



<http://ui.ac.ir/en>

Journal of Researches in Linguistics

E-ISSN: 2322-3413

Document Type: Research Paper

12(1), 1-22

Received: 02.12.2019 Accepted: 29.02.2020

Does Perceptual Processing of Initial Consonant Clusters Lead to the Perception of Illusive Vowels in Persian?

Janane Shalpush

MA in Linguistics, Imam Khomeini International University, Iran

jshalpush@yahoo.com

Vahid Sadeghi*

Associate Professor, Department of English and Linguistics, Imam Khomeini International University, Iran

*Corresponding author, email: vsadeghi@hum.ikiu.ac.ir

Abstract

In this article, we investigated the role of Persian phonotactic parameters on the perception of Persian initial consonant clusters (ungrammatical) in the context of an experimental study. Given the limitations of the Persian syllable construction of having only one consonant at the beginning of the syllable, the purpose of this study was to answer the question whether Persian speakers perceive a vowel in nonsensically manipulated C1V1C2 sequences in which V1 is deleted in different degrees from the speech signal or they are found to be sensitive to the magnitude of vowel deletion in the sequences in which they differentiate between sequences in which V1 is deleted and those in which V1 is preserved in the speech signal. To conduct the study, twenty non-words with C1V1C2V2C3 syllable pattern were designed so that they contained one of the six vowels of the Persian sound system. The target non-words were re-synthesized in the Praat software after being produced and recorded in such manner that V1 in each non-word was deleted gradually from the speech signal in several steps. The results showed that the perception of an illusive vowel does not occur systematically for all C1C2 sequences and all Persian vowels. Thus, the hypothesis that Persian speakers perceive C1C2 sequences in initial consonant clusters as C1VC2 at the beginning of non-words, due to phonotactic constraints in the Persian sound system, is not valid.

Keywords: Initial Clusters, Illusive Vowel, Speech Perception, Non-words, Auditory Sensitivity.

Introduction

The phenomenon of perceptual illusions particularly that of illusory vowels has become a very useful probe to understand both loan word patterns and the speech perception mechanism. Previous research has unearthed a variety of factors that affect where illusory vowels are perceived. They are perceived in stimuli that contain: (a) consonantal sequences that are phonotactically illicit in the native language of the speaker (Dupoux et al., 1999; Durvasula and Kahng, 2016); (b) consonantal sequences that violate putative language universals such as the Sonority Sequencing Principle (Davidson, 2007), and (c) specific acoustic cues in the stimuli, such as strong consonantal release bursts, increased voicing amplitudes, etc. (Davidson, 2007).

What evidence exists that phonotactic constraints play a role in perception? Adults have rather clear intuitions about permissible sequences. For example, English speakers know that “mba” is not a possible English word. Davidson (2007) has shown that nine month old infants are sensitive to the phonotactic patterns of the words in their language and some researchers have argued that such regularities could be useful in helping the child to discover words. Mattingley et al. (2015) investigated the influence of phonotactic constraints on phoneme perception. They used the fact that /sri/ and /ʃli/ are not allowed in English while /sli/ and /ʃli/ are allowed. They synthesized a series of stimuli ranging from [s] to [ʃ] and



This is an open access article under the CC BY-NC-ND/4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).



<http://dx.doi.org/10.22108/jrl.2020.120165.1425>



<https://dori.net/dor/20.1001.1.20086261.1399.12.1.1.3>

presented them to participants in the / li/ and / ri/ context. There was a significant shift in the identification functions between the two contexts, demonstrating that participants tend to hear segments that respect the phonotactics of their language.

This paper addressed the role of Persian phonotactic constraints on the perception of initial consonant clusters. Given that initial consonant clusters are not permitted in the Persian sound system, we investigated whether Persian listeners perceive an illusive vowel in C1V1C2 sequences in which V1 is deleted in different degrees from the speech signal, or they are found to be sensitive to the magnitude of vowel deletion in the sequences.

Materials and Methods

Twenty non-words with C1V1C2V2C3 structure were designed such that they contained one of the six vowels of the Persian sound system. The target non-words were re-synthesized in the Praat software such that V1 was deleted from the speech tokens in magnitude degrees.

All acoustic measurements were made using Praat acoustic software (Boersma and Weenink, 2010). The stimuli were digitized on a PC Compatible computer using an OROS AU22 A/D board. Five different files were then created from each original item by splicing out pitch periods of the target vowel, V1, at zero crossings. Stimulus 1 contained little or no vowel (most of the transitions in and out of the vowel were also removed). Stimulus 2 contained the two most extreme pitch periods of the vowel (i.e., one from the transition of the first consonant to the vowel, and another from the end part of V1 into the following consonant). Stimulus 3 contained the four most extreme pitch periods (two on each side), and similarly, Stimulus 4 six pitch periods, and Stimulus 5 eight pitch periods. Stimulus 6 was the original stimulus in which the number of pitch periods varied from 10 to 13 across items (10.7 periods in average.) The average overall duration of one pitch period in V1 in each item was 9.06 ms. There were a total of 60 stimuli in one session.

Discussion of Results & Conclusions

Results suggested that listeners do not consistently perceive an illusive vowel in C1C2 sequences. Thus, our study does not provide sufficient evidence for confirming the hypothesis generally assumed in the phonetic literature that C1C2 sequences in languages which do not permit initial consonant clusters are perceived as C1VC2.

References

- Berent, I., T. Lennertz, J. Jun, M. Moreno, & P. Smolensky (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences* (105), 5321– 5325.
- Boersma, P., & D. Weenink (2018). Praat: doing phonetics by computer [Computer program]. Version 4.3.01, Retrived from <http://www.praat.org/>.
- Davidson, L. (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* 24 (2), 261–286.
- Davidson, L. & J. A. Shaw (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics* 40 (2), 234–248. ISSN 0095-4470. Doi: <http://dx.doi.org/10.1016/j.wocn.2011.11.005>.
- Dupoux, E., K. Kakehi, Y. Hirose, Ch. Pallier, & J. Mehler (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25 (6), 1568–1578. ISSN 1939 1277(ELECTRONIC);0096-1523(PRINT). Doi: 10.1037/0096-1523.25.6.1568.
- Dupoux, E., E. Parlato, S. Frota, Y. Hirose, & Sh. Peperkamp (2011). Where do illusory vowels come from? *Journal of Memory and Language* 64 (3), 199–210.
- Durvasula, K. & J. Kahng (2015). Illusory vowels in perceptual epenthesis: the role of phonological alternations. *Phonology* 32 (03), 385–416. ISSN 1469-8188. Doi: 10.1017/ S0952675715000263.
- Durvasula, K. & J. Kahng (2016). The role of phrasal phonology in speech perception: What perceptual epenthesis shows us. *Journal of Phonetics* 54, 15– 34. ISSN 0095-4470.

- Durvasula, K., H. H. Huang, S. Uehara, Q. Luo, & Y. H. Lin (2018). Phonology modulates the illusory vowels in perceptual illusions: Evidence from Mandarin and English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9 (1), 1-27.
- Fant, G., & K. T. Högskolan (1957). *Modern instruments and methods for acoustic studies of speech*. Stockholm: Royal Institute of Technology, Division of Telegraphy-Telephony.
- Guevara-Rukoz, A., I. Lin, M. Morii, Y. Minagawa, E. Dupoux, & Sh. Peperkamp (2017). Which epenthetic vowel? Phonetic categories versus acoustic detail in perceptual vowel epenthesis. *The Journal of the Acoustical Society of America* 142 (2), 35-59. EL211–EL217.
- Halle, M. (1998). The stress of English words. *Linguistic Inquiry* 29(4): 539-568.
- Jusczyk, P. W., A. D. Friederici, J. M. Wessels, V. Y. Svenkerud, & A. M. Jusczyk (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of memory and language* 32(3), 402-420.
- Jusczyk, P. W., P. A. Luce, & J. Charles-Luce (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33(5), 630-645.
- Kabak, B. & W. J. Idsardi (2007). Perceptual distortions in the adaptation of English consonant clusters: syllable structure or consonantal contact constraints? *Language and Speech* 50 (1), 23–52.
- Lewis, E., & M. Tatham (2001). Automatic Segmentation of recorded speech into syllables for speech synthesis. *Proceedings of Eurospeech 1: 1703-1707*, edited by J. Hirschberg & T. M. Mitchell. Denmark: Aalborg University.
- Mattingley, W., E. Hume, & K. C. Hall (2015). The influence of preceding consonant on perceptual epenthesis in Japanese. Paper presented at the 18th *International Congress of Phonetic Sciences*. 29-33. Glasgow: University of Edinburgh.
- Monahan, Ph. J., E. Takahashi, Ch. Nakao, & W. Idsardi (2009). Not All Epenthetic Contexts are Equal: Differential Effects in Japanese Illusory. In *Proceedings of the 17th Annual Japanese / Korean Linguistics Conference* edited by I. Shoichi, H. Hoji, P. M. Clancy, & S. O. Sohn. 391–405. Tokyo, Japan: University of Tokyo.
- Moreton, E. (2002). Structural constraints in the perception of English stop–sonorant clusters. *Cognition* 84, 55–71.
- Nearey, T. M. (1992). Context effect in a double-weak theory of speech Perception. *Language and speech* 35 (2): 153-171.
- Nearey, T. M. (1997). Speech Perception as pattern recognition. *Journal of Acoustic society of America* 101 (6), 3241 -3254.
- Peperkamp S., E. Dupoux (2002). A typological study of stress “deafness”. In *Laboratory Phonology 7* edited by C. Gussenhoven, & N. Warner. 40-203. Berlin: Mouton de Gruyter.
- Rahmani, H., T. Rietveld, & C. Gussenhoven (2015). Stress “Deafness” reveals absence of lexical marking of stress or tone in the adult grammar. *PLoS ONE* 10(12), 3-46.
- Rahmani, H., M. Bijankhan, & M. Ghajargar, (2012). Stress deafness in Persian speakers. *Audiology* 21 (3), 86: 102.(inpersian).
- Stevens, K. N. (1989). On the Quantal Nature of Speech. *Journal of Phonetics* 17:3-46.
- Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). Cambridge, MA: MIT press.
- Wilson, C., & L. Davidson (2013). Bayesian analysis of non-native cluster production. In *Proceedings of NELS 40* (2): 265–278, edited by N. LaCara, Y. Park & Y. Fainleib. Cambridge, MA: MIT Press.

