



Effect of Vowel Context on the Recognition of Initial Consonants in Kannada

Mohan Kumar Kalaiah and Jayashree S Bhat

Department of Audiology and Speech Language Pathology, Kasturba Medical College, Manipal University, Mangalore, India

Received May 23, 2017

Revised July 19, 2017

Accepted August 25, 2017

Address for correspondence

Mohan Kumar Kalaiah, MSc

Department of Audiology and

Speech Language Pathology,

Kasturba Medical College,

Manipal University,

Mangalore 575001, India

Tel +91-88840-45381

Fax +91-824-2428379

E-mail mohan.kumark@manipal.edu

Background and Objectives: The present study was carried out to investigate the effect of vowel context on the recognition of Kannada consonants in quiet for young adults. **Subjects and Methods:** A total of 17 young adults with normal hearing in both ears participated in the study. The stimuli included consonant-vowel syllables, spoken by 12 native speakers of Kannada. Consonant recognition task was carried out as a closed-set (fourteen-alternative forced-choice). **Results:** The present study showed an effect of vowel context on the perception of consonants. Maximum consonant recognition score was obtained in the /o/ vowel context, followed by the /a/ and /u/ vowel contexts, and then the /e/ context. Poorest consonant recognition score was obtained in the vowel context /i/. **Conclusions:** Vowel context has an effect on the recognition of Kannada consonants, and the vowel effect was unique for Kannada consonants.

J Audiol Otol 2017;21(3):146-151

KEY WORDS: Consonant recognition · Vowel context · Kannada consonants.

Introduction

Consonants and vowels are smallest segments or units of speech, also known as phonemes. In human communication, speech sounds are not produced individually as isolated phonemes, instead consonants, and vowels are combined to form syllables, words, phrases, and sentences. When speech is produced continuously, the articulatory movements of phonemes overlap or changes the articulation of neighboring phonemes. This phenomenon in which one phoneme affects the production of preceding and upcoming phonemes is referred as coarticulation. The coarticulatory effects observed during speech production affects the acoustic signal of speech, which in turn has an influence on the perception of speech sounds. Several investigations have demonstrated coarticulatory effects of consonant contexts on vowel targets [1-3], vowel contexts on consonant targets [4,5], and vowel contexts on vowel targets [6].

Many investigators have studied the effect of vowel context on the perception of place of articulation of consonants [7-11].

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.

These studies have shown that perception of the place of articulation is dependent on the subsequent vocalic information for stop consonants [7,8], and fricatives [9-11] such as /j/ and /s/. Stop consonants with noise burst centered at 1,600 Hz are perceived as /k/ in the vowel context /a/, and in the vowel context /i/ or /u/ perceived as /p/ [7]. Additionally, studies have also investigated the effect of vocalic context on categorical perception in the selective adaptation paradigm [12,13]. These investigations have found an effect of vowel context on the categorical boundaries for place of articulation and voice onset time. Cooper [12] investigated ability to identify /ba/ vs. /pa/ in the continuum /bi-/pi/ and /ba-/pa/ with and without alternating adaptors, which included /da/ and /ti/. The results showed a shift in the category boundary towards voiceless /pi/ for continuum /bi-/pi/ in the presence of adaptor /ti/, and towards /ba/ for continuum /ba-/pa/ in presence of adaptor /da/.

Consonant recognition in various vowel contexts have been investigated by several investigators [14-18]. These investigations have also found an effect of the vowel context on the identification of consonants, while the patterns of vowel effects are not consistent across studies. This discrepancy in the findings has been attributed to the differences in the phonetic

