

# Relationship between Working Memory and Identification of a Few Native Phonetic Contrasts

Usha Shastri<sup>a</sup>, Keerthana Kulath Purath Raj<sup>b</sup>, Mable Mathew<sup>b</sup>, Mohan Kumar Kalaiah<sup>a</sup>, Ajith Kumar Uppunda<sup>c</sup>

<sup>a</sup>Department of Audiology and Speech Language Pathology, Kasturba Medical College, Mangalore; and Manipal Academy of Higher Education, Manipal, Karnataka, India

<sup>b</sup>Nitte Institute of Speech and Hearing, Mangalore, Karnataka, India

<sup>c</sup>Department of Audiology, All India Institute of Speech and Hearing, Mysuru, Karnataka, India

**Correspondence:** Mohan Kumar Kalaiah, PhD  
Department of Audiology and Speech-Language Pathology, Kasturba Medical College, Light House Hill Road, Mangalore 575001, India  
Tel +91-8884-045-381  
Fax +91-824-2428379  
E-mail: mohan.kumark@manipal.edu

Received: August 24, 2018  
Revised: February 23, 2019  
Accepted: February 23, 2019

**Objectives:** Large individual variability is documented for identification performance of native phones, especially in challenging situations. It is not known whether the ability to utilize cues available for phone identification is facilitated by cognitive abilities, thereby explaining a proportion of the individual variability. This study investigated the relationship between working memory capacity and identification of a few Malayalam phones in the absence of contextual cues among native listeners. **Methods:** Forty native listeners of Malayalam, aged between 18 and 25, participated in this study. Participants identified 8 Malayalam phones embedded in nonsense words. Working memory capacity was measured using tasks such as reading span, operation span, digit forward span, and digit backward span. Identification score for each phone, total phone identification score (average identification score from 8 phones), and reaction time during identification were obtained. **Results:** Phone identification score of participants ranged from 57.8% to 99%. Pearson product moment correlation analysis showed a significant positive correlation between all measures of working memory capacity and total phone identification score, indicating that working memory capacity play a role in the identification of phones. Reaction time showed a significant negative correlation with digit backward span and operation span. The measures of working memory capacity accounted for 24.7% of the variability in phone identification score. **Conclusion:** Identification of phones in the absence of contextual cues increases the cognitive load. Therefore, higher working memory capacity might aid in native phone identification in difficult situations. This study reveals the top down influence of cognition on native speech perception.

**Keywords:** Native speech perception, Working memory capacity, Top down influence, Native phone identification, Cognition, Individual variability

Identification of phonemes present in the native language has been considered to be near perfect and effortless. This is true especially when redundant linguistic cues are present in stimuli, such as words and sentences. However, in the absence of redundant cues such as in the context of nonsense words and nonsense syllables, identification or discrimination of native phones has been found to be difficult (Kalaiah, Thomas, Bhat, & Ranjan, 2016; Kong & Edwards, 2011; López-Zamora, Luque, Álvarez, & Cobos, 2012;

Shastri, Mythri, & Kumar, 2014). Evidence for less than perfect identification or discrimination performance among native listeners is available from many investigations of cross language speech perception (Broersma, 2010; Chang & Mishler, 2012; Tsao, Liu, & Kuhl, 2006). Studies of cross language speech perception have compared native and non-native listeners' abilities for the perception of various contrasts. For example, use of preceding vowel duration as a perceptual cue for final fricative voicing (Broersma, 2010), per-

ception of unreleased stops (Chang & Mishler, 2012) in English, and alveo-palatal affricate vs. fricative contrast in Mandarin Chinese (Tsao et al., 2006). In these investigations, identification/discrimination scores of native listeners for a variety of stimuli have been reported to range from 32% to 92.5%. Thus, it appears that less than perfect identification score is universal to listeners of any native language.

In addition to less than perfect identification or discrimination scores, native listeners also show large amount of individual variability in identification or discrimination performance, especially when the stimuli do not contain contextual cues (Kong & Edwards, 2011; López-Zamora et al., 2012; Shastri et al., 2014). A few studies have explored the factors contributing to individual variability in phoneme identification scores among native listeners. Physiology of the auditory system and cognitive ability in terms of utilization of cues available for perception are thought to contribute to individual variability (Makashay, 2003). To study contribution of the physiology of the auditory system towards individual variability, our earlier investigation (Shastri et al., 2014) studied the top down influence of the descending auditory pathway on individual variability for the identification of native phonetic contrasts without contextual cues. Top down influence of the descending auditory pathway was studied by measuring contralateral inhibition of transient evoked otoacoustic emissions. Relationship between the amount of contralateral inhibition of transient evoked otoacoustic emissions and identification of a few native Malayalam phones (/l/, /ɭ/, /r/, /ɻ/, /ʈ/, /t/, /ɳ/, and /n/) in nonsense words was noted. The amount of contralateral inhibition of transient evoked otoacoustic emissions explained about 30% of the variance in phone identification. Hence, it was suggested that the top down connections mediated via the medial olivocochlear bundle may influence an individual's ability to identify acoustically similar native phones.