آیا پردازش ادراکی خوشه‌های همخوانی آغازی در زبان فارسی منجر به درج واکه خیالی می‌شود؟

* جانانه شالیپوش

** وحید صادقی

چکیده

در این مقاله، نقش واج‌آرایی زبان فارسی در درک خوشه‌های آغازی (غیرمجاز) فارسی در چارچوب مطالعه آزمایشگاهی بررسی شده است. هدف از انجام این پژوهش، پاسخ‌گویی به این پرسش است که آیا با توجه به محدودیت‌های ساخت‌هیجایی زبان فارسی مبنی بر وجود تنها یک همخوان در آغاز هجا، فارسی‌زبانان در رشته‌های آوایی دست‌کاری‌شده بی‌معنی CIVIC2 واکه می‌شنوند که در آنها V1 به درجه‌های مختلف از سیگنال آوایی گفتار حذف شده است. یا آنکه به درجه‌های مختلف حذف V1 در سیگنال گفتار حساسیت شنیداری نشان می‌دهند؛ بنابراین، بین صورت‌های آوایی که در آنها V1 حذف شده و صورت‌هایی که در آنها V1 حذف نشده است، تمایز می‌گذارند. برای انجام این پژوهش، تعداد ۲۰ ناواژه با الگوی هجایی CIVIC2V2C3 طراحی شد؛ به طوری که ناواژه‌ها هر ۶ واکه زبان فارسی را در خود داشتند. ناواژه‌ها پس از تولید و ضبط در نرم‌افزار پرت بازسازی شدند؛ به این صورت که واکه اول درون هر ناواژه، به طور تدریجی طی چندین گام از سیگنال آوایی ناواژه حذف شد. نتایج به دست آمده نشان داد درک واکه‌های خیالی به طور نظام‌مند برای تمامی توالی‌های C1C2 و تمامی واکه‌های فارسی روی نمی‌دهد؛ بنابراین، این فرضیه که فارسی‌زبانان هجاهای حاوی خوشه‌های آغازی C1C2 در آغاز ناواژه‌ها را با توجه به محدودیت‌های نظام آوایی زبان فارسی به صورت توالی‌های آوایی CIVIC2 درک می‌کنند، فرضیه معتبری نیست.

کلیدواژه‌ها

خوشه‌های آغازی، واکه خیالی، درک گفتار، ناواژه‌ها، حساسیت شنیداری



۱. مقدمه

در نظریه صوت‌شناختی درک گفتار، درک پیام زبانی از رهگذر پردازش شنیداری الگوهای صوت‌شناختی در حوزه شنوایی و نگاشت الگوهای صوت‌شناختی در حوزه شنوایی و نگاشت الگوهای صوت‌شناختی بر مشخصه‌ها یا واحدهای آوایی ناپیوسته صورت می‌گیرد (Stevens, 1989:244-253; Nearey, 1997:3-10). براساس این نظریه، الگوهای صوت‌شناختی هر واج یا مشخصه واجی به صورت الگوهای آوایی مرجع در بازنمایی آوایی آن ذخیره می‌شود و درک آن واج یا مشخصه واجی از رهگذر انطباق الگوهای صوت‌شناختی درونداد با الگوهای آوایی مرجع در بازنمایی آوایی آن واحد واجی صورت می‌گیرد (Stevens, 1989:244-253; 2000:1-2)؛ از این رو، شنونده ابتدا امواج صوتی حاصل از فعالیت دستگاه گفتار را به نوعی بازنمود ادراکی می‌نگارد که دربرگیرنده مجموعه‌ای از پارامترهای آوایی ناپیوسته (دو ارزشی) است و سپس پیام زبانی را از رهگذر انطباق بین بازنمود ادراکی و الگوهای واجی درک می‌کند (Fant & Högskolan, 1957:2-6; Stevens, 1989:244-253). در چارچوب نظریه صوت‌شناختی درک گفتار، هر واحد واجی به مثابه یک واحد ادراکی در فضای صوت‌شناختی خود یک حوزه مغناطیسی ایجاد می‌کند و تمامی نمونه‌های وابسته به خود را جذب می‌کند (Nearey, 1992: 1-4; 1997:3-10). به عبارت روشن‌تر، الگوهای صوت‌شناختی مرجع در بازنمایی آوایی واحدهای واجی به گونه‌ای مشخص شده‌اند که شنونده می‌تواند مرز ادراکی میان آنها را تشخیص دهد و پارامترهای صوت‌شناختی را در نواحی ادراکی خاص خود بنگارد.

درک آواهای غیربومی در پژوهش‌های شنیداری به طور نظام‌مند بررسی شد. نتایج این پژوهش‌ها نشان می‌دهد هنگامی که یک آوا یا زنجیره آوایی ناآشنا دریافت می‌شود، سعی می‌کنیم آن را با نظام آوایی زبان خود همگون کنیم. یعنی در آن تغییرهایی ایجاد کنیم تا با ساخت واجی زبان مان هماهنگ شود. پیرکمپ^۱ و دوپو^۲ (2002) معتقدند این تغییرها، تغییرهای آوایی کمینه‌ای هستند که به هنگام درک گفتار ایجاد می‌شوند. از نظر آنها، نظام درک شنیداری گفتار شامل یک رمز گذارنده آوایی^۳ است که واحدهای آوایی غیربومی یا غیرمجاز را به نزدیک‌ترین مقوله آوایی ممکن در نظام آوایی زبان بومی می‌نگارد. به این ترتیب، صورت‌های واجی غیردستوری و بدساخت به نزدیکترین صورت‌های واجی خوش ساخت^۴ در واج‌شناسی بومی نگاشته می‌شوند.^۵

در پژوهش حاضر نقش واج‌آرایی زبان فارسی را در درک و پردازش شنیداری زنجیره‌های آوایی غیرمجاز فارسی در چارچوب یک مطالعه آزمایشگاهی بررسی کرده‌ایم. با توجه به محدودیت‌های ساخت هجایی زبان فارسی مبنی بر وجود تنها یک همخوان در آغاز هجا، هدف اصلی پژوهش پاسخ‌گویی به این پرسش است که آیا فارسی‌زبانان رشته‌های آوایی دست‌کاری‌شده بی‌معنی CIVIC2V2C3 را که در آنها V1 به درجه‌های مختلف از سیگنال آوایی گفتار حذف می‌شود، به صورت CIVIC2V2C3 درک می‌کنند. یعنی V1 را با وجود حذف کامل از زنجیره آوایی به یک واحد واکه‌ای در بازنمود واجی رشته آوایی پخش شده می‌نگارند. یا آنکه شنونده‌ها به درجه‌های مختلف حذف V1 در سیگنال گفتار حساسیت

1 S. Peperkamp

2. E. Dupoux

3. phonetic decoder

4. well-formed

5. map

شنیداری نشان می‌دهند؛ بنابراین، بین صورت‌های آوایی که در آنها V1 حذف شده و صورت‌هایی که در آنها V1 حذف نشده است، تمایز شنیداری برقرار می‌کنند.

مطابق با پیشینه پژوهش‌های انجام‌شده، فرض می‌شود دانش واجی فارسی‌زبانان بر پردازش درکی زنجیره‌های صوتی غیرمجاز تأثیرگذار است؛ بنابراین، چون در زبان فارسی هجاهای C1C2 وجود ندارد، فارسی‌زبانان این توالی آوایی را به صورت C1VC2 می‌شنوند. به بیان دیگر، شنونده‌های فارسی، خوشه‌های آوایی غیرمجاز C1C2 را با درج یک واکه خیالی ترمیم می‌کنند.

۲. پیشینه پژوهش: درک واکه خیالی

پژوهش‌های پیشین درباره درک توالی‌های آوایی غیرمجاز منجر به شکل‌گیری یافته‌های مهمی درباره سازوکارهای دخیل در پردازش ادراکی گفتار شده است: (۱) شنونده‌ها از همبسته‌های آکوستیکی مشخصه‌های واجی انتزاعی بهره می‌برند (Moreton, 2002)؛ (۲) اصول و الگوهای واجی انتزاعی مانند اصل توالی رسایی^۱ در درک گفتار تغییرهایی ایجاد می‌کند (Berent et al., 2008)؛ (۳) قواعد و الگوهای آوایی ساخت هجانقش مهم‌تری را در مقایسه با الگوهای آوایی مربوط به سطح واج ایفا می‌کند (Kabak & Idsardi, 2007)؛ (۴) به‌طور کلی، گویشوران توالی همخوان‌هایی که از لحاظ آوایی غیرمجاز هستند را به شیوه‌های مختلفی ترمیم می‌کنند و تنها روش ممکن، درج واکه‌های خیالی نیست (Davidson, 2007; Davidson & Shaw, 2007, 2013; Halle, 1998 & Wilson & Davidson, 2013).

جادشیک^۲ و همکاران (1993) و جادشیک و همکاران (1994) شواهدی ارائه کرده‌اند که نوزادان نُه‌ماهه به الگوهای آوایی زبان مادری‌شان حساسیت نشان می‌دهند. پژوهشگران بر این باورند که چنین الگوهای در بازیابی و ذخیره‌سازی واژه‌ها به کودکان کمک می‌کنند.