environment, in the language being investigated, or the consonants being investigated. English consonants are shown to be more accurately identified in the environment of /a/ than /i/ or /u/ [14]. In addition, information transfer analysis has shown that place of articulation information was identified less accurately before /i/ than before other vowels in English but not in other languages [15]. Studies have also investigated the effect of vowel context on the position of consonant in consonant-vowel-consonant (CVC) and consonant-vowel (CV)/ vowel-consonant (VC) syllables and have shown a vowel dependent effects on the identification of consonants [16-18]. In general, results show that initial consonants are identified more accurately than final consonants in the vowel context /a/, while final consonants are identified more accurately after the vowel /i/. Further, in addition to consonant recognition in auditory alone based paradigm, studies have also investigated the effect of vowel context on the perception of stop consonants in audio-visual paradigm or McGurk paradigm [19-22]. McGurk effect is a perceptual phenomenon in which ‘when humans are presented with conflicting visual and auditory speech stimuli (e.g., hear-ba/see-ga) listeners’ report of hearing a completely novel speech sound (e.g., da).’ These studies have found that, vowel context also has an influence on the magnitude of the McGurk effect [20,21]. McGurk effect was found to be largest in the /i/ context, moderate in the /a/ context, and almost nonexistent in the /u/ context [20,22].

From the findings of the above investigations it is apparent that the recognition of consonants is influenced by the vowel context although the extent of the vowel context effect varies across the languages [23-26]. Based on these findings, we believe it is essential to understand the vowel context effects on the recognition of consonants in all the languages. The present study was carried out to investigate the effect of the vowel context on the recognition of Kannada consonants.

Subjects and Methods

Participants

Seventeen young adults (4 males, 13 females) aged between 21 to 28 years (mean=22.6, standard deviation=1.8) participated in the study. All participants had hearing sensitivity within normal limits in both ears, and pure-tone threshold was less than 15 dB HL at octave frequencies from 250 Hz to 8,000 Hz. Immittance evaluation also showed A-type of tympanogram with acoustic reflex thresholds at normal levels in all the participants, suggesting normal middle ear functioning. None of the participants had otologic or neurologic problems, exposure to hazardous noise or ototoxic medication and difficulty understanding speech in noise. The study

was approved by Institutional Ethics Committee (IECKMC-MLR-02-13/29) and informed consent was obtained from all participants prior to their participation in the study.

Stimuli

Isolated CV syllables were used to investigate consonant perception abilities of the participants. The CV syllables composed of 14 consonants (/k/, /g/, /ʃ/, /dʒ/, /t/, /d/, /t/, /d/, /n/, /p/, /b/, /m/, /j/, and /s/) followed by one of the five vowels (/a/, /i/, /u/, /e/, and /o/). Thus, the present study included a total of 70 CV syllables (14 consonant×5 vowels). These syllables were spoken by 12 native speakers of Kannada (6 males and 6 females), and the utterances were recorded using Computerized Speech Lab (Model 4150, Ver.3.2.1., Hoya Co., Tokyo, Japan). The utterances were digitally recorded using 16 bit analog-to-digital converter at a sampling rate of 44,100 Hz. The recorded utterances was reviewed by two audiologists to ensure the intelligibility of the syllables, and when the intelligibility of the syllables was judged to be poor such utterances were replaced with new recordings. The stimulus set included a total of 840 utterances, 70 syllables spoken by 12 talkers. These CV syllables used in the present study formed a subset of syllables used in an earlier investigation [27].

Procedure

Consonant identification task was carried out as a closed set identification task. The participants were instructed to identify the consonants in the CV syllables and respond by clicking the corresponding button labelled with an individual consonant sound, shown on the computer screen. The CV syllables were presented monaurally to right ear of the participants using Sennheiser HD 380 pro circum-aural headphone. Once the response was obtained, next syllable was presented following a short pause of 1.5 seconds. All participants completed the consonant identification task in one session. Responses of all the participants were stored separately in the form of confusion matrix. The syllables were presented in a random order across the speakers, at the most comfortable level of the participants.