In cognitive ability, working memory capacity is an important individual-differences variable and is known to account for a significant portion of a variance in general intellectual ability (Conway, Kane, & Engle, 2003; Kane et al., 2004). A positive relationship exists between working memory capacity and speech recognition in noise in individuals with normal hearing sensitivity (Gordon-Salant & Cole, 2016). Speech comprehension in noise appears to depend on working memory and executive-control processes (Heald

& Nusbaum, 2014). The influence of working memory on speech perception could be because of the individual differences in the ability to use contextual information effectively (Davis, Ford, Kherif, & Johnsruide, 2011; Janse & Jesse, 2014). However, the influence of working memory capacity on individual variability in the identification of native phones in quiet situations, in the absence of contextual cues is not known. Executive functions allow listeners to direct attention to integrate the acoustic signal with previous knowledge and inhibit irrelevant information (Tamati, Gilbert, & Pisoni, 2013; Woods, Kalluri, Pentony, & Nooraei, 2013). Hence, it can be postulated that working memory capacity might positively relate with the ability to identify difficult native phone contrasts. This knowledge is crucial in understanding the speech and language processing in native listeners, especially in difficult listening situations. In this context, the purpose of this study was to investigate the relationship between working memory capacity (reading span, operation span, digit forward span, and digit backward span) and identification of a few native phonetic contrasts in the absence of contextual cues by native listeners of Malayalam. Among four measures of working memory, digit forward and backward span were auditory based, whereas reading and operation span were visual based tasks.

Malayalam was chosen as it contains rich phonetic inventory and has fine phonetic contrasts that can be challenging to identify even for native listeners (Kumari, 1972; Kumar, Hegde, & Mayaleela, 2010; Shastri et al., 2014), especially when contextual cues are not available (Shastri et al., 2014). The present study used four pairs of Malayalam phones: /l-/ɭ/, /r-/ɻ/, /ʈ-/t/, and /ɳ-/n/ which are the same phones as were used in our previous study (Shastri et al., 2014), which showed that these phones are difficult to identify within pairs. Following are the phonetic descriptions of the phones: /l/-alveolar lateral approximant; /ɭ/-retroflex lateral approximant; /r/-alveolar trill; /ɻ/-retroflex approximant; /ʈ/-unvoiced retroflex stop; /t/-unvoiced alveolar stop; /ɳ/-dental nasal; /n/-alveolar nasal (Kumari, 1972). Among the four phone pairs, /l-/ɭ/ and /r-/ɻ/ occur in singleton syllable onsets in the medial position and contrast meaning and hence, are phonemic pairs in Malayalam (Asher & Kumari, 1997; Kumari, 1972; Mohanan & Mohanan, 1984). The /ʈ-/t/ pair contrast the meaning as geminates in the word medial position. Apart from the medial position, the phone pair /ʈ-/t/ does

not contrast in meaning in any other word positions or in singleton. Thus, this pair can be considered as a phonemic pair only in the word medial position. The phone pair / $\eta$ /-/n/ contrasts in meaning when it occur as a geminate in the word medial position (Asher & Kumari, 1997; Kumari, 1972; Mohanan & Mohanan, 1984). The occurrence of phones / $\eta$ / and /n/ are complementary, such that / $\eta$ / occurs only in word initial position, whereas, /n/ occurs in the word medial and final position. In this respect, the / $\eta$ /-/n/ pair can be considered as an allophonic variation in the initial position. Furthermore, we chose nonsense syllables for an identification task similar to that used in our earlier study (Shastri et al., 2014). These phone pairs in the absence of contextual cues are more challenging to identify, thereby we expected that our stimuli might engage more cognitive resources during the identification task.

## METHODS

### Participants

Forty native listeners of Malayalam, aged between 18 and 25, with normal hearing sensitivity participated in the study. The hearing threshold at octave frequencies between 250 Hz and 8,000 Hz in air conduction mode and at octave frequencies from 250 Hz to 4,000 Hz in bone conduction mode was less than or equal to 15 dB HL in both ears. The mean pure tone average was 4.94 ( $\pm$  3.28) dB HL in the right ear and 4.19 ( $\pm$  3.15) dB HL in the left. None of the participants reported of any history of otological problems. All the tests were done in a quiet room free from any distractions. The study was approved by the scientific committee of the institute and ethical clearance was obtained (NISH/SCICOM/2016-17/08). The purpose and nature of the study was explained to participants and informed consent was obtained.

### Malayalam Phone Contrast Identification Test (MPCIT)

As mentioned earlier, stimuli included four pairs of Malayalam phones, /l/-/l/, /r/-/r/, /t/-/t/, and / $\eta$ /-/n/ which are considered to be challenging for native listeners in the absence of linguistic cues (Shastri et al., 2014). These 8 phones were embedded in nonsense disyllabic words of the form  $C_1V_1C_2V_2$ . Initial consonant ( $C_1$ ) always contained the target phone (one of the /l/, /l/, /r/, /r/, /t/, /t/, / $\eta$ /, or /n/). The second consonant ( $C_2$ ) was one among the randomly

chosen phones /k/, /d/, /s/, and /p/. First vowel ( $V_1$ ) was either of /a/, /i/ or /u/ and second vowel ( $V_2$ ) was /a/. MPCIT contained 10 nonsense words for each of the 8 target phones with a combination of different  $V_1$ , and  $C_2$ . In total there were 320 nonsense words from four talkers, two males and two females (4 talkers  $\times$  8 phones  $\times$  10 nonsense words). Nonsense words were preferred in MPCIT over meaningful words as they do not provide any additional higher linguistic cues (such as semantic or contextual cues) that might aid in the identification of the phone, thus making the task challenging. Recorded material developed by Shastri (2015) was used. Stimulus presentation and measurement of reaction time was done using DMDX software (Forster & Forster, 2003).

