 2020).



**Figure 2.** Individual children's VOT and F0 coefficients from the three mixed-effects regression models: Fortis-Asp., Fortis-Lenis, and Lenis-Asp. models.

**Table 1.** Summary of fixed effect coefficients (model estimates and std. errors)

	Estimate (SE)		
	Lenis-Asp.	Fortis-Asp.	Fortis-Lenis
(Intercept)	-0.30 (1.09)	-0.48 (0.48)	-0.23 (0.44)
VOT	<b>1.73 (0.45)</b>	<b>6.69 (1.26)</b>	<b>4.82 (0.90)</b>
F0	<b>9.09 (1.65)</b>	0.68 (0.62)	<b>-2.23 (0.89)</b>
VOT×Region	<b>-0.97 (0.56)</b>	0.30 (1.31)	0.26 (1.02)
F0×Region	2.58 (2.03)	1.02 (0.87)	-0.34 (1.05)
VOT×Gender	-0.55 (0.58)	0.12 (1.48)	1.43 (1.30)
F0×Gender	-1.47 (2.04)	0.65 (0.95)	1.31 (1.14)
VOT×Region×Gender	0.35 (0.74)	-0.38 (1.93)	-0.88 (1.66)
F0×Region×Gender	-2.24 (2.75)	-0.63 (1.25)	-1.07 (1.52)

Bold indicates  $p<.05$ , italics  $p<.1$ . Region (ref.=GK), Gender (ref.=F).

#### 3.2. Executive Function Tasks: Response Time (RT) and Accuracy

Table 1 summarizes the descriptive statistics of children's performance of digit 1-back and DCCS tasks. While we employed both RTs (i.e., log-transformed time difference between target condition and control condition) and accuracy counts to represent cognitive ability, the two measures were not always correlated. RTs of digit 1-back were negatively correlated with accuracy counts [ $r(39)=-0.56$ ,  $p<.0005$ ], suggesting that children with higher accuracy scores responded faster than others. However, RTs and accuracy counts from DCCS were not meaningfully correlated

[ $r(39)=0.24, p=.11$ ].

**Table 2.** Descriptive statistics (mean and SD) of digit 1-back and DCCS

		Mean (SD)	Mean (SD)
DCCS		RT	Accuracy
Gyeongsang	Female	0.103 (0.204)	9.727 (1.902)
	Male	0.099 (0.179)	8.428 (1.718)
Seoul	Female	0.151 (0.188)	9.384 (1.445)
	Male	0.031 (0.296)	8.800 (2.201)
Digit 1-back		RT	Accuracy
Gyeongsang	Female	0.021 (0.127)	6.363 (2.419)
	Male	0.053 (0.180)	6.285 (3.251)
Seoul	Female	-0.007 (0.089)	6.461 (2.503)
	Male	0.063 (0.173)	5.700 (2.584)

### 3.3. Relationship between EF Scores and Acoustic Cues

Based on the results of no significant inter-group differences in production, we performed correlation tests over all children ignoring dialect and gender. Table 2 presents a summary (i.e., correlation coefficients and  $p$ -values) of the partial correlation tests over the 24 combinations of EF scores and acoustic variables. After a Bonferroni correction (a conservative adjustment of a significance level for the hypothesis test to countercheck multiple comparisons,  $0.05/24=0.002$ ), there was only one variable pair whose correlation coefficient was marginally significant: in the lenis-aspirated model, DCCS accuracy was positively correlated with F0 coefficients when the VOT coefficient was controlled [ $r(41)=.46, p=.002$ ]. This means that children with better mental flexibility utilized F0 more than others in producing the lenis-aspirated stop contrast. Further analyses over subject subsets also revealed the same relationship: female children [ $r(24)=.56, p=.005$ ] and Gyeongsang children [ $r(24)=.69, p=.002$ ] with greater DCCS accuracy used F0 more than those with lower DCCS accuracy in realizing the lenis-aspirated contrast.