Data analysis

Consonant identification score was computed, in percentage, for each participant across the vowel contexts, and the percent correct score was transformed in to rationalized arc-sine units (RAU). The transformed RAU score was subjected to repeated measure analysis of variance (ANOVA) to investigate the effect of talker gender and vowel context on consonant identification. All statistical analysis were performed using Statistical Package for the Social Sciences software ver-

sion 16.0 (SPSS Inc., Chicago, IL, USA). Confusion matrices obtained from all the participants were added separately, across the conditions, to obtain combined confusion matrix. The combined matrices were subjected to information transfer analysis [16], using feature information xfer software. The features of consonants such as place of articulation, manner of articulation, and voicing were used for information transfer analysis. The values for place of articulation was bilabial, alveolar, palatal, dental, retroflex, glottal, and velar; for manner of articulation the values used were stop, affricate, fricative, glide, liquid, and nasal. Voicing had two values, voiced and voiceless.

Results

Overall consonant identification

Fig. 1 shows the average percent correct consonant identification scores for male and female talkers and the combined scores for male and female talkers. The consonant identification score was poorest in the /i/ vowel context compared to /a/, /u/, /e/, or /o/ vowel contexts. Further, the identification score was superior in the /o/ vowel context, followed by the /a/ and /u/ vowel contexts, and then the /e/ vowel context. In addition performance was superior for consonants spoken by female talkers than for consonants spoken by male talkers in all the vowel contexts. To evaluate the data in greater detail, the percent correct score was transformed to RAU and trans-

formed score was subjected to further statistical analysis. A two-way repeated measure ANOVA was performed with vowel context (/a/, /i/, /u/, /e/, and /o/) and talker gender (male, female, and combined) as repeated measures. The ANOVA showed a significant main effect of vowel context [$F(4,64)=15.778, p<0.001$] and talker gender [$F(1.1,17.3)=11.041, p=0.003$], while the interaction between talker gender and vowel context was not significant [$F(3.6,57.2)=1.818, p=0.144$]. Bonferroni multiple comparison post hoc test showed significantly poorer score for consonants presented in the /i/ vowel context ($p<0.05$) compared to consonants in other vowel contexts. Further, identification score in the /e/ vowel context was significantly poorer compared to the /o/ context ($p<0.05$). Overall consonant identification score was significantly better for female talker compared to male talkers ($p=0.05$), while the absolute difference between scores was only 1.9%.

Feature and phoneme analysis

Fig. 2 shows the proportion of transmitted information scores for consonant features, place of articulation, manner of articulation and voicing, as a function of vowel context. Overall the manner of articulation feature was transmitted better than features place of articulation and voicing. While in the vowel context /a/, the manner and voicing features were transmitted better than the place feature. Further, the effect of vowel context was present for all the consonant features. All conso-