A phone-to-picture association paradigm was used for the identification task which is similar to that used by other researchers (Chandrasekaran, Sampath, & Wong, 2010; Kumar et al., 2010; Shastri et al., 2014; Wong & Perrachione, 2007). In this, each of the 8 phones selected for the study was associated with a picture, which was arbitrary. For example, /la/ was associated with the picture of a butterfly, /la/ was associated with the picture of a chair, and so on. During phone-to-picture association learning, participants heard a recorded CV syllable and the picture to be associated with that phone (for example, when /la/ was heard, picture of butterfly was shown on the screen). This phone-to-picture association was repeated as many number of times the participant asked for it. Following this, to ensure that participants had learned the phone-to-picture association successfully, a practice trial of 16 items was given. During this, participants saw three pictures on the screen and simultaneously heard one nonsense word. The task was to identify the picture associated with the target phone in the nonsense syllable heard. If a participant attained 90% accuracy in the practice trial, phone-to-picture association learning was considered successful. Every participant of this study could successfully learn the phone-to-picture association and it took around 15 minutes to learn the association and complete the practice trial. MPCIT was administered after ensuring successful learning of phone-to-picture association. During testing, the participant listened to nonsense words binaurally through Sennheiser HD-100 headphones at comfortable listening level. Three pictures were presented simultaneously on the screen of a laptop. The task of the participant was to identify the picture associated with the target phone by pressing

the key corresponding to correct picture on the keyboard. Performance in MPICT reflected the identification of target phones and one point was given for each correct response. The MPCIT took approximately 30 minutes for each participant.

### Working Memory Measures

Stimuli developed by Shastri (2015) were used to measure operation span, reading span and auditory digit span in Malayalam. The order of administering these working memory measures and MPCIT was randomized among the participants. Each participant was able to complete all working memory measures in approximately 30 minutes.

#### Auditory digit span

Stimuli included recorded disyllabic digits from one to nine of equal syllable length. The test included three lists, each having six levels. Level 1, which contained three digits was the easiest and level 6, which contained eight digits was the toughest. The inter-digit-interval was 250 ms, and the order of the digits was randomized to eliminate familiarity effects (Rönnberg, 1990). In the forward digit span, clusters of digits were presented binaurally through headphones and the participants were asked to repeat the digits they heard in the same order. In the backward digit span, the participants were asked to repeat numbers in the reverse order. The highest level at which digits were repeated correctly in the desired order at least two out of three runs was noted. The number of digits at that level was considered as the digit span of the participant.

#### Operation span task

Participant's ability to remember the target stimuli which was interleaved with a secondary processing task (to verify a mathematical problem) was evaluated. Each element consisted of a mathematical operation and a word to be remembered in Malayalam—e.g.,  $(3 \times 5) - 4 = 4$ , yes or no? Apple. Combination of a number of elements was defined as a trial and its length varied from two to five. Three trials of each length (two, three, four, and five) were presented, for a total of 12 trials (4 lengths  $\times$  3 trials).

The test procedure was similar to one used by Kane et al. (2004). An element consisting of a mathematical problem was displayed on the computer screen followed by a word to be remembered. The

participant read the mathematical equation aloud, verified whether the given answer was correct, and then read the word. Soon after this, the next element was presented. After all the elements in a trial were presented, participants were asked to repeat words in the trial in correct serial order. Scoring was done according to the guidelines provided by Kane et al. (2004) and Conway et al. (2005). One point was provided for each element recalled in the correct serial order. Further, proportion correct score for each trial was calculated. These proportions were added across all 12 trials to obtain the final score, which is the operation span of the participant.

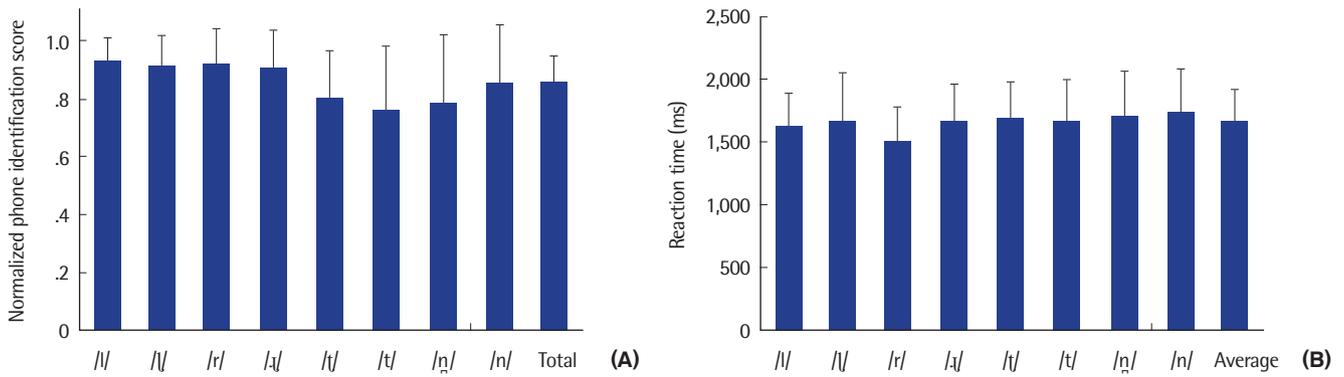
#### Reading span task

Participants' ability to remember the target stimuli, which was interleaved with a secondary processing task (verifying semantic/pragmatic correctness of a sentence) was evaluated. Each element consisted of a sentence and a syllable to be remembered in Malayalam—e.g. "Ramu is going to school. /ka/". Similar to operation span, three trials of each length (two, three, four, and five) were presented for a total of 12 trials (4 lengths  $\times$  3 trials). The test procedure and scoring was similar to that used for operation span task.

## RESULTS

### Performance in Malayalam Phone Contrast Identification Test

A normalized phone identification score was obtained for each phone separately, and by averaging identification score from all 8 phones a total phone identification score was obtained and is shown in Figure 1A. Figure 1A shows that phoneme pairs (/l/-/l/, /r/-/r/) were identified with greater accuracy, followed by allophones (/ɳ/-/n/) and acoustically similar phones (/t/-/t/). Here, it is important to note here that though some phones were easy to identify and some were difficult, native listeners of Malayalam identified all the 8 phones at well above chance levels. One-way repeated measures ANOVA with phones as the within subject factor revealed that there was a significant main effect of phone ( $F_{(7,273)} = 7.81, p < .001$ ). Pair wise comparison using Bonferroni test revealed a significant difference for identification scores between certain phones; results are shown in Table 1. It was seen that acoustically similar phone phones had significantly lower mean identification scores than



**Figure 1.** (A) Normalized phone identification score of 8 phones and total phone identification score along with one standard deviation averaged across the participants. (B) Average reaction time during identification of 8 phones and the overall reaction time averaged across all the phones.