پژوهش‌هایی صورت گرفته در رابطه با درک توالی‌های غیرمجاز آوایی، بیشتر متکی به پدیده درج واکه‌ها است. واکه‌های خیالی در محرک‌هایی درک می‌شوند که دارای توالی‌های آوایی ذیل باشند: (۱) توالی همخوان‌هایی که از لحاظ آوایی در زبان بومی گویشور غیرمجاز هستند؛ (۲) توالی همخوان‌هایی که جهانی‌های پذیرفته‌شده واج‌آرایی مانند اصل توالی رسایی را نقض می‌کنند؛ (۳) محرک‌هایی که در آنها سرنخ‌های آکوستیکی مشخصی وجود داشته باشد، مانند رهش قوی همخوان‌ها و افزایش دامنه واکه‌داری.

دوپو و همکاران (1999) درج واکه‌های خیالی در زبان ژاپنی را بررسی کردند و ثابت کردند الگوهای آوایی زبان مادری نقش مهمی در شناخت و پردازش درکی صورت‌های آوایی زبان غیربومی دارد. آنها در پژوهش خود چهار مطالعه آزمایشگاهی جداگانه انجام دادند و محرک‌های شنیداری را بر روی دو جامعه آماری بررسی کردند: (۱) گویشوران بومی زبان ژاپنی؛ (۲) گویشوران بومی زبان فرانسوی. لازم به توضیح است که در زبان ژاپنی توالی همخوان‌ها در آغاز، وسط و پایان کلمه غیرمجاز است. در مقابل، در زبان فرانسوی توالی همخوانی در وسط کلمه تظاهر آوایی دارد. از سوی دیگر، در زبان ژاپنی کشش واکه نقش تقابلی دارد و واژه‌ها می‌توانند شامل ترکیب متوالی چهار واکه متفاوت باشند. اما در زبان فرانسوی هیچ جفت‌واژه‌ای وجود ندارد که در کشش واکه با یکدیگر تفاوت داشته باشند. در آزمایش‌های اول و دوم داده‌هایی ارائه شد که برخی از آنها دارای

1. Sonority Sequencing Principle

2. P. W. Jusczyk

واکه و برخی دیگر فاقد واکه بودند؛ مانند Ebzo و Ebuzo. نتایج این آزمایش نشان داد ژاپنی‌ها برخلاف فرانسوی‌ها، در بین توالی‌های همخوانی فاقد واکه، یک واکه خیالی به صورت [u] می‌شنوند و این در حالی بود که هیچ واکه‌ای در محرک‌های مدنظر وجود نداشت. در آزمایش سوم و چهارم شرکت‌کنندگان ژاپنی در تشخیص توالی‌های [VCuCV] و [VCCV] (مانند واژه‌های Ebuzo و Ebuzo) از یکدیگر دچار مشکل شدند. در مقابل، شرکت‌کنندگان فرانسوی در تشخیص محرک‌هایی مشکل پیدا کردند که طول واکه‌هایشان با هم تفاوت داشت؛ چراکه زبان فرانسوی دارای ساخت هجایی پیچیده است و کشش واکه در این زبان تمایزدهنده نیست. یافته‌های دوپو و همکاران (1999) این فرضیه را تأیید کرد که درک گفتار به شدت تحت تأثیر دانش واجی است؛ بنابراین، دانش واجی نه تنها در دسته‌بندی واج‌ها تأثیرگذار است، بلکه در درک آواهای خیالی که هیچ همبسته آکوستیکی ندارند، نیز مؤثر است. این واقعیت نشان می‌دهد روشی که براساس آن زنجیره پیوسته گفتار به واج‌های انتزاعی افزای می‌شود، روش جهانی نیست و به نظام آوایی همخوان‌ها و واکه‌ها در زبان مدنظر بستگی دارد؛ بنابراین، هنگامی که آواهای یک زبان غیربومی را دریافت می‌کنیم، نه تنها تمایل داریم آنها را با مقوله‌های زبانی خودمان همگون کنیم، بلکه حتی ممکن است مؤلفه‌هایی را تغییر دهیم تا آنها را با الگوها و قواعد آوایی زبان خود منطبق سازیم.

در پژوهشی دیگر موناها^۱ و همکاران (2009) واکه‌های خیالی در بافت‌های آوایی مختلف در زبان ژاپنی را بررسی کردند. آنها طبق بافت واجی و ام‌واژه‌ها در زبان ژاپنی و بررسی‌های انجام‌شده بر روی واژه‌های مختلف نتیجه گرفتند که واکه خیالی در مجاورت همخوان‌های غیرتیغه‌ای به صورت واکه [u] و در مجاورت همخوان‌های تیغه‌ای به صورت واکه [o] درک می‌شود. همچنین، آنها دریافتند گویشوران ژاپنی بیشتر از گویشوران انگلیسی تمایل به درک واکه خیالی [u] در مجاورت همخوان‌های غیرتیغه‌ای دارند و درک واکه خیالی [o] در مجاورت همخوان‌های تیغه‌ای تمایلی غالب و مسلط نیست.

دوپو و همکاران (2011) درباره کیفیت واکه‌های خیالی بحث کرده‌اند و نشان داده‌اند واکه خیالی کوتاه‌ترین واکه در نظام آوایی یک زبان است. او با استناد به پژوهش‌های آکوستیکی روی واکه‌های زبان کره‌ای نشان داد که کوتاه‌ترین واکه در این زبان [u] است (دیرش این واکه در ابتدای عبارت‌ها حدود ۱۴۴ میلی‌ثانیه است، در حالی که دیرش واکه‌های [i] و [u] در جایگاهی یکسان به ترتیب ۱۶۰ و ۱۶۵ میلی‌ثانیه است). از نظر دوپو و همکاران (2011) شواهد کافی برای اثبات [u] به‌مثابه واکه خیالی مهیا می‌شود؛ چرا که کوتاه‌ترین واکه است و دارای کمترین میزان دیرش در مقایسه با واکه‌های دیگر است.

دوپو و همکاران (2011) نشان دادند گویشوران ژاپنی در محرک‌هایی که توالی همخوان‌هایشان از لحاظ آوایی غیرمجاز است، واکه خیالی [u] را درک می‌کنند و این در حالی است که گویشوران زبان پرتغالی (لهجه برزیلی) در توالی‌های همخوانی غیرمجاز، واکه خیالی [i] را درک می‌کنند. آنها نشان دادند که این واکه‌ها جزو کوتاه‌ترین واکه‌ها در زبان‌های ژاپنی و لهجه برزیلی زبان پرتغالی هستند و احتمالاً به همین دلیل است که به‌مثابه واکه‌های خیالی در توالی‌های همخوانی غیرمجاز درج می‌شوند. همچنین، آنها نشان دادند گویشوران ژاپنی با شنیدن برخی محرک‌ها که در آنها توالی همخوان‌های غیرمجاز از طریق درج واکه میانی به یک توالی مجاز تبدیل می‌شود، واکه درج‌شده را درک کردند، نه واکه خیالی [u] را. بنابراین، واکه‌های درک‌شده در چنین بافت‌هایی در حقیقت، واکه‌های خیالی نیستند و باید این دو الگوی درکی را از هم تمییز دهیم. مطالعه دوپو و همکاران (2011) در حقیقت، نشان داد وقتی محرک‌ها از سیگنال گفتار طبیعی استخراج می‌شوند و مصنوعی یا بازسازی شده نیستند، واکه‌های خیالی متفاوتی در زبان‌های مختلف درک می‌شوند.

دوپو و همکاران (2011) در مطالعه‌ای آزمایشگاهی خوشه‌های همخوانی و توالی‌های مجاز و غیرمجاز در زبان ژاپنی را بررسی کردند (چنان که گفته شد، خوشه‌های همخوانی در زبان ژاپنی غیرمجازند و از این رو، در وام‌واژه‌ها با درج واکه^۱ شکسته می‌شوند). آنها برای تبیین پدیده درج واکه در زبان ژاپنی استدلال‌های مختلفی را مدنظر قرار دادند. اول آنکه، پدیده درج واکه در گفتار رخ می‌دهد که دلیل آن ناتوانی ژاپنی‌ها در تولید خوشه‌های همخوانی است. پس علت درج واکه سهولت در تولید گفتار است. دلیل دوم، مربوط به نظام نوشتاری^۲ زبان ژاپنی است. برخی توالی‌ها در نظام نوشتاری کانجی (خط رسمی ژاپنی) غیرمجازند. مثلاً در ژاپنی هیچ واژه‌ای وجود ندارد که به صورت /sfinks/ نوشته شده باشد و چنین توالی‌ای غیرمجاز است و با نظام نوشتاری کانجی مغایرت دارد. دوپو و همکاران (2011) دلیل سوم را نیز مطرح کردند و آن اینکه درج واکه در سطح شنیداری رخ می‌دهد؛ یعنی شنونده بیشتر تمایل دارد واحدهایی را بشنود و درک کند که با قواعد آوایی و واجی زبان خودش هماهنگ باشد. آنها با بیان استدلال‌هایی نشان دادند دلیل سوم در مقایسه با دو دلیل دیگر، کفایت تبیینی بیشتری دارد؛ بنابراین، باید واقعیت درکی درج واکه را مدنظر قرار داد.

متینگلی^۳ و همکاران (2015) در مطالعه‌ای آزمایشگاهی، درک واکه‌های خیالی گویشوران ژاپنی را بررسی کردند. آنها بر مبنای بافت واجی وام‌واژه‌ها در زبان ژاپنی، انتظار تظاهر واکه‌های خیالی متفاوتی را در بافت‌های همخوانی مختلف داشتند و از این رو، بافت‌های آوایی بیشتری را بررسی کردند. نتایج این مطالعه نشان داد با اینکه شرکت‌کنندگان ژاپنی در توالی‌های همخوانی مختلف، واکه‌های متفاوتی را درک می‌کنند، واکه‌هایی که درک می‌کنند مبتنی بر بافت واجی وام‌واژه‌ها نیست.

دورواسولا^۴ و کانگ^۵ (2015) نحوه درک شنیداری توالی‌های همخوانی غیرمجاز گویشوران کره‌ای و انگلیسی را بررسی کردند. آنها نشان دادند گویشوران کره‌ای، برخلاف انگلیسی‌زبان‌ها، واکه‌های خیالی مختلفی را در بافت‌های همخوانی متفاوت درک می‌کنند و حتی در مواردی واکه‌های خیالی متفاوتی را در یک بافت آوایی یکسان دریافت می‌کنند.