Unlike the findings from Kong & Yoo (2017) exploring children's perception, the production data in the present study did not yield any robust relationship between EF scores and acoustic variables for the fortis-aspirated stop model. Similarly, acoustic coefficients from the fortis-lenis stop model were not significantly associated with EF scores, either. Although greater VOT coefficients used for the fortis-lenis stops were weakly correlated with slower RTs (i.e., poor EF capacity), the partial correlation coefficients between VOT and the RT of digit 1-back and DCCS did not reach the conservatively adjusted level of significance (digit 1-back RT:  $r=.375, p<.01$ , DCCS RT:  $r=.335, p<.05$ ; Table 3).

**Table 3.** Summary of partial correlation tests (estimates and  $p$ -values) between EF tasks (digit 1-back and DCCS) and acoustic coefficients (VOT and F0)

Partial correlation	Estimate ( $p$ -value)	Estimate ( $p$ -value)
<b>1-Back and VOT with F0 controlled.</b>		
	RT	Accuracy
Fortis-Lenis	<b>0.375 (0.01)</b>	-0.167 (0.30)
Fortis-Asp.	-0.026 (0.87)	-0.033 (0.83)
Lenis-Asp.	-0.069 (0.66)	0.079 (0.62)
<b>1-Back and F0 with VOT controlled.</b>		
	RT	Accuracy
Fortis-Lenis	-0.116 (0.47)	-0.174 (0.28)
Fortis-Asp.	0.122 (0.45)	-0.037 (0.81)
Lenis-Asp.	0.075 (0.64)	0.147 (0.36)
<b>DCCS and VOT with F0 controlled.</b>		
	RT	Accuracy
Fortis-Lenis	<b>0.335 (0.03)</b>	-0.121 (0.45)
Fortis-Asp.	0.021 (0.89)	-0.059 (0.71)
Lenis-Asp.	-0.104 (0.52)	0.054 (0.73)
<b>DCCS and F0 with VOT controlled.</b>		
	RT	Accuracy
Fortis-Lenis	-0.196 (0.22)	-0.153 (0.34)
Fortis-Asp.	0.019 (0.90)	-0.092 (0.57)
Lenis-Asp.	0.132 (0.41)	<b>0.468 (0.002)</b>

Bold indicates  $p<.05$ .

## 4. Discussion & Conclusion

The present study explored the variable relationship between children's Korean stop production and their executive function task performance to understand whether and how one's domain-general cognitive ability accounts for individual variability in the use of multiple cues in speech production. As Kong & Yoo (2017) demonstrated that children's limited cognitive resources could restrict their use of redundant cues in speech perception, we were interested in finding matching evidence in speech production. Specifically looking into the Korean stop productions, however, the present study failed to find a robust correlation between EF scores and the use of primary or secondary acoustic cues to the stop contrast. Even for the fortis-aspirated stop pair where F0 is undoubtedly redundant acoustic information, there was no consistent association between greater EF scores and less use of F0. On the one hand, this may indicate that domain-general cognitive ability does not necessarily moderate speech production and perception in the same way, exhibiting its loose association with the production mode of speech. On the other hand, these null results may simply denote a methodological problem of the current study in pursuing the research question.

Indeed, the absence of a consistent relationship between EF and F0 use for the fortis-aspirated stop production may be attributable to the task employed in the current study, i.e., a self-paced word repetition task. This repetition task might have been too easy for the elementary school children to exhaust cognitive resources in order to successfully complete the task. When prior studies successfully showed a robust association between EF and speech production performance, they carried out a verbal fluency task with time pressure (e.g., Amunts et al., 2020; Filippi et al., 2022) or a word-naming task with younger children and speakers with speech disorder (e.g., Eaton & Ratner, 2016). The relatively easy task used in the present study might have provided little room for cognitive resources to act as a mediating factor in children's utilizing multiple

acoustic cues for the stop contrast. Admittedly, future research needs to be carefully designed so that one's cognitive control can be fully engaged in speech production task.

With this limitation being said, we may interpret children's F0 use for the lenis-aspirated stop production, which was marginally correlated with EF, as the speakers' (long-term rather than spontaneous) adaptation to the innovative phonetic form of the target sounds in the context of sound change. Although marginal in the statistical tests, child listeners who used innovative acoustic information, F0, for the lenis-aspirated stops more than others had better mental flexibility, the sub-component of EF capacity that DCCS measures. This tendency suggests that children with better executive function are more sensitive to subtle acoustic variation under the sound change, and may actively employ new variants in their stop production. Despite the aforementioned methodological limitation as well as a small sample size, this at least suggests the possibility that individuals' cognitive ability plays a role in spreading and stabilizing the sound change.