**Table 1.** Result of post hoc pair wise comparisons for identification score across each phone

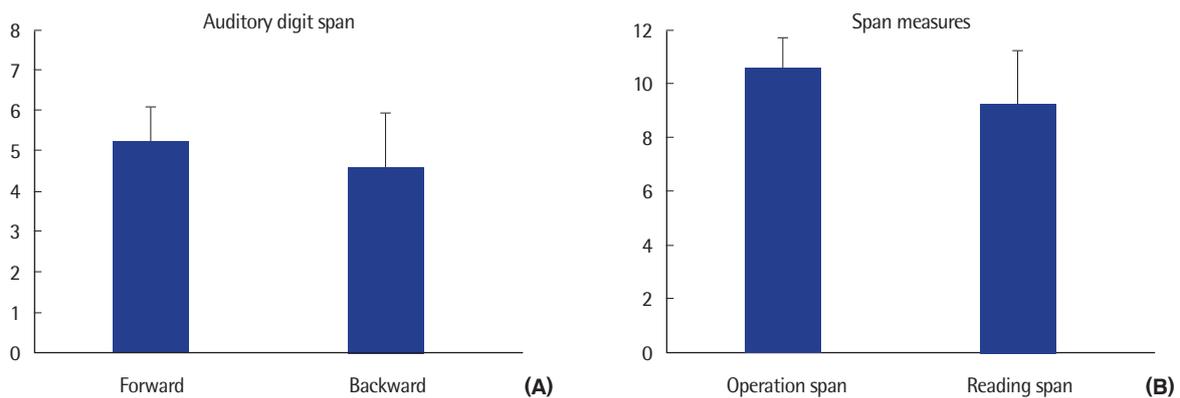
	/l/	/ʌ/	/r/	/ɹ/	/t/	/tʰ/	/n̥/	/n/
/l/								
/ʌ/	NS							
/r/	NS	NS						
/ɹ/	NS	NS	NS					
/t/	S*	S*	S*	S*				
/tʰ/	S*	S*	S*	S*	NS			
/n̥/	NS	NS	NS	NS	NS	NS		
/n/	NS	NS	NS	NS	NS	S*	NS	

S=significant; NS=not significant.  
\* $p < .05$ .

**Table 2.** Result of post hoc pair wise comparisons for reaction time across each phone

	/l/	/ʌ/	/r/	/ɹ/	/t/	/tʰ/	/n̥/	/n/
/l/								
/ʌ/	NS							
/r/	NS	NS						
/ɹ/	NS	NS	S*					
/t/	NS	NS	S*	NS				
/tʰ/	NS	NS	S*	NS	NS			
/n̥/	NS	NS	S*	NS	NS	NS		
/n/	NS	NS	S*	NS	NS	NS	NS	

S=significant; NS=not significant.  
\* $p < .05$ .



**Figure 2.** (A) Mean auditory forward and backward digit span along with one standard deviation. (B) Mean operation span and reading span along with one standard deviation.

phonemic and allophonic pairs. Figure 1A also shows normalized total phone identification scores (last bar), and the scores varied between 57.81% and 99.06% across the participants. Thus, native listeners showed large individual variability in phone identification ability. For the purpose of statistical analysis, phone identification

scores were transformed to arcsine units. Figure 1B shows the mean reaction time for each of the 8 phones and overall reaction time averaged across the phones. Similar to identification scores, reaction time data also revealed slight variations across phones and also individual variability. A one-way repeated measures AN-

OVA with phones as within subject factor revealed a significant main effect of phone on reaction time ( $F_{(7,273)} = 4.73, p < .001$ ). Results of pair wise comparison using Bonferroni test is shown in Table 2, which revealed that the phone /r/ had a significantly shorter reaction time than phones /ɹ/, /l/, /t/, /ŋ/, and /n/. All other reaction times were not significantly different from each other.