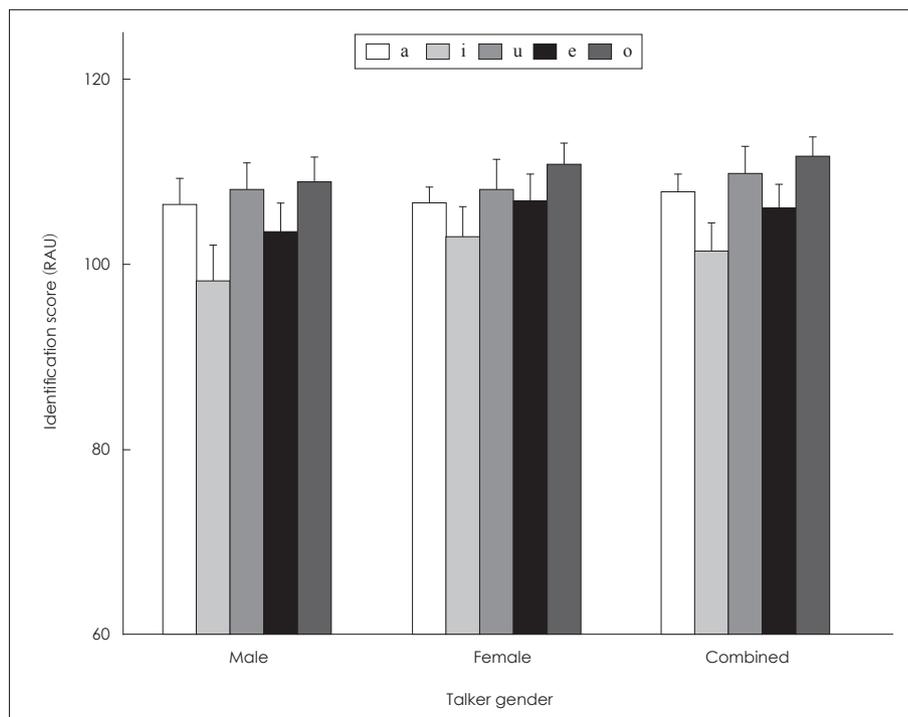


Fig. 1. Mean consonant identification score (in RAU) as a function of vowel context and talker gender. Error bars represent 1 standard deviation of the mean. RAU: rationalized arcsine units.

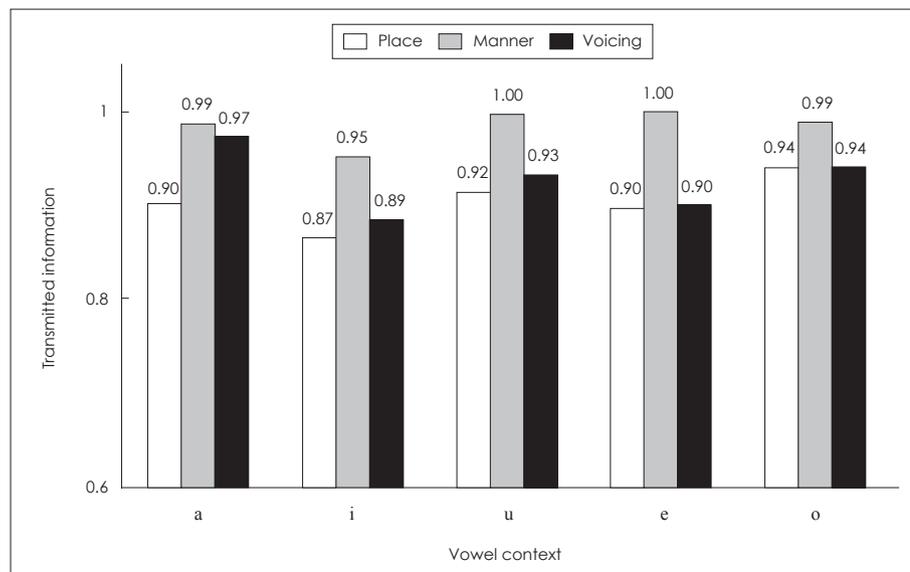


Fig. 2. Proportion of transmitted information scores for consonant features of voicing, place, and manner of articulation as a function of vowel context and talker gender.