در پژوهش دیگری، گوورا روکاز^۶ و همکاران (2017) درک واکه‌های خیالی در دو بافت واجی [VhpV] و [VkpV] در زبان ژاپنی را بررسی کردند. نتایج این پژوهش نشان داد سرنخ‌های صوتی نقش مهمی در درک شنیداری و شناسایی واکه‌های خیالی دارد. اما کشف و استخراج سرنخ‌های صوتی مربوط به کیفیت واکه‌های خیالی بسیار دشوار است و بررسی‌های بینابانی فراوانی باید در این زمینه صورت بگیرد.

دورواسولا و کانگ (2016) در پژوهشی آزمایشگاهی این موضوع را بررسی کردند که آیا شنونده‌ها دانش واجی خود را در حین درک گفتار در سطوح بالاتر از واژه به کار می‌برند یا خیر. آنها به‌طور خاص عملکرد شنیداری شنونده‌های زبان‌های کره‌ای و انگلیسی را در به‌کارگیری دانش واجی در سطح گروه آهنگ در حین درک گفتار بررسی کردند. نتایج به‌دست آمده نشان داد شنونده‌های کره‌ای واکه‌های خیالی بیشتری را در محرک‌های بی‌معنی‌ای که از توالی انفجاری‌های واکدار و همخوان خیشومی تشکیل شده است، دریافت می‌کنند تا محرک‌هایی که توالی‌ای از انفجاری‌های بی‌واک و همخوان خیشومی دارند. با این حال، درک واکه درج‌شده نمی‌تواند نتیجه رفتار آوایی حوزه‌های کوچک‌تر از گروه آهنگ باشد، به دلیل اینکه هر دو

1. vowel insertion
2. orthography system
3. W. Mattingley
4. K. Durvasula
5. J. Kahng
6. A. Guevara-Rukoz

توالی در حوزه‌های کوچک‌تر از گروه آهنگ، غیرمجاز تلقی می‌شود. آنها با استناد به نتایج خود این گونه بحث کردند که کار اصلی شنونده در درک گفتار، استنتاج معکوس^۱ است. یعنی شنونده با بهره‌گیری از اطلاعات آکوستیکی سیگنال آوایی، زیرساخت مدنظر در پاره‌گفتار را شناسایی می‌کند. هنگامی که شنونده کره‌ای توالی بی‌معنی همخوان انفجاری «بی‌واک + همخوان خیشومی» را می‌شنود، تشخیص می‌دهد که چنین توالی دارای مرزی در گروه آهنگ است و در نتیجه نیازی نیست تا چنین توالی را با درج واکه خیالی ترمیم کند. در مقابل، اگر شنونده کره‌ای توالی بی‌معنی انفجاری‌های واگذار و همخوان خیشومی را در سطح گروه آهنگ (و نه مرز بین دو گروه آهنگ) بشنود، به‌سرعت درک می‌کند که این توالی از لحاظ واجی غیرمجاز است و هیچ مرزی را نمی‌تواند شناسایی کند تا براساس آن چنین توالی به‌مثابه یک توالی آوایی مجاز ظاهر یابد.

دورواسولا و همکاران (۲۰۱۸) درک واکه خیالی را حاصل استنتاج معکوس از اطلاعات آکوستیکی روساختی در نظر می‌گیرند. لازمه این دیدگاه شناخت تناوب‌های آوایی، قواعد و الگوهای آوایی حاکم بر استنتاج بازنمایی آوایی از بازنمایی واجی است. بر این اساس، در رابطه با کیفیت واکه‌های خیالی، شنونده باید تلاش کند تا توالی‌های آوایی غیرمجاز را با واکه‌ای ترمیم کند تا ساخت واجی غیردستوری محرک آوایی با زیرساخت واجی زبان هماهنگ شود. زمانی که هیچ تناوب آوایی شنونده‌ها را به‌سمت واکه مشخصی در بافت واجی مدنظر سوق ندهد، بهترین واکه‌ای که می‌تواند نقض آوایی صورت گرفته را ترمیم کند یا براساس سرنخ‌های صوتی مربوطه انتخاب می‌شود و یا براساس کوتاه‌ترین واکه موجود در زبان. اما زمانی که تناوب‌های آوایی مربوطه، شنونده‌ها را به‌سوی درک واکه خاصی سوق دهند، انتخاب بهترین واکه به ویژگی‌های محرک آوایی مدنظر و تناوب‌های آوایی مربوطه وابسته است.

رحمانی و همکاران (۱۳۹۱) در یک مطالعه آزمایشگاهی در چارچوب انگاره‌ناشنوایی تکیه و با بهره‌گیری از محرک‌های مصنوعی، توانایی درکی عمومی تکیه از سوی گویشوران فارسی را بررسی کردند. آنها ادعا کردند که طبق انگاره‌ناشنوایی تکیه، چون محل وقوع تکیه در زبان فارسی قابل پیش‌بینی است، اطلاعات مربوط به تکیه در واژگان زبان فارسی ذخیره نمی‌شود.^۲ شرکت‌کنندگان این پژوهش شامل سه گروه مختلف بودند: الف) فارسی‌زبانان بومی بدون آشنایی با زبان انگلیسی؛ ب) فارسی‌زبانانی که حداقل ۵ سال در یک کشور انگلیسی‌زبان زندگی کرده بودند؛ و ج) فارغ‌التحصیلان مقطع کارشناسی ارشد و دکتری رشته زبان‌شناسی. یافته‌های این پژوهش نشان داد: ۱) حساسیت کلی فارسی‌زبانان نسبت به تکیه ضعیف است؛ ۲) دانش آواشناسی در درک همبسته‌های تکیه گویشوران فارسی تأثیری ندارد و تفاوت میزان درک شرکت‌کنندگان دارای دانش آواشناسی با شرکت‌کنندگان فاقد این دانش، به‌لحاظ آماری معنی‌دار نیست؛ ۳) قرار گرفتن فارسی‌زبانان در محیط زبانی انگلیسی، موجب افزایش حساسیت آنها نسبت به تکیه می‌شود به‌صورتی که عملکرد کلی تشخیص تکیه در آنها تا ۵۰ درصد افزایش می‌یابد؛ ۴) با بررسی الگوی درک هر یک از همبسته‌های آکوستیکی تکیه (فرکانس پایه، دیرش و شدت انرژی) به‌طور مجزا از هم، نتیجه فوق‌درباره پایین بودن حساسیت نسبت به تکیه در فارسی‌زبانان تأیید شد. فارسی‌زبانانی که به‌صورت مستمر در معرض زبان انگلیسی نبودند، هر سه همبسته را در مقایسه با آنهايي که در مواجهه طولانی‌مدت با زبان انگلیسی بودند،

1. reverse inference

۲. از منظر الگوی ناشنوایی تکیه، زبان فارسی در گروه زبان‌های دارای تکیه ثابت قرار می‌گیرد؛ اما باید توجه داشت که محل وقوع این تکیه همواره در لبه پاره‌گفتار نیست و فرایندهای صرفی و نحوی در مشخص کردن محل وقوع آن بی‌تأثیر نیست. به بیان جامع‌تر، برجستگی نوایی در زبان فارسی مستقل از عوامل دستوری نیست؛ یعنی زبان فارسی در دسته‌بندی زبان‌هایی قرار می‌گیرد که پارامتر تکیه در آن بعد از حصول توانایی کودک در تقطیع کلمه‌های نقشی یا محتوایی از زنجیره‌گفتار، ارزش‌دهی می‌شود.

ضعیف‌تر تشخیص دادند؛ و بالاخره اینکه (فارسی‌زبانان در تشخیص همبسته دیرش مشکل دارند و بهترین سطح تشخیص، متعلق به همبسته فرکانس پایه است، هر چند که به لحاظ آماری تفاوتی میان میزان درک فرکانس پایه و شدت انرژی وجود ندارد. رحمانی و همکاران (۱۳۹۱) با توجه به نتایج فوق چنین بحث کردند که براساس الگوی ناشنوایی تکیه، حساسیت درکی به مشخصه‌های نوایی، تابع نظام واج‌شناسی نوایی زبان اول است و با قرار گرفتن در معرض زبان دوم تغییر چندانی نمی‌کند یا بدون تغییر باقی می‌ماند. پس هر قدر تکیه در زبان اول بیشتر پیش‌بینی‌پذیر باشد، حساسیت درکی گویشور به تکیه تضعیف خواهد شد؛ زیرا الگوهای منظم تکیه نیاز به رمزگذاری و ازگانی ندارند. در واقع، توانایی درک تکیه زمانی تقویت می‌شود که تکیه در بازنمایی ذهنی وجود داشته باشد.

رحمانی و همکاران (۲۰۱۵) در پژوهشی نحوه درک تکیه از سوی گویشوران فرانسوی، اسپانیایی، فنلاندی، لهستانی و مجارستانی را بررسی کردند. آنها در آزمایش شنیداری خود مشاهده کردند که گویشوران فرانسوی در پردازش تقابل‌های تکیه‌ای موجود در محرک‌های آزمایشی بی‌معنی، مشکلات زیادی دارند. فنلاندی‌ها هم ضمن آزمایش با چنین مشکلاتی مواجه بودند و عملکرد تشخیصی آنها به‌طور کلی ضعیف بود (لازم به توضیح است که در دو زبان فرانسوی و فنلاندی، تکیه نقش تقابلی ندارد). اما گویشوران اسپانیایی که تکیه در آن نقش تقابلی دارد، مشکلی در تشخیص جایگاه تکیه در محرک‌ها نداشتند. عملکرد گویشوران مجارستانی و لهستانی در این آزمایش به‌گونه‌ای بود که به لحاظ رتبه‌بندی، میان فرانسه / فنلاندی و اسپانیایی قرار می‌گرفتند (باید اشاره کرد که در زبان‌های مجارستانی و لهستانی، تکیه به لحاظ قابلیت تمایزدهندگی در حد فاصل بین زبان‌های فرانسوی و اسپانیایی قرار می‌گیرد). بر این اساس، رحمانی و همکاران (۲۰۱۵) صورتی از الگوی ناشنوایی تکیه پیشنهاد کردند که طبق آن، زبان‌های دارای تکیه ثابت، براساس میزان پیش‌بینی‌پذیری تکیه، رتبه‌بندی می‌شوند. هر قدر تکیه در زبانی پیش‌بینی‌پذیرتر باشد، برای گویشوران این زبان تشخیص تکیه مشکل‌تر خواهد بود.