### Performance in Working Memory Measures

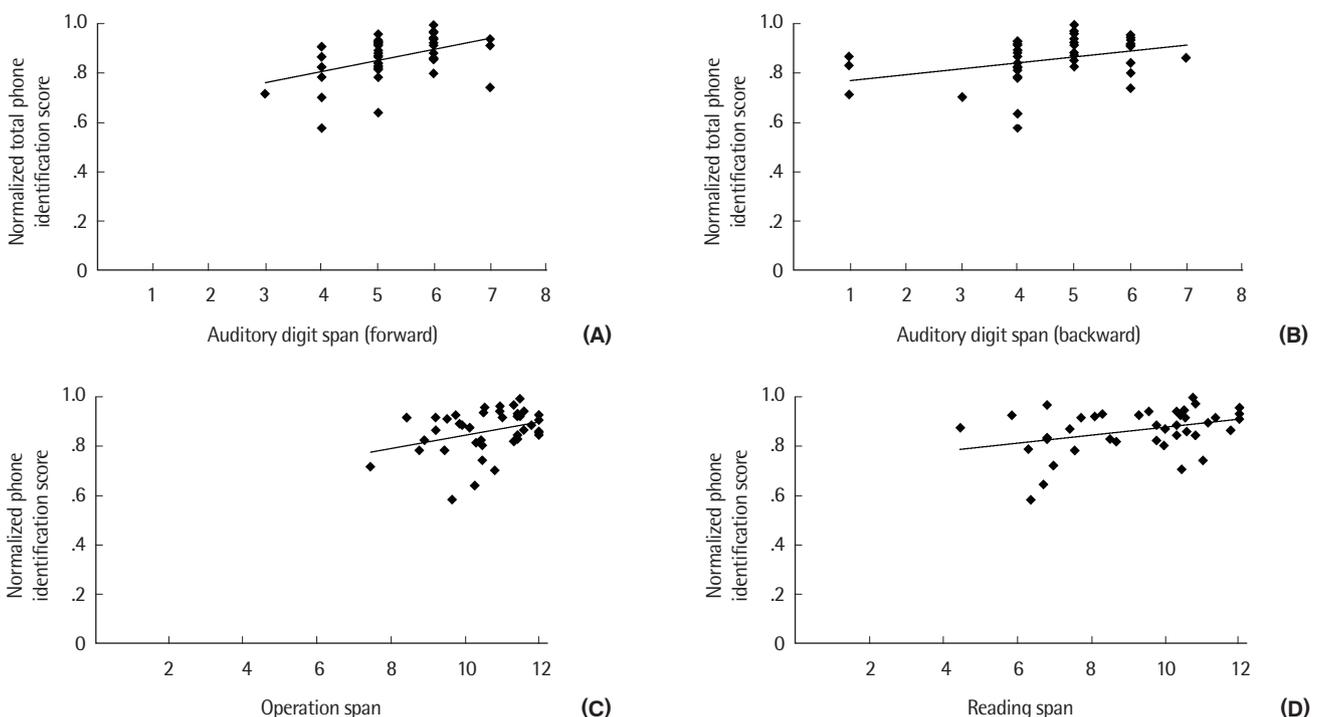
Figure 2A shows the mean auditory digit forward and backward span along with one standard deviation. It shows that the mean score for forward digit span was better than backward digit span, which is expected. Figure 2B shows the mean score for operation span and reading span, where the score for operation span was slightly better compared to the score for reading span. Individual variability was present in scores of measures of working memory capacity.

### Relationship between Phone Identification Score, Reaction Time, and Working Memory Measures

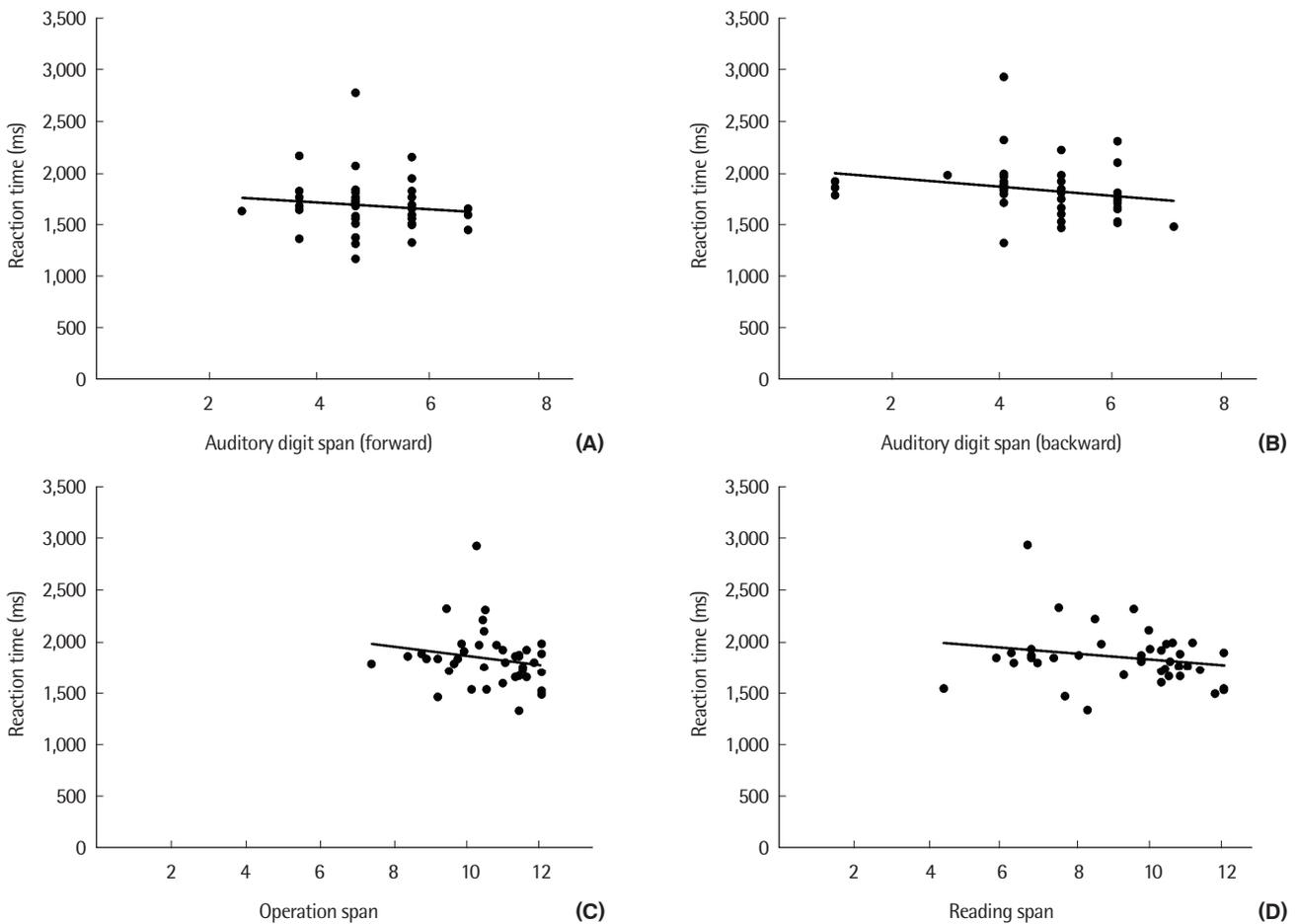
The relationship between different measures of working memo-