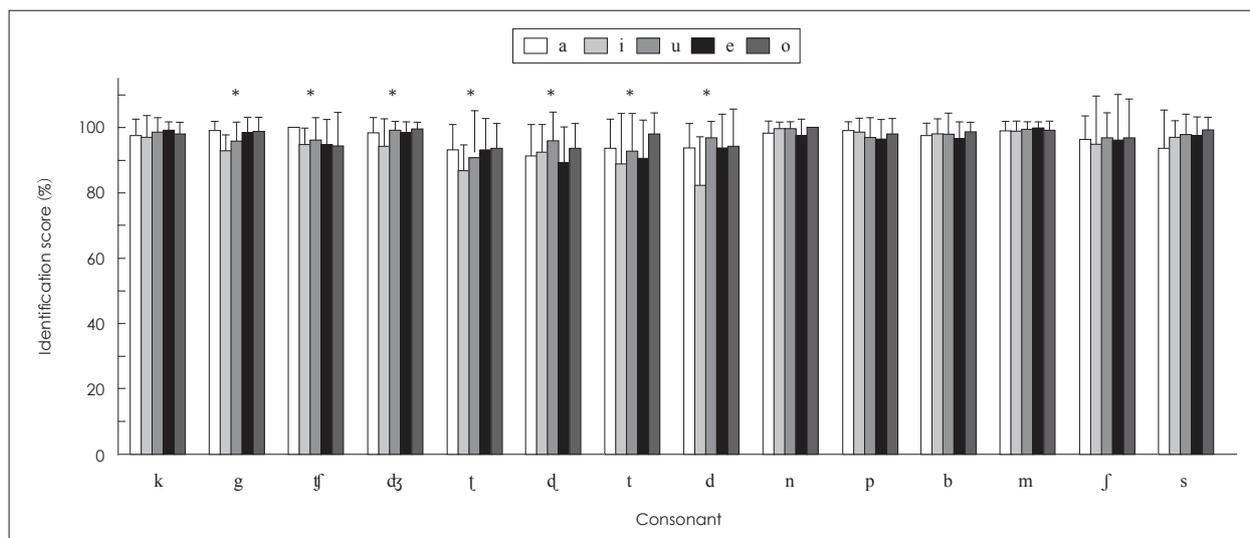


Fig. 3. Mean recognition score of individual consonants (in percent) as a function of vowel context. Error bars represent 1 standard deviation of the mean. *maximum difference in the consonant recognition score across the vowel context is greater than chance level.

nant features had weak effect of the vowel context, and best score was observed in the vowel context /o/ for place of articulation, /u/ and /e/ for manner of articulation, and /a/ for voicing. Further, least score was observed in the vowel context /i/ for all consonant features. Among these consonant features the strongest effect of the vowel context was observed for voicing feature.

To further investigate the vowel context effects on individual consonants, the percent-correct data was calculated for each consonant. Fig. 3 shows mean identification score for each consonant across the vowel context. Among all the consonants, nasal consonants /n/ and /m/, stop consonants /k/, /p/, and /b/, and fricatives /ʃ/ and /s/ were least affected by the vowel context. Maximum difference in the consonant recog-

inition score across the vowel context for these consonants was less than chance level (7%). Consonants /g/, /ʃ/, /dʒ/, and /d/ were slightly affected by the vowel context and the maximum difference in scores was between 7 to 14%. Consonants /t/, /t/, and /d/ were greatly affected with a maximum difference in scores of greater than 21%. Correct recognition score for the consonant /t/ was poorest in the vowel context /i/ and /u/, and it was confused with consonants /p/, /t/, and /d/ in the /i/ vowel context and with /t/ and /d/ in the vowel context /u/. Identification of consonant /t/ was maximally affected in the vowel context /i/, /u/, and /e/. In contrast, recognition of the consonant /d/ was maximally affected in the vowel context /i/ and was confused with consonants /d/ and /t/. Among consonants /g/, /ʃ/, /dʒ/, and /d/ which were slightly affected by the

vowel contexts, consonants /g/ and /dʒ/ were maximally affected in the /i/ vowel context, /d/ was greatly affected in the /e/ vowel context, and /tʃ/ had poor scores in the vowel context /o/. Consonants /g/ and /dʒ/ were confused with each other in the vowel context /i/, the consonant /d/ was confused with /d/ and /b/, and /t/ was confused with /t/ and with /t/ and /d/ in the /i/ and /e/ vowel context respectively. From these findings it is evident that each consonant is affected differently in the context of different vowels.

Discussion

Results of the present study shows an effect of vowel context on the recognition of Kannada consonants. This finding was expected and is in consonance with the findings of several investigations [14-18]. Present study also revealed a significant effect of talker gender on the recognition of consonants, and is consistent with the findings of previous investigations [28,29]. Here, although the mean difference was significantly different statistically the magnitude of difference was not relevant for clinical purpose. Further, in the present study, it was found that the recognition of consonants was significantly poorer in the vowel context /i/ compared to other vowel contexts. Similar finding has been reported by various investigators for English [14-17] and Japanese [15] consonants. While, in contrast to the findings of present investigation, no vowel context effect (/a/ and /i/) has been found for Arabic consonants [15]. Further, a reverse pattern has been reported for Hindi consonants, vowel context /i/ had better identification score compared to the vowel context /a/ [15].