۳. روش پژوهش: آزمایشگاهی

با توجه به محدودیت‌های نظام آوایی فارسی در رابطه با وجود خوشه‌های همخوانی در آغاز هجا، درک واکه خیالی در هجاهای دارای خوشه‌های آغازی C1C2 در رشته‌های آوایی بی‌معنی بررسی شد. برای ساخت داده‌های صوتی پژوهش، ابتدا تعداد ۲۰ ناواژه با الگوی هجایی C1V1C2V2C3 طراحی شد به طوری که V1 هر ۶ واکه زبان فارسی را دربرمی‌گرفت. ناواژه‌ها را دو گویشور فارسی زبان (یک زن و یک مرد) با محدوده سنی ۲۰ تا ۲۵ سال، دارای تحصیلات دانشگاهی و با گویش فارسی معیار به صورت داوطلبانه در محیط آزمایشگاهی تولید کردند. به این ترتیب، ۴۰ محرک صوتی بی‌معنی (ناواژه صوتی) به دست آمد.

ضبط داده‌ها در اتاق آزمایشگاه آواشناسی در سکوت کامل انجام شد. داده‌ها با استفاده از میکروفون بیبردینامیک با پاسخ بسامدی پنجاه تا پانزده هزار هرتز با فرکانس نمونه‌برداری ۱۱۰۲۵ هرتز بر روی کارت صوتی کریتو مدل ساند بلاستر X-Fi 5.1 یک رایانه شخصی ضبط شدند. ناواژه‌ها روی صفحه نمایشگر رایانه، به خط فارسی، برای آزمودنی‌ها (دو گوینده مرحله اول) نمایش داده شد. از آزمودنی‌ها خواسته شد ۲۰ ناواژه را یکی پس از دیگری به صورت طبیعی و با مکث کوتاه بخوانند.

جدول ۱. ناواژه‌های استفاده‌شده در آزمایش درک واکه خیالی (ساخت هجایی همه ناواژه‌ها به صورت C1V1C2V2C3 است)

Table 1. Non-words used in the perception experiment of illusive vowels (the structure of all syllables is C1V1C2V2C3V)

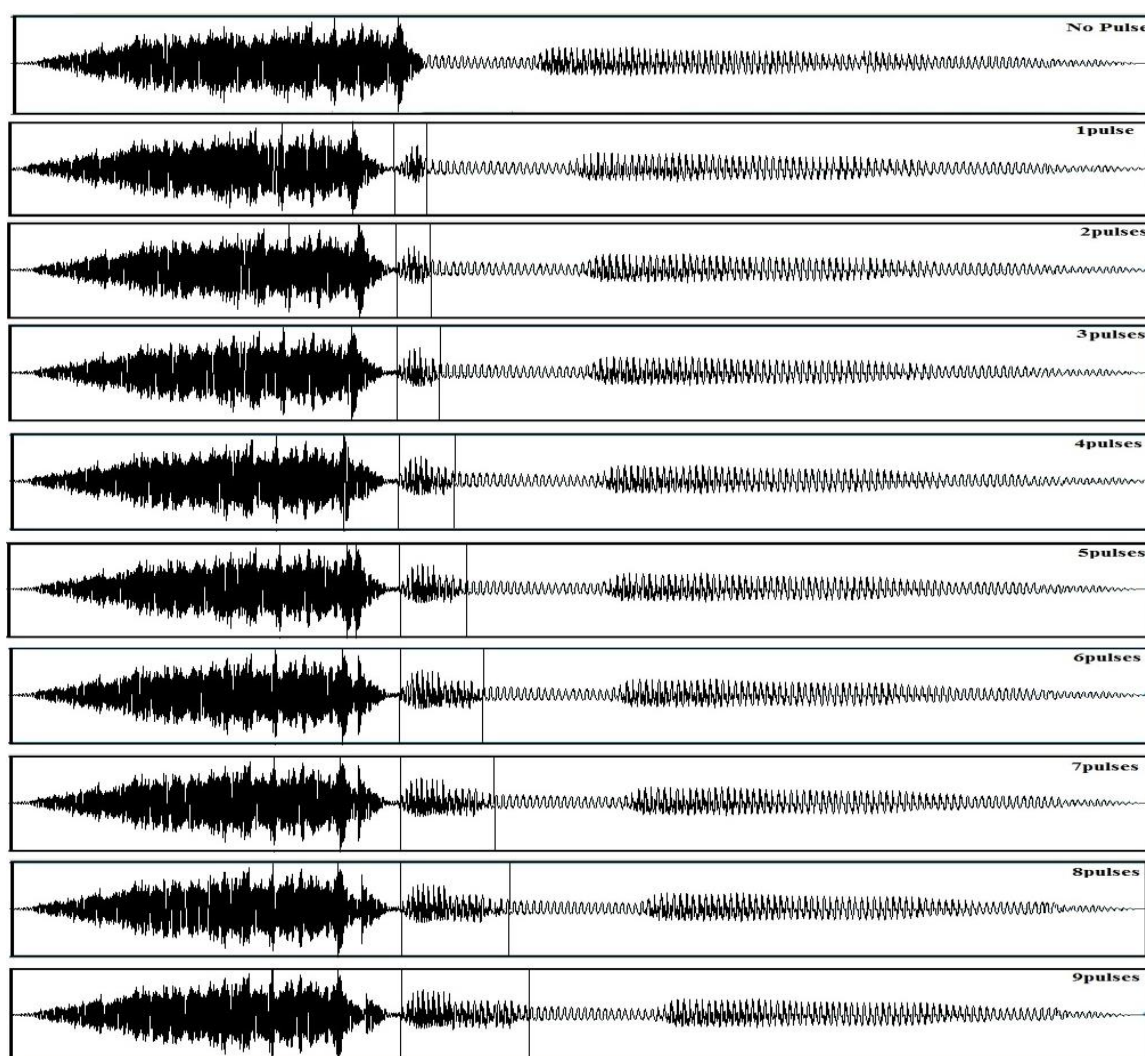
واکه	ناواژه	واکه	ناواژه
i	/tizor/, /sidæf/, /zikæt/	u	/bufan/, /tusad/, /fudaf/
e	/meræk/, /defæs/, /seram/, /pebat/	o	/noris/, /lobar/, /soban/
æ	/lædet/, /sælæb/, /færuʃ/, /tæsæl/	ɑ	/salum/, /zasam/, /maʒub/

در مرحله بعد، مقادیر دیرش واکه V1 در محرک‌ها در نرم‌افزار پرت، ویرایش ۴,۳,۰۱ (بورزما^۱ و وینینک^۲، ۲۰۱۸) بازسازی شد. بازسازی دیرش V1 طی ده گام انجام شد. در گام اول سیگنال صوتی واکه به کلی از زنجیره آوایی محرک‌ها حذف گردید. در گام دوم، ۱ تناوب (چاکنایی) واکه‌ای در ابتدا و ۱ تناوب واکه‌ای در انتهای واکه V1 در زنجیره آوایی محرک‌ها حفظ و مابقی سیگنال آوایی واکه V1 از محرک‌ها حذف گردید. در گام سوم، ۲ تناوب واکه‌ای در ابتدا و ۲ تناوب واکه‌ای در انتهای واکه V1 در زنجیره آوایی محرک‌ها حفظ و بقیه سیگنال آوایی V1 از محرک‌ها حذف شد. به همین ترتیب، در گام‌های چهارم تا دهم به تدریج تعداد بیشتری تناوب واکه‌ای در ابتدا و انتهای واکه V1 در زنجیره‌های آوایی هدف حفظ و مابقی سیگنال آوایی V1 از محرک‌ها حذف شدند (مثلاً برای گام چهارم تعداد ۳ تناوب واکه‌ای در ابتدا و ۳ تناوب واکه‌ای در انتهای واکه V1 و برای گام دهم تعداد ۹ تناوب واکه‌ای در ابتدا و ۹ تناوب واکه‌ای در انتهای V1 حفظ و بقیه سیگنال V1 از زنجیره‌های صوتی حذف شدند).

برای تعیین مرز هجا در زنجیره‌های آوایی V.C از معیارهای ارائه‌شده در لوئیز^۳ و تاتهام^۴ (۲۰۰۱) استفاده شد. بر این اساس، در زنجیره آوایی V. Stop مرز هجا جایی در نظر گرفته شد که دامنه ارتعاش تارآواها برای واکه در هجای اول به میزان درخور توجهی کاهش می‌یابد. این ناحیه ناظر بر آغاز محدوده بست همخوان انفجاری است. برای همخوان‌های انفجاری بی‌واک تشخیص پایان واکه آسان است؛ چون به دنبال واکه یک دوره بی‌واکی کامل همراه با سکوت و نوفه ظاهر می‌شود. ولی برای انسدادی‌های واگذار این تشخیص سخت‌تر است؛ چون ارتعاش تارآواها برای واکه در هجای اول به مدت ۲۰ تا ۳۰ میلی‌ثانیه پس از شروع هجای دوم ادامه می‌یابد. الگوی تقطیع هجا برای V. Fricative مشابه V. stop است با این تفاوت که برای سایشی‌های واگذار نوفه و واک در پایان واکه به‌طور همزمان تظاهر می‌یابد. در زنجیره آوایی V. Liquid چون تظاهر همخوان‌های روان روی طیف‌نگاشت به صورت امواج پیچ‌درپیچ و مدور است، شروع این ناحیه به‌مثابه نقطه شروع هجای دوم در نظر گرفته می‌شود. تعیین لحظه شروع این ناحیه روی طیف‌نگاشت آسان نیست. معمولاً حداقل میزان دامنه این ناحیه محل شروع هجای دوم محسوب می‌شود.