To conclude, we examined whether and how domain-general cognitive ability correlates with elementary school children's production of the Korean stop contrast, and compared this with its role in children's use of multiple cues in speech perception. After conducting a word-repetition task and two executive function tasks (digit 1-back and DCCS), we did not successfully find consistent correlations between working memory or mental flexibility (two components of executive function) and children's use of redundant acoustic cues for the fortis-aspirated stop productions. This is a result incompatible with the prior finding from speech perception (Kong & Yoo, 2017). This may suggest that the two modes of speech (production and perception) interact with cognitive ability in different ways. However, we may also need richer experimental evidence to argue for it, acknowledging the methodological limitations of the study.

## Acknowledgements

The current study explored a subset of participants from the larger speech production database presented in Holliday et al. (2023). We report original data on executive function task scores and statistical analyses examining their relationship with speech production.

## References

Amunts, J., Camilleri, J. A., Eickhoff, S. B., Heim, S., & Weis, S. (2020). Executive functions predict verbal fluency scores in healthy participants. *Scientific Reports*, *10*, 11141.

Boersma, P., & Weenink, D. (2023). Praat: Doing phonetics by computer (version 6.3.14) [Computer program]. Retrieved from <http://www.praat.org>

Chandrasekaran, B., Sampath, P. D., & Wong, P. C. M. (2010). Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, *128*(1), 456-465.

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

Clayards, M. (2018). Differences in cue weights for speech perception are correlated for individuals within and across contrasts. *The*

*Journal of the Acoustical Society of America*, *144*(3), EL172-EL177.

Eaton, C. T., & Ratner, N. B. (2016). An exploration of the role of executive functions in preschoolers' phonological development. *Clinical Linguistics & Phonetics*, *30*(9), 679-695.

Filippi, R., Ceccolini, A., & Bright, P. (2022). Trajectories of verbal fluency and executive functions in multilingual and monolingual children and adults: A cross-sectional study. *Quarterly Journal of Experimental Psychology*, *75*(1), 130-147.

Francis, A. L., & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics*, *71*, 1360-1374.

Han, M. S., & Weitzman, R. S. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/. *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*(3), 263-272.

Holliday, J. J., Kong, E. J., & Lee, H. (2023, August). The effects and interaction of age and region on the acoustic realization of the Korean stop contrast. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*. Prague, Czech.

Idemaru, K., Holt, L. L., & Seltman, H. (2012). Individual differences in cue weights are stable across time: The case of Japanese stop lengths. *The Journal of the Acoustical Society of America*, *132*(6), 3950-3964.

Kang, K. H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, *124*(6), 3909-3917.

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.

Kim, M., & Stoel-Gammon, C. (2009). The acquisition of Korean word-initial stops. *The Journal of the Acoustical Society of America*, *125*(6), 3950-3961.

Kim, M. R. (2000). *Segmental and tonal interactions in English and Korean: A phonetic and phonological study* (Doctoral dissertation). University of Michigan, Ann Arbor, MI.

Kim, S. (2015). Ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, *22*(6), 665-674.

Kim, S. H. (1999). Multiple acoustic cues of three-way phonemic contrast in stop consonants. *Studies in Phonetics, Phonology and Morphology*, *5*(1), 79-103.

Kong, E. J., Beckman, M. E., & Edwards, J. (2011). Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics*, *39*(2), 196-211.

Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics*, *59*, 40-57.

Kong, E. J., & Yoo, J. (2017). Speech processing strategy and executive function: Korean children's stop perception. *Phonetics and Speech Sciences*, *9*(3), 57-65.

Lee, H. (2020). Cross-generational acoustic comparisons of tonal Kyungsang Korean stops. *The Journal of the Acoustical Society of America*, *148*(2), EL172-EL178.

Lee, H., Holliday, J. J., & Kong, E. J. (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, *14*(7), e12374.

- Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, 42(2), 145-169.
- Lee, H., & Kong, E. J. (2016). Relationship between individual traits and variability in processing Korean stops: Implication on mechanism of sound change. *Language & Linguistics*, 71, 215-236.
- 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.
- Netelenbos, N., Gibb, R. L., Li, F., & Gonzalez, C. L. R. (2018). Articulation speaks to executive function: An investigation in 4- to 6-year-olds. *Frontiers in Psychology*, 9, 172.
- Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, 39(1), 59-67.
- Ou, J., Xiang, M., & Yu, A. C. L. (2023). Individual variability in subcortical neural encoding shapes phonetic cue weighting. *Scientific Reports*, 13(1), 9991.
- Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1), 46-59.
- Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication*, 13(1-2), 109-125.
- Posit Team. (2023). RStudio IDE. Retrieved from <https://posit.co/downloads/>
- Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23(2), 287-308.
- Son, G. (2018). Production and perception of F0 by three-year-old children. *The Journal of Studies in Language*, 34(3), 505-518.
- Wright, J. D. (2007). *Laryngeal contrast in Seoul Korean* (Doctoral dissertation). University of Pennsylvania, Philadelphia, PA.
- Yu, A. C. L. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PLOS ONE*, 5(8), e11950.
- Yu, A. C. L. (2022). Perceptual cue weighting is influenced by the listener's gender and subjective evaluations of the speaker: The case of English stop voicing. *Frontiers in Psychology*, 13, 840291.
- Yu, A. C. L. (2023). The actuation problem. *Annual Review of Linguistics*, 9, 215-231.
- Yu, A. C. L., Grove, J., Martinovic, M., & Sonderegger, M. (2011, August). Effects of working memory capacity and "autistic" traits on phonotactic effects in speech perception. *Proceedings of the 17th International Congress of the Phonetic Sciences* (pp. 2236-2239). Hong Kong, China.
- Yu, A. C. L., & Lee, H. (2014). The stability of perceptual compensation for coarticulation within and across individuals: A cross-validation study. *The Journal of the Acoustical Society of America*, 136(1), 382-388.
- Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., ... Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, 68(3), i-151.
- **Eun Jong Kong**  
 Professor, School of Liberal Arts and Sciences  
 Korea Aerospace University  
 76 Hanggongdaehang-ro, Deogyang-gu, Goyang 10540, Korea  
 Tel: +82-2-300-0338  
 Email: ekong@kau.ac.kr  
 Fields of interest: Speech production & perception, Laboratory phonology, Language acquisition
- **Hyunjung Lee**, Corresponding author  
 Associate Professor, Dept. of English Language Education  
 Incheon National University  
 12 Gaebeol-ro, Songdo-dong, Yeonsu-gu, Incheon 21999, Korea  
 Tel: +82-32-835-8992  
 Email: hjlee123@inu.ac.kr  
 Fields of interest: Speech production & perception, Laboratory phonology, Language acquisition
- **Jeffrey J. Holliday**  
 Assistant Professor, Dept. of Linguistics  
 University of Kansas  
 1541 Lilac Lane, Lawrence, KS 66045, USA  
 Tel: +1-785-864-6021  
 Email: jjh@ku.edu  
 Fields of interest: Phonetics, Second language acquisition, Sociolinguistics

## Appendix 1.

The reading list used in the production task

Orthography	Transcription	Consonant	Vowel	Accent type
그림	ki.lim	k	i	HH
까기	k'a.ki	k'	a	HH
칼이	k <sup>h</sup> al.i	k <sup>h</sup>	a	HH
발이	pal.i	p	a	HH
빨다	p'al.ta	p'	a	HH
판이	p <sup>h</sup> an.i	p <sup>h</sup>	a	HH
달이	tal.i	t	a	HH
딸이	t'al.i	t'	a	HH
틈이	t <sup>h</sup> im.i	t <sup>h</sup>	i	HH
거미	kΛ.mi	k	a	HL
깡이	k'anj.i	k'	a	HL
칸이	k <sup>h</sup> an.i	kh	a	HL
방이	paŋ.i	p	a	HL
뺑이	p'anj.i	p'	a	HL
팔이	p <sup>h</sup> al.i	p <sup>h</sup>	a	HL
답이	tak.i	t	a	HL
따귀	t'a.kwi	t'	a	HL
탐이	t <sup>h</sup> am.i	t <sup>h</sup>	a	HL