Information transfer for Kannada consonants, in the present study, showed that transfer of information was highest for manner of articulation, and lowest for place of articulation. This finding obtained in the present study is comparable to the results of earlier investigations for Kannada consonants [27], but, contrasts the results of investigations with English consonants [16,17,30,31]. Here, it could be noted that, although the pattern of vowel context effect are similar for overall consonant recognition score, for both Kannada and English consonants, transfer of consonant feature information are different. For English consonants, the transfer of information is largest for consonant feature voicing and lowest for place of articulation. This contrasting finding may be attributed to the different cues available for listeners in Kannada and English consonants for perception. English voiceless consonants are produced with marked aspiration, and it serves as an additional cue for perception of voicing differences for English consonants. Thus, transfer of information for consonant feature voicing might be greater for English consonants

than Kannada consonants.

Perception of nasal consonants /n/ and /m/ were least affected, this finding obtained for Kannada consonants in the present study is comparable to the findings reported for English, Kannada, Hindi, Arabic, and Japanese consonants [14,15,27,30]. Further, the vowel context had least effect or no effect on the recognition of nasal consonants, fricative consonants, and stop consonants (/k/, /p/, and /b/). This finding obtained in the present study for nasal and fricative consonants is in consonance with the earlier investigations [15]. Further, affricates and stop consonants (/g/, /tʃ/, /dʒ/, /t/, and /d/) showed an effect of vowel context. The vowel context effect observed for Kannada consonants was found to be different from English consonants. Studies have shown that perception of English affricates are better in the vowel context /i/ [14], while for Kannada consonants the vowel context /i/ resulted in lowest correct recognition scores. The contrasting findings may be attributed to the differences in the acoustical properties of speech sounds, as a consequence of differences in the articulation of Kannada and English consonants. In addition, the observed difference may also be attributed to differences in phoneme inventory between the languages, which also influences speech perception [32]. Feature analysis revealed that, majority of the perceived confusions or errors were related to the place of articulation, while the perception of manner of articulation and voicing information was accurate. This finding in the present study is in agreement with the findings of other investigations.

Findings of the present study shed light on the impact of vowel context on the perception of Kannada consonants, this information could be significant while developing nonsense syllable based speech tests in Kannada, for assessing speech recognition. Kannada consonants /g/, /tʃ/, /dʒ/, /d/, /t/, /t/, and /d/ had errors above chance level thus should be measured in different vowel contexts. Further, this information would be essential during development of auditory training/rehabilitation programs for listeners with hearing impairment. Auditory training could be designed to improve identification of these more difficult consonants, which could improve comprehension of speech and reduce listening effort in everyday listening conditions. In addition, the finding of the present study would serve as baseline for comparison of consonant recognition in noise and among hearing impaired listeners.

To conclude, the result of the present study showed an effect of vowel context on the recognition of Kannada consonants in quiet listening condition. For overall consonant recognition score, the vowel context effect observed for Kannada consonants was similar to English consonants while it was different compared to Arabic and Hindi consonants. Feature

analysis showed that the transfer of information for Kannada consonants was different compared to English consonants although both had similar vowel effect for consonant recognition score.

Conflicts of interest

The authors have no financial conflicts of interest.