1. P. Boersma
2. D. Weenink
3. E. Lewis
4. M. Tatham

برای بازسازی محرک‌ها از نرم‌افزار پرت ویرایش ۴,۳,۰۱ (Boersma & Weenink, 2018) استفاده شد. برای بازسازی محرک‌های دیرش، ابتدا علامت آوایی هر پاره‌گفتار از طریق گزینه دستکاری به یک بازنمود دوبعدی از فرکانس پالس‌های حنجره در واحد زمان تبدیل و از خروجی لایه دیرش این بازنمود برای بازسازی و تنظیمات مقادیر دیرش استفاده شد. در مرحله بعد، از ۲۰ شرکت‌کننده (۱۰ زن و ۱۰ مرد) با محدوده سنی ۲۰ تا ۳۵ سال، دارای تحصیلات دانشگاهی و با گویش فارسی معیار خواسته شد تا به صورت تصادفی به ناواژه‌های تولیدشده گوش بدهند و پس از شنیدن هر محرک، قضاوت کنند که آیا در ناواژه تولیدشده، واکه‌ای در حد فاصل بین دو همخوان اول C1C2 درک می‌کنند یا خیر. هر محرک دو بار برای هر شرکت‌کننده پخش شد و به‌ازای هر محرک، درصد پاسخ هر شرکت‌کننده به هر گزینه محاسبه شد.



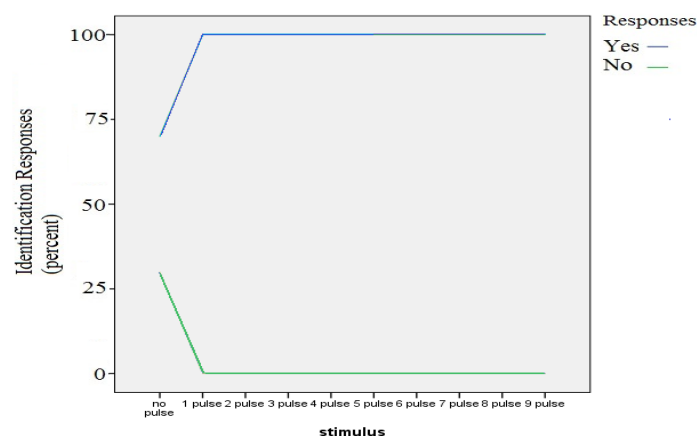
شکل ۱. محرک‌های صوتی بازسازی‌شده از ناواژه zikæt. در گام اول (no pulse) سیگنال صوتی واکه V1 به کلی از زنجیره آوایی حذف شده است و در گام‌های بعد به تدریج بین ۱ تا ۹ تناوب واکه‌ای در ابتدا و انتهای سیگنال واکه‌ای حفظ و مابقی سیگنال آوایی V1 از محرک‌ها حذف شده است.

Figure 1. Synthesized tokens for the non-word zikæt. In the first step (no pulse), V1 was completely deleted from the speech signal and in the following steps, one to nine glottal pulses were preserved at the beginning and end of the target vowel, V1, and the rest were deleted.

۴. نتایج

نتایج آزمون شناسایی در شکل‌های ۲ تا ۷ نشان داده شده است. شکل‌ها درصد پاسخ‌های شناسایی به محرک‌های پیوستار تعداد تناوب‌های واکه‌ای را نشان می‌دهد. در هر شکل محور X محرک‌های آوایی هدف آزمایش را نشان می‌دهد که از No pulse (نبود هیچ تناوب واکه‌ای در محرک صوتی) تا 9th pulse (حضور ۹ تناوب واکه‌ای در آغاز و ۹ تناوب واکه‌ای در پایان محرک صوتی) مرتب شده است. محور Y درصد پاسخ‌های شناسایی را نشان می‌دهد. منحنی آبی‌رنگ پاسخ‌های شناسایی مثبت یا «بلی» به حضور واکه و منحنی سبز رنگ پاسخ‌های شناسایی منفی یا «خیر» به حضور واکه را نشان می‌دهد. به بیان دیگر، پاسخ‌های شناسایی مثبت به معنای درک واکه در سیگنال محرک صوتی پخش شده و پاسخ‌های شناسایی منفی به معنای درک نکردن واکه در سیگنال محرک آوایی مدنظر است. این دو منحنی را به اختصار منحنی خوانش واکه‌ای (منحنی آبی‌رنگ) و منحنی خوانش غیرواکه‌ای (منحنی سبز رنگ) می‌گوییم.

ابتدا شکل (۲) را در نظر می‌گیریم. این شکل پاسخ‌های شناسایی به گام‌های تناوب چاکنایی را برای محرک CiCC نشان می‌دهد. چنانکه در شکل مشاهده می‌شود توزیع پاسخ‌های شناسایی برای CiCC بر روی تمامی گام‌های صوتی برای خوانش واکه‌ای از خوانش غیرواکه‌ای به‌طور درخور توجهی بیشتر است. نسبت پاسخ‌های واکه‌ای به غیرواکه‌ای برای گام‌های دوم تا دهم ۱۰۰ به صفر و برای گام اول ۷۰ به ۳۰ درصد است. به‌طوری که مشاهده می‌شود پاسخ‌های خوانش غیرواکه‌ای به محرک صوتی اول، یعنی ناحیه No pulse حدود ۳۰ درصد است؛ اما از گام دوم تا دهم، درصد پاسخ‌های خوانش غیرواکه‌ای کاهش پیدا کرده و سطح اطمینان آنها به صفر درصد رسیده است.



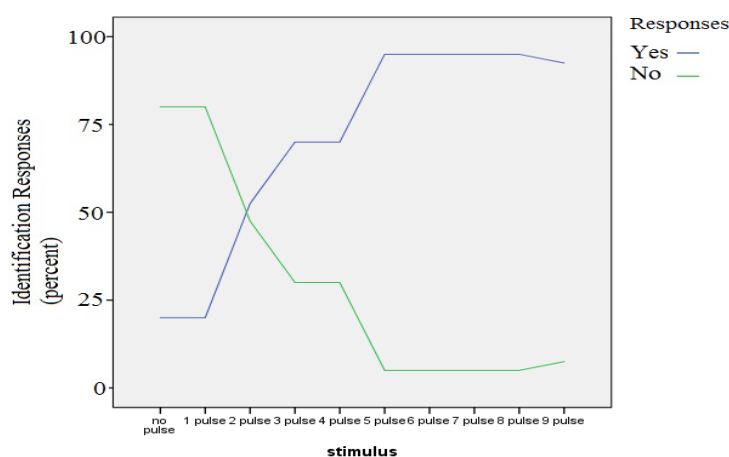
شکل ۲. نمودار خطی پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای برای محرک CiCC

Figure 2. Line graph of identification responses (in percent) to steps of glottal pulses for CiCC.

در مقابل، پاسخ‌های خوانش واکه‌ای بر روی گام نخست سطح اطمینانی حدوداً ۷۰ درصدی دارد و در نُه گام پایانی سطح اطمینان این پاسخ‌ها به ۱۰۰ درصد رسیده است. اگر منحنی خوانش واکه‌ای را مبنای ارزیابی الگوی پردازش ادراکی محرک‌های صوتی مدنظر بدانیم، در آن صورت باید گفت شنونده‌های فارسی محرک‌های CiCC را که در آنها واکه /i/

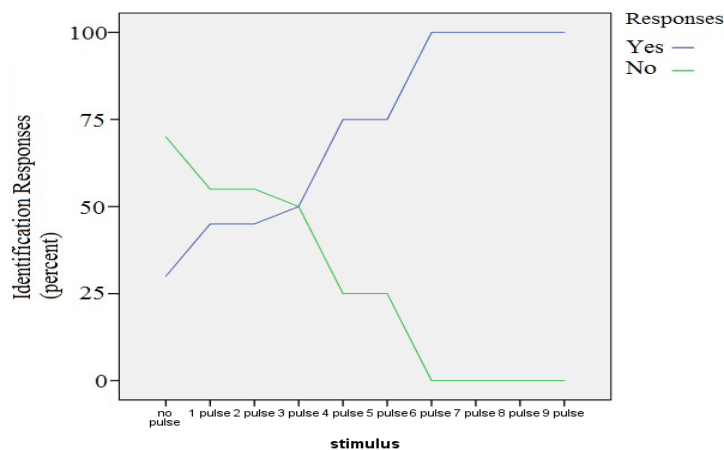
دیرشی به اندازه یک یا تعداد بیشتری تناوب چاکنایی داشته باشد، با سطح اطمینان بسیار مطلوبی به صورت CiCC می‌شنوند، طوری که افزایش تعداد تناوب‌های چاکنایی در سیگنال آوایی CiCC بر فراوانی پاسخ‌های شناسایی مربوطه تأثیرگذار نیست. شنونده‌ها حتی محرک‌هایی را که در آنها /i/ به کلی از سیگنال آوایی حذف شده است، نیز با سطح اطمینان پذیرفتنی به صورت CiCC می‌شنوند.