ry and total phone identification scores are shown as scatterplots in Figure 3, and the relationship between different working memory measures and overall reaction time are shown in scatterplots in Figure 4. Pearson product-moment correlation analysis was carried out to investigate the relationship between total phone identification, overall reaction time, and working memory capacity. Pearson product-moment correlation coefficient ( $r$ ) is a linear correlation, which is the most commonly used measure of correlation, finds the degree of the association of two sets of variables quantitatively (Paler-Calmorin & Calmorin, 1997). Results are shown in Table 3; which shows slight to moderate (Paler-Calmorin & Calmorin, 1997) significant positive correlation between phone identification score and all measures of working memory capacity. In addition to this, the total phone identification score also showed significant negative correlation with the overall reaction time ( $r = -.398, p < .05$ ). From Table 3, it can also be inferred that overall reaction time had a significant negative correlation with backward digit span and operation span.

To further analyze the effect of working memory capacity on phone identification abilities, regression analysis was carried out.



**Figure 3.** Scatter plot showing the relationship between normalized total phone identification score and working memory measures: (A) forward digit span, (B) backward digit span, (C) operation span, and (D) reading span.



**Figure 4.** Scatter plot showing the relationship between average reaction time and working memory measures: (A) forward digit span, (B) backward digit span, (C) operation span, and (D) reading span.

**Table 3.** Pearson product moment correlation coefficient values ( $r$ ) between total phone identification score, reaction time and working memory measures

	Forward digit span	Backward digit span	Operation span	Reading span
Arcsine transformed total phone identification score	.45*	.33*	.34*	.36*
Average reaction time	-.13	-.45*	-.35*	-.26

\* $p < .05$ .

Before this, all four working memory measures were subjected to factor reduction by conducting Principal Component Analysis (PCA). PCA is one of the most widely used techniques to reduce the dimensionality of the dataset and maximize the variability (Jolliffe & Cadima, 2016). PCA gives new variables that are linear functions of those in the original dataset, that successively maximize variance and that are uncorrelated with each other (Jolliffe & Cadima, 2016). Hence, instead of using four working memory measures as predictor variables in the regression analysis, we used PCA

to see whether four working memory measures can be reduced to lesser number of variables. The result of PCA showed that all four measures could be reduced to one single factor, and this new variable was considered as the predictor variable for linear regression analysis with arcsine transformed total phone identification score as the dependent variable. The results revealed that overall working memory measures used in this study could account for 24.7% of the variance in total phone identification score ( $p < .005$ ).

## DISCUSSION & CONCLUSION

Results of the present study showed that among the four phone pairs tested, some pairs were easier to identify than others. Specifically, phoneme pairs (/l/-l/, /r/-r/) were identified with greater accuracy than allophones (/ɫ/-n/) and acoustically similar phones (/t/-t/). This finding is in consonance with results of our previous investigation (Shastri et al., 2014), which showed that phonemes are identified better than allophones and acoustically similar phones. Further, other studies have also shown that discrimination of allophones is difficult compared to phonemes (Harnsberger, 2001; Whalen, Best, & Irwin, 1997). Results of phone identification score was complimented by the reaction time. The reaction time was shortest for identification of phoneme pairs (/l/-l/, /r/-r/) as compared to allophones (/ɫ/-n/) and acoustically similar phones (/t/-t/). All the phones tested in this study were identified well above the chance level. Thus, even in challenging situations, native phones are identified reliably.

Individual variability was observed for identification scores of native phones, where the identification score ranged from 57.81% to 99.06%. This finding was consistent with literature where several studies have reported a large individual variability for identification score of native phones (Kong & Edwards, 2011; López-Zamora et al., 2012; Shastri et al., 2014). Though native phones are identified reliably in the absence of linguistic cues, not everyone had near perfect performance. The reaction time was also variable across participants. Furthermore, results of the present study showed individual variability on the performance of all measures of working memory capacity. Similar large individual variability on measures of working memory capacity has been reported across life span (Mella, Fagot, Lecerf, & de Ribaupierre, 2015). Working memory measures used in the present study have been widely used in other investigations to capture individual differences in working memory capacity (Daneman & Carpenter, 1983; Daneman & Merikle, 1996; King & Just, 1991; Wilhelm, Hildebrandt, & Oberauer, 2013).

Correlation analysis indicated significant positive correlation between total phone identification score and all measures of working memory capacity. In addition, a significant negative correlation was found between the reaction time and two measures of working memory capacity. These results suggest that individuals

with higher working memory capacity could identify native phones with greater accuracy and shorter reaction time in challenging situations. Thus, results of the present study suggests that working memory capacity also plays a role during identification of native phone contrasts in the absence of contextual cues. Further, the strength of correlation between phone identification score and auditory based digit span measures were similar to that between the phone identification score and visual based operation span and reading span measures. Results of PCA also showed that all four working memory measures could be reduced to a single factor. These findings suggest that both auditory based as well as visual based working memory measures shared similar variability in the data. This view is supported by other researchers who reported that the performance is similar in auditory based and visual based working memory measures in healthy individuals (Huguelet, Zanello, & Nicastro, 2000; Lee, 2018).