REFERENCES

- 1) Holt LL, Lotto AJ, Kluender KR. Neighboring spectral content influences vowel identification. *J Acoust Soc Am* 2000;108:710-22.
- 2) Lindblom BE, Studdert-Kennedy M. On the role of formant transitions in vowel recognition. *J Acoust Soc Am* 1967;42:830-43.
- 3) Nearey TM. Static, dynamic, and relational properties in vowel perception. *J Acoust Soc Am* 1989;85:2088-113.
- 4) Holt LL. Auditory constraints on speech perception: an examination of spectral contrast [dissertation]. Madison(WI): University of Wisconsin-Madison;1999.
- 5) Mann VA, Repp BH. Influence of preceding fricative on stop consonant perception. *J Acoust Soc Am* 1981;69:548-58.
- 6) Fowler CA. Production and perception of coarticulation among stressed and unstressed vowels. *J Speech Hear Res* 1981;24:127-39.
- 7) Liberman AM, Delattre P, Cooper FS. The role of selected stimulus-variables in the perception of the unvoiced stop consonants. *Am J Psychol* 1952;65:497-516.
- 8) Kewley-Port D. Measurement of formant transitions in naturally produced stop consonant-vowel syllables. *J Acoust Soc Am* 1982;72:379-89.
- 9) Mann V, Soli SD. Perceptual order and the effect of vocalic context on fricative perception. *Percept Psychophys* 1991;49:399-411.
- 10) Mann VA, Repp BH. Influence of vocalic context on perception of the zh to f distinction. *Percept Psychophys* 1980;28:213-28.
- 11) Whalen DH. Effects of vocalic formant transitions and vowel quality on the English [s]-[ʃ] boundary. *J Acoust Soc Am* 1981;69:275-82.
- 12) Cooper WE. Contingent feature analysis in speech perception. *Percept Psychophys* 1974;16:201-4.
- 13) Miller JL, Eimas PD. Studies on the selective tuning of feature detectors for speech. *J Phon* 1976;4:119-27.
- 14) Dubno JR, Levitt H. Predicting consonant confusions from acoustic analysis. *J Acoust Soc Am* 1981;69:249-61.
- 15) Singh S, Black JW. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *J Acoust Soc Am* 1966;39:372-87.
- 16) Wang MD, Bilger RC. Consonant confusions in noise: a study of perceptual features. *J Acoust Soc Am* 1973;54:1248-66.
- 17) Woods DL, Yund EW, Herron TJ, Ua Cruadhlaioch MA. Consonant identification in consonant-vowel-consonant syllables in speech-spectrum noise. *J Acoust Soc Am* 2010;127:1609-23.
- 18) Redford MA, Diehl RL. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *J Acoust Soc Am* 1999;106(3 Pt 1):1555-65.
- 19) McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature* 1976;264:746-8.
- 20) Green KP, Kuhl PK, Meltzoff AN. Factors affecting the integration of auditory and visual information in speech: the effect of vowel environment. *J Acoust Soc Am* 1988;84:S155.
- 21) Burnham D. Language specificity in the development of auditory-visual speech perception. In: Campbell R, Dodd B, Burnham D, editors. *Hearing eye II: advances in the psychology of speechreading and auditory-visual speech*. Hove: Psychology Press Ltd;1998. p.27-60.
- 22) Shigeno S. Influence of vowel context on the audio-visual speech perception of voiced stop consonants. *Jpn Psychol Res* 2000;42:155-67.
- 23) Manuel SY, Krakow RA. Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskin Lab Status Reports Speech Res* 1984;SR-77/78:69-78.
- 24) Magen H. Vowel-to-vowel coarticulation in English and Japanese. *J Acoust Soc Am* 1984;75:S41.
- 25) Manuel SY. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *J Acoust Soc Am* 1990;88:1286-98.
- 26) Boyce SE. Coarticulatory organization for lip rounding in Turkish and English. *J Acoust Soc Am* 1990;88:2584-95.
- 27) Kalaiah MK, Thomas D, Bhat JS, Ranjan R. Perception of consonants in speech-shaped noise among young and middle-aged adults. *J Int Adv Otol* 2016;12:184-8.
- 28) Bradlow AR, Torretta GM, Pisoni DB. Intelligibility of normal speech I: global and fine-grained acoustic-phonetic talker characteristics. *Speech Commun* 1996;20:255-72.
- 29) Hazan V, Markham D. Acoustic-phonetic correlates of talker intelligibility for adults and children. *J Acoust Soc Am* 2004;116:3108-18.
- 30) Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 1955;27:338-52.
- 31) Phatak SA, Allen JB. Consonant and vowel confusions in speech-weighted noise. *J Acoust Soc Am* 2007;121:2312-26.
- 32) Wagner A, Ernestus M. Identification of phonemes: differences between phoneme classes and the effect of class size. *Phonetica* 2008;65:106-27.