حال به بررسی شکل (۳) می‌پردازیم که پاسخ‌های شنیداری برای محرک CeCC را نشان می‌دهد. چنانکه مشاهده می‌شود پاسخ‌های خوانش غیرواکه‌ای بیشتر در نواحی ابتدایی پیوستار بر روی گام‌های اول و دوم و پاسخ‌های خوانش واکه‌ای بیشتر در نواحی میانی و بالای پیوستار یعنی بر روی گام‌های چهارم تا دهم توزیع شده‌اند. گام سوم که دقیقاً منطبق بر 2 pulse است نقطه تلاقی ۵۰ درصدی است که ناحیه مبهم پیوستار صوتی مدنظر است. نکته مهم آنکه پاسخ‌های شناسایی در برخی نواحی، تغییرهایی پیوسته و مدرج دارد. از گام دوم تا گام ششم، توزیع پاسخ‌های شناسایی مثبت در منحنی خوانش واکه‌ای به‌طور تدریجی افزایش یافته و به سطح اطمینان ۱۰۰ درصد رسیده است. بعد از گام ششم تا گام پایانی (یعنی تناوب پنجم تا تناوب نهم) سطح اطمینان پاسخ‌ها در همان میزان ۱۰۰ درصد باقی مانده و تغییر محسوسی نداشته است. بر این اساس، در حالی که درصد پاسخ‌های مثبت خوانش واکه‌ای در نواحی میانی پیوستار صوتی CeCC (گام دوم تا ششم) با افزایش تعداد تناوب‌های واکه‌ای /e/ به‌طور تدریجی بیشتر شده است، افزایش تناوب‌های واکه‌ای در نواحی بالای پیوستار تغییرهای محسوسی را در منحنی پاسخ‌های شناسایی ایجاد نکرده است. بر مبنای الگوی کلی توزیع پاسخ‌ها این گونه نتیجه‌گیری می‌شود که شنونده‌ها گام‌های اول و دوم محرک‌های صوتی CeCC را با سطح اطمینان پذیرفتنی (حدود ۸۰ درصد) به صورت CCC و گام‌های ششم تا دهم را با سطح اطمینان بسیار مطلوبی (حدود ۱۰۰ درصد) به صورت CeCC می‌شنوند. در این میان درک آنها از محرک‌های سوم، چهارم و پنجم در کی مدرج است که به صورت تابعی از تعداد تناوب‌های واکه‌ای در محرک‌های صوتی CeCC تغییر می‌کند. به این ترتیب، شنونده‌ها در گام سوم پیوستار CeCC با سطح اطمینانی نزدیک به ۵۰ درصد، در گام چهارم با سطح اطمینانی نزدیک به ۶۵ درصد و در گام پنجم با سطح اطمینانی نزدیک به ۷۵ درصد، واکه /e/ می‌شنوند.



شکل ۳. نمودار خطی پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای برای محرک CeCC
Figure 3. Line graph of identification responses (in percent) to steps of glottal pulses for CeCC.

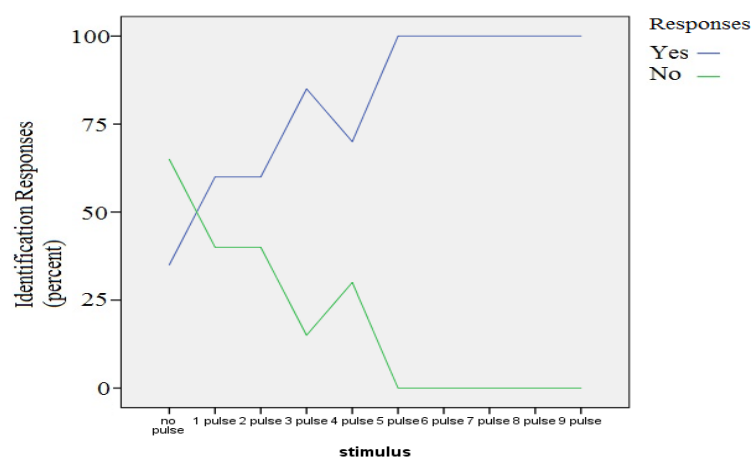
حال شکل (۴) را در نظر می‌گیریم. این شکل پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای مربوط به محرک CæCC را نشان می‌دهد. همان‌طور که مشاهده می‌شود، الگوی کلی توزیع پاسخ‌ها برای CæCC بسیار شبیه CeCC است. پاسخ‌های خوانش غیرواکه‌ای بیشتر در ابتدای پیوستار بر روی گام اول، یعنی محرک No pulse و پاسخ‌های خوانش واکه‌ای بیشتر بر روی گام‌های بالای پیوستار یعنی گام‌های ششم تا دهم توزیع شده‌اند. گام‌های دوم، سوم و چهارم با توجه به نزدیکی به نقطه تلاقی ۵۰ درصدی سه ناحیه مبهم پیوستار صوتی مدنظر هستند که در آنها توزیع پاسخ‌های شناسایی برتری مشخصی را به نفع هیچ کدام از دو خوانش واکه‌ای یا غیرواکه‌ای نشان نمی‌دهد. پاسخ‌ها در منحنی خوانش واکه‌ای بر روی گام‌های پنجم تا دهم توزیع درخور توجهی دارد؛ اما سطح اطمینان پاسخ‌ها برای چهار گام پایانی پیوستار، یعنی گام‌های هفتم تا دهم (تناوب‌های ششم تا نهم) به‌طور درخور توجهی بیشتر از دو گام قبل از آن، یعنی گام‌های پنجم و ششم (تناوب‌های چهار و پنج) است. در حالی که پاسخ‌های شناسایی برای چهار گام پایانی به سطح اطمینان ۱۰۰ درصد نزدیک است، سطح اطمینان پاسخ‌ها برای پالس چهار و پالس پنج حدوداً ۷۵ درصد است. بر این اساس، با افزایش تعداد تناوب‌های چاکنایی /æ/ در سیگنال آوایی محرک CæCC، فراوانی پاسخ‌های شناسایی در نواحی میانی به تدریج افزایش یافته است و به سطح اطمینان ۱۰۰ درصد نزدیک می‌شود. بعد از آن سطح اطمینان پاسخ‌ها بر روی گام‌های پایانی در همین میزان باقی مانده است و تغییری نمی‌کند.



شکل ۴. نمودار خطی پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای برای محرک CæCC
Figure 4. Line graph of identification responses (in percent) to steps of glottal pulses for CæCC.

شکل (۵) پاسخ‌های شناسایی شرکت کنندگان به گام‌های تناوب واکه‌ای در محرک CuCC را نشان می‌دهد. روند کلی توزیع پاسخ‌های شناسایی برای CuCC تا حد زیادی مشابه CeCC و CæCC است. در این شکل مشاهده می‌شود که پاسخ‌های خوانش غیرواکه‌ای در ابتدای پیوستار بر روی گام No pulse توزیع بیشتری دارند و پاسخ‌های خوانش واکه‌ای بر روی پنج گام پایانی پیوستار یعنی گام‌های ششم تا دهم توزیع درخور توجهی دارند. پاسخ‌ها بر روی گام‌های دوم و سوم توزیع مبهمی دارند و برتری مشخصی را به نفع هیچ‌یک از دو خوانش واکه‌ای یا غیرواکه‌ای نشان نمی‌دهند. پاسخ‌های خوانش واکه‌ای

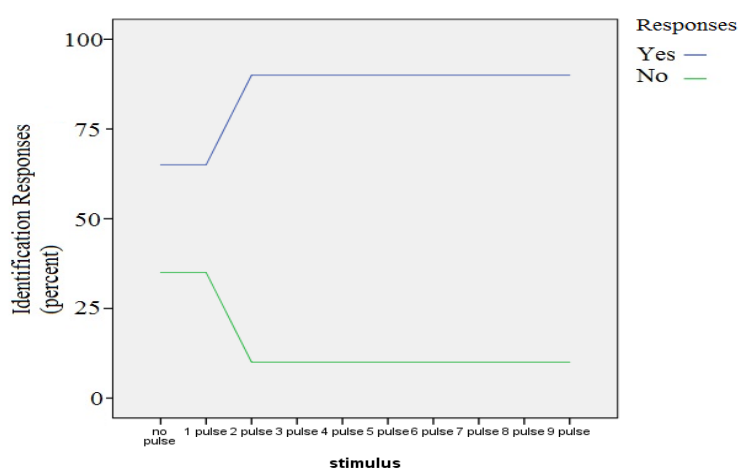
بر روی گام چهارم در مقایسه با گام‌های قبل از آن افزایش یافته است (گام چهارم نسبت به گام پیش از آن حدوداً ۲۰ درصد افزایش داشته است)؛ اما روند تدریجی افزایش پاسخ‌های خوانش واکه‌ای بر روی گام بعد از آن تداوم نداشته است. بر روی گام ششم، منحنی خوانش واکه‌ای به جای سیر صعودی، برخلاف انتظار، سیر نزولی به خود گرفته است طوری که فراوانی پاسخ‌ها بر روی این گام در مقیسه با گام قبل از آن یعنی گام پنجم حدوداً ۱۵ درصد کمتر شده است. بعد از گام ششم، پاسخ‌های خوانش واکه‌ای با شیبی تند دوباره افزایش یافته و به سطح اطمینان ۱۰۰ درصد رسیده است و تا پایان (گام دهم) در همین میزان باقی مانده است.



شکل ۵. نمودار خطی پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای برای محرک CuCC

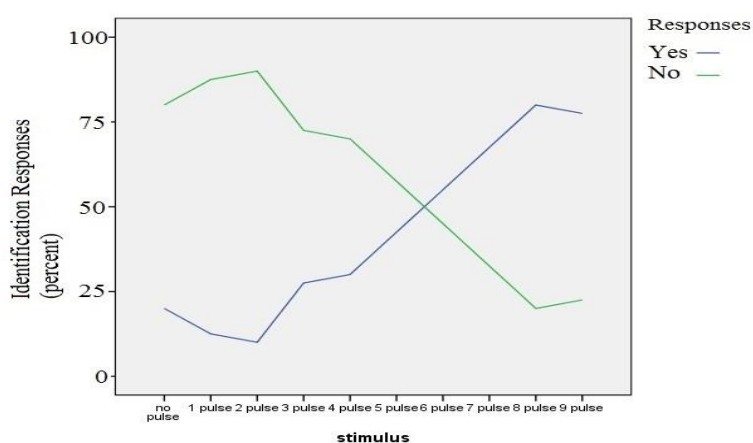
Figure 5: Line graph of identification responses (in percent) to steps of glottal pulses for CuCC.