Regression analysis revealed that nearly 25% of the variation in the identification score of native phone contrasts could be attributed to working memory capacity. It is possible that cognitive load is increased during identification of native phones in the absence of contextual cues. Thus, larger working memory capacity probably facilitates native speech perception in difficult listening situations. It is said that working memory measures reflect the ability to keep task-relevant information actively maintained in the face of interference (Engle, 2002). In situations where the signal is degraded and interpretation is therefore ambiguous, working memory may contribute to the ability to use sentence context to guide and constrain interpretation to compensate for increased processing demands (Rodd, Davis, & Johnsrude, 2005; Rodd, Johnsrude, & Davis, 2012; Zekveld, Rudner, Johnsrude, Heslenfeld, & Rönnberg, 2012). Thus, individuals with greater working memory capacity may be better able to compensate for degraded listening conditions. Similarly, in challenging listening conditions like identification of native phones in stimuli that does not contain contextual cues, large working memory capacity might aid in better utilization of the available acoustic cues for the identification of phones, thereby leading to better performance. Overall, results of the current study show the importance of top down influence on the perception of native phone contrasts. However, the results of the current study are limited to only correlational analysis and it is known

that correlation need not mean causation.

Existing literature throws light on the influence of working memory capacity on various aspects of speech perception. For example, individual differences in the perception of acoustic cues may vary as a function of working memory (Francis & Nusbaum, 2009). Ou & Law (2017) studied attention and working memory, and showed that individual differences in underlying cognitive factors give rise to individual difference in speech perception. Thus, for accurate phone identification, the way in which acoustic cues are used is important, which is influenced by the higher level cognitive processes (Ou & Law, 2017). The present study advances knowledge by showing the influence of working memory measures on identification of native phone contrasts in challenging situation. Accurate identification of phone contrasts is crucial for understanding speech. Hence, we can imply from the results of the present study that individuals with greater working memory capacity could become successful communicators in real life, difficult situations. Better working memory capacity might also aid to alleviate the deteriorating effects of hearing loss on the identification of phone contrasts.

Another point to be noted is that, though results showed a significant correlation, the strength of correlation was found to be weak or moderate. In addition, the amount of individual variability explained by working memory measures was 24.7%. It can be recollected here that Shastri et al. (2014) observed that contralateral inhibition of otoacoustic emissions explained around 30% of the variability in phone identification score. Phone identification task was identical in both studies. Thus, from the present study and Shastri et al. (2014), it can be suggested that the amount of individual variability in phone identification explained by contralateral inhibition of otoacoustic emissions is greater than that explained by working memory capacity. Hence, it appears that the top down influence of medial olivocochlear bundle in the auditory system, as well as working memory capacity, are important for native phone identification in challenging situations, with greater emphasis on the former.

To conclude, we investigated the relationship between identification of a few native phone contrasts in the Malayalam language with four measures of working memory capacity. Significant positive correlation was found between the identification score of na-

tive phones and working memory capacity. Further, working memory capacity explained a significant proportion of the variance in phone identification performance. These findings suggests that working memory capacity plays a role during perception of native phone contrasts, in adverse listening conditions.

## REFERENCES

- Asher, R. E., & Kumari, T. C. (1997). *Malayalam*. London: Routledge.
- Broersma, M. (2010). Perception of final fricative voicing: native and nonnative listeners' use of vowel duration. *The Journal of the Acoustical Society of America*, 127(3), 1636-1644.
- Chandrasekaran, B., Sampath, P. D., & Wong, P. C. (2010). Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, 128(1), 456-465.
- Chang, C. B., & Mishler, A. (2012). Evidence for language transfer leading to a perceptual advantage for non-native listeners. *The Journal of the Acoustical Society of America*, 132(4), 2700-2710.
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.
- Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547-552.
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 561-584.
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422-433.
- Davis, M. H., Ford, M. A., Kherif, F., & Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. *Journal of Cognitive Neuroscience*, 23(12), 3914-3932.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19-23.
- Forster, K. I., & Forster, J. C. (2003). DMDX: a Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116-124.
- Francis, A. L., & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psycho-*

- physics*, 71(6), 1360-1374.
- Gordon-Salant, S., & Cole, S. S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear and Hearing*, 37(5), 593-602.
- Harnsberger, J. D. (2001). The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: a multidimensional scaling analysis. *Journal of Phonetics*, 29(3), 303-327.
- Heald, S., & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience*, 8, 35.
- Huguelet, P., Zanello, A., & Nicastro, R. (2000). A study of visual and auditory verbal working memory in schizophrenic patients compared to healthy subjects. *European Archives of Psychiatry and Clinical Neuroscience*, 250(2), 79-85.
- Janse, E., & Jesse, A. (2014). Working memory affects older adults' use of context in spoken-word recognition. *The Quarterly Journal of Experimental Psychology*, 67(9), 1842-1862.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Kalaiah, M. K., Thomas, D., Bhat, J. S., & Ranjan, R. (2016). Perception of consonants in speech-shaped noise among young and middle-aged adults. *Journal of International Advanced Otolaryngology*, 12(2), 184-188.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189-217.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language*, 30(5), 580-602.
- Kong, E., & Edwards, J. (2011). Individual differences in speech perception: evidence from visual analogue scaling and eye-tracking. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, Hong Kong, China, 1126-1129.
- Kumar, A. U., Hegde, M., & Mayaleela. (2010). Perceptual learning of non-native speech contrast and functioning of the olivocochlear bundle. *International Journal of Audiology*, 49(7), 488-496.
- Kumari, B. (1972). *Malayalam phonetic reader*. Mysore: Central Institute of Indian Languages.
- Lee, S. J. (2018). The relationship between hearing impairment and cognitive function in middle-aged and older adults: a meta-analysis. *Communication Sciences & Disorders*, 23(2), 378-391.
- López-Zamora, M., Luque, J. L., Álvarez, C. J., & Cobos, P. L. (2012). Individual differences in categorical perception are related to sublexical/phonological processing in reading. *Scientific Studies of Reading*, 16(5), 443-456.
- Makashay, M. J. (2003). *Individual differences in speech and non-speech perception of frequency and duration* (Doctoral dissertation). The Ohio State University, Columbus, OH, USA.
- Mella, N., Fagot, D., Lecerf, T., & De Ribaupierre, A. (2015). Working memory and intraindividual variability in processing speed: a lifespan developmental and individual-differences study. *Memory & Cognition*, 43(3), 340-356.
- Mohanan, K. P., & Mohanan, T. (1984). Lexical phonology of the consonant system in Malayalam. *Linguistic Inquiry*, 15(4), 575-602.
- Ou, J., & Law, S. P. (2017). Cognitive basis of individual differences in speech perception, production and representations: the role of domain general attentional switching. *Attention, Perception, & Psychophysics*, 79(3), 945-963.
- Paler-Calmorin, L., & Calmorin, M. A. (1997). *Statistics in education and the sciences*. Manila: Rex Bookstore Inc.
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, 15(8), 1261-1269.
- Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cerebral Cortex*, 22(8), 1761-1773.
- Rönnerberg, J. (1990). Cognitive and communicative function: The effects of chronological age and "handicap age". *European Journal of Cognitive Psychology*, 2(3), 253-273.
- Shastri, U. (2015). *Auditory perceptual training of non-native speakers: role of auditory and cognitive factors* (Doctoral dissertation). University of Mysore, Mysore, India.
- Shastri, U., Mythri, H. M., & Kumar, U. A. (2014). Descending auditory pathway and identification of phonetic contrast by native listeners. *The Journal of the Acoustical Society of America*, 135(2), 896-905.
- Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: a first report. *Journal of the American Academy of Audiology*, 24(7), 616-634.
- Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: cross-language tests on adults and infants. *The Journal of the Acoustical Society of America*, 120(4), 2285-2294.
- Whalen, D. H., Best, C. T., & Irwin, J. R. (1997). Lexical effects in the percep-

- tion and production of American English /p/ allophones. *Journal of Phonetics*, 25(4), 501-528.
- Wilhelm, O., Hildebrandt, A. H., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, 4, 433.
- Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4), 565-585.
- Woods, W. S., Kalluri, S., Pentony, S., & Nooraei, N. (2013). Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario. *The Journal of the Acoustical Society of America*, 133(6), 4268-4278.
- Zekveld, A. A., Rudner, M., Johnsrude, I. S., Heslenfeld, D. J., & Rönnerberg, J. (2012). Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain and Language*, 122(2), 103-113.

## 국문초록

### 모국어 말소리 대조 시 작업기억과 변별능력 간의 연관성

Usha Shastri<sup>1</sup>, Keerthana Kulath Purath Raj<sup>2</sup>, Mable Mathew<sup>2</sup>, Mohan Kumar Kalaiah<sup>2</sup>, Ajith Kumar Uppunda<sup>3</sup>

<sup>1</sup>Department of Audiology and Speech Language Pathology, Kasturba Medical College, <sup>2</sup>Nitte Institute of Speech and Hearing, <sup>3</sup>Department of Audiology, All India Institute of Speech and Hearing

**배경 및 목적:** 모국어의 말소리(phone)를 변별할 때, 특히 도전적 상황에서 이를 수행할 때에는 매우 큰 개인차가 존재한다. 말소리 변별 시 유용한 단서를 활용할 수 있는 능력이 인지능력에 의해 촉진되는지, 이를 통해 개인차를 어느 정도 설명해 줄 수 있는지에 대해서는 잘 알려져 있지 않다. 본 연구는 말라얄람어(Malayalam)를 모국어로 사용하는 청자가 말라얄람어 말소리를 맥락 단서 없이 구분할 때 작업기억능력과 변별능력 간에 어떤 연관성이 있는지 알아보고자 하였다. **방법:** 말라얄람어를 모국어로 사용하는 18-25세 청자 40명이 본 연구에 참여하였다. 참여자들로 하여금 무의미단어 사이에 삽입된 말라얄람어 8개 말소리를 변별하도록 하였다. 읽기폭 과제, 조작폭 과제, 숫자 바로외우기 과제, 숫자 거꾸로외우기 과제 등을 이용하여 작업기억능력을 측정하였고, 각 말소리의 변별점수, 전체 말소리 변별점수(8개 말소리로부터 얻은 평균변별점수)와 변별 시 반응시간을 함께 측정하였다. **결과:** 참여자의 말소리 변별점수는 57.8%-99%의 범위를 보였다. 피어슨 적률상관분석 결과 모든 작업기억능력 측정치와 전체 말소리 변별점수 간에는 유의한 정적 상관이 나타나 작업기억능력이 말소리 변별에 중요한 역할을 하는 것으로 나타났다. 작업기억능력의 측정치는 말소리 변별점수 다양성의 24.7%를 설명할 수 있었다. **논의 및 결론:** 맥락 단서가 없는 상황에서의 말소리 변별은 인지 부담을 높인다. 그러므로 높은 능력은 어려운 상황에서 모국어 말소리를 변별하는 데 도움을 준다. 이 연구는 모국어 말소리 지각에서 인지가 미치는 하향식 영향력을 제시하고 있다.

**핵심어:** 모국어 말소리 지각, 작업기억능력, 하향식 영향력, 모국어 말소리 변별, 인지, 개인차

## ORCID

Usha Shastri (<http://orcid.org/0000-0003-1172-3549>); Keerthana Kulath Purath Raj (<http://orcid.org/0000-0002-4229-9798>); Mable Mathew (<http://orcid.org/0000-0001-6878-1198>); Mohan Kumar Kalaiah (<http://orcid.org/0000-0001-9984-9175>); Ajith Kumar Uppunda (<http://orcid.org/0000-0002-1368-9834>)