شکل (۶) پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای /o/ را در محرک CoCC نشان می‌دهد. الگوی کلی توزیع پاسخ‌ها برای CoCC تا حدی زیادی مشابه CiCC است. چنانکه مشاهده می‌شود پاسخ‌های خوانش واکه‌ای حتی در گام‌های ابتدایی پیوستار هم توزیع درخور توجهی دارند به طوری که در گام نخست که واکه به صورت کامل از سیگنال آوایی CoCC جدا شده، سطح اطمینان پاسخ‌های خوانش واکه‌ای در حدود ۶۵ درصد است. این میزان فراوانی برای گام دوم نیز حفظ شده و از گام سوم به بعد سطح اطمینان پاسخ‌ها به حدود ۹۰ درصد رسیده است. آنچه مهم است آنکه روند کلی توزیع پاسخ‌های خوانش واکه‌ای برای CoCC همانند CiCC در هیچ ناحیه‌ای بر روی پیوستار تناوب چاکنایی محرک مربوطه روندی تدریجی و پیوسته ندارد. یعنی این گونه نیست که همانند داده‌های در کی محرک‌های CeCC و CæCC با افزایش گام‌های تناوب چاکنایی به ویژه در نواحی میانی پیوستار به تدریج بر میزان پاسخ‌های خوانش واکه‌ای افزوده شود؛ شنونده‌ها دو گام اول پیوستار CoCC را با سطح اطمینانی نزدیک به ۶۵ درصد و سایر گام‌ها را با سطح اطمینانی نزدیک به ۹۰ درصد به صورت CoCC می‌شنوند.



شکل ۶. نمودار خطی پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای برای محرک CoCC
 Figure 6. Line graph of identification responses (in percent) to steps of glottal pulses for CoCC.

شکل (۷) نتایج پاسخ‌های شناسایی برای محرک CaCC را نشان می‌دهد. شیب منحنی پاسخ‌های شناسایی در این شکل همانند CeCC و CaCC شیئی تدریجی است که در آن با افزایش تعداد تناوب‌های چاکنایی واکه /a/ در CaCC درصد پاسخ‌های خوانش واکه‌ای به تدریج بیشتر شده است. در این شکل، گام هفتم نقطه تلاقی ۵۰ درصدی است که توزیع پاسخ‌ها برتری مشخصی را به نفع دو خوانش واکه‌ای یا غیرواکه‌ای نشان نمی‌دهد. شنونده‌ها سه گام ابتدایی پیوستار را با سطح اطمینانی بین ۸۰ تا ۹۰ درصد به خوانش غیرواکه‌ای و دو گام پایانی پیوستار را با سطح اطمینانی نزدیک به ۸۰ درصد به خوانش واکه‌ای نسبت می‌دهند. بین این دو ناحیه بر روی پیوستار، یعنی گام سوم تا نهم، پاسخ‌های خوانش واکه‌ای با شیئی کند و تدریجی به صورت تابعی از تعداد تناوب‌های چاکنایی /a/ در CaCC افزایش یافته است.



شکل ۷. نمودار خطی پاسخ‌های شناسایی به گام‌های تناوب واکه‌ای برای محرک CaCC
 Figure 7. Line graph of identification responses (in percent) to steps of glottal pulses for CaCC.

۵. بحث و نتیجه‌گیری

پژوهش‌های آزمایشگاهی نشان داده است که شنونده‌ها سازگاری‌هایی را در حین درک شنیداری گفتار غیربومی در مقوله‌های آوایی که با زبان خودشان تطابق ندارد، به وجود می‌آورند. شنونده‌ها اغلب از دانش واجی زبان خود، یعنی الگوها و قواعد آوایی برای انطباق ساخت عناصر آوایی زبان غیربومی با زبان خود استفاده می‌کنند. یافته‌های شنیداری نشان می‌دهد که اگر شنونده با واژه‌ای بی‌معنی مواجه شود که توالی همخوان‌های آن واژه، الگوهای آوایی زبان‌اش را نقض کند، آنگاه یک واکه خیالی در میان چنین توالی‌ای درج می‌کند تا ساخت واجی این توالی منطبق با ساخت آوایی زبان گویشور شود. بر این اساس، درک شنیداری گفتار تا حد زیادی تابع محدودیت‌های واج‌آرایی است. دانش واج‌آرایی زبان نه تنها بر شیوه طبقه‌بندی واحدهای واجی نظارت می‌کند، بلکه منجر به درک آوایی می‌شود که هیچ همبسته صوتی ندارند (آواهای خیالی)؛ بنابراین، شیوه پردازش درکی امواج صوتی گفتار و تبدیل آنها به واحدهای آوایی ناپیوسته و مقوله‌ای (واج‌ها) شیوه‌ای جهان‌شمول نیست، بلکه وابسته به الگوی واج‌آرایی و شیوه تعامل واجی همخوان‌ها و واکه‌ها در زبان مدنظر است. به این ترتیب، شنونده در درک آوایی غیربومی نه تنها آنها را با مقوله‌های آوایی بومی زبان خود همگون می‌سازد، بلکه آوایی را به زنجیره آوایی غیربومی اضافه و یا از آن حذف می‌کند تا صورت آوایی غیربومی با واج‌آرایی زبان مقصد هماهنگ شود.

در این پژوهش به بررسی درک واکه خیالی در خوشه‌های همخوانی آغازی فارسی پرداختیم. با توجه به ممنوعیت خوشه‌های همخوانی آغازی در زبان فارسی، درک واکه خیالی در هجاهای دارای خوشه‌های آغازی C1C2 در رشته‌های آوایی بی‌معنی بررسی شد. برای این منظور، تعداد ۲۰ ناواژه با الگوی هجایی (C1V1C2V2C3) طراحی شد، به طوری که ناواژه‌ها هر ۶ واکه زبان فارسی را در خود داشتند. سپس، هر ناواژه تولیدشده از سوی دو گویشور فارسی زبان (یک زن و یک مرد) تولید و ضبط شدند. سپس، دیرش واکه V1 در ناواژه‌ها در نرم‌افزار پرت بازسازی شد. برای بازسازی دیرش V1 ابتدا سیگنال آوایی کل واکه از زنجیره آوایی محرک مدنظر به کلی جدا شد و سپس به تدریج در گام‌های بعدی بین ۱ تا ۹ تناوب واکه‌ای در ابتدا و انتهای واکه V1 در زنجیره آوایی محرک هدف آزمایش حفظ و مابقی سیگنال آوایی واکه V1 از محرک حذف گردید. در گام بعد، از بیست شرکت‌کننده (ده زن و ده مرد) خواسته شد تا به صورت تصادفی به ناواژه‌های تولیدشده گوش بدهند و پس از شنیدن هر محرک، قضاوت کنند که آیا در ناواژه تولیدشده، واکه‌ای درک می‌کنند یا خیر.

نتایج به دست آمده نشان داد توزیع پاسخ‌های شنیداری برای واکه‌های مختلف در این پژوهش از دو الگوی متفاوت پیروی می‌کند. برای صورت‌های C1iC2V2C3 و C1oC2V2C3 الگوی توزیع پاسخ‌های شنیداری به منحنی فرضی درک واکه‌های خیالی بسیار نزدیک است. شنونده‌ها واکه‌های /i/ و /o/ را در تمامی رشته‌های آوایی C1iC2V2C3 و C1oC2V2C3 می‌شنوند؛ حتی زمانی که این واکه‌ها به کلی از رشته‌های آوایی مربوطه حذف شده باشند. پاسخ‌های خوانش واکه‌ای به تمامی محرک‌های بازسازی شده C1iC2V2C3 و C1oC2V2C3 فراوانی درخور توجهی دارد. فراوانی پاسخ‌ها برای محرک‌های بدون واکه C1iC2V2C3 و C1oC2V2C3 کمتر از محرک‌های دیگر است؛ ولی شیب توزیع پاسخ‌های خوانش واکه‌ای کُند است و روند توزیع پاسخ‌ها بین دو سر ابتدا و انتهای پیوستار دیرش واکه (تناوب‌های واکه‌ای) به آهستگی انجام می‌شود. بر این اساس، یافته‌های مربوط به C1iC2V2C3 و C1oC2V2C3 در این پژوهش با تأیید فرضیه درک

واکه‌های خیالی نشان می‌دهد که درک واکه‌های /i/ و /o/ در محرک‌های بدون واکه C1oC2V2C3 و C1iC2V2C3 ناشی از نبود خوشه‌های همخوانی آغازی در نظام واجی زبان فارسی است.

با این حال، الگوی پاسخ‌های مربوط به واکه‌های /i/ و /o/ برای محرک‌های بازسازی‌شده واکه‌های دیگر تکرار نشد. فراوانی پاسخ‌های خوانش واکه‌ای برای محرک‌های بازسازی‌شده C1eC2V2C3، C1æC2V2C3، C1uC2V2C3 و C1aC2V2C3 بسته به تعداد تناوب واکه‌ای در هر محرک (دیرش واکه) تا حدی زیادی متفاوت بود. پاسخ‌های واکه‌ای به محرک‌های واقع در ابتدای پیوستار به‌طور درخور توجهی کمتر از محرک‌های پایانی بود. فراوانی پاسخ‌ها در نواحی میانی بین دو سر ابتدایی و انتهایی پیوستار نیز به‌صورت تابعی از دیرش واکه تغییر می‌کرد. به بیان دیگر، شنونده‌ها برای تشخیص حضور یا نبود واکه در محرک‌های C1eC2V2C3، C1æC2V2C3، C1uC2V2C3 و C1aC2V2C3 به دیرش واکه حساسیت زیادی نشان می‌دادند. نبود تناوب واکه‌ای یا تعداد اندک تناوب‌های واکه‌ای (دیرش بسیار کوتاه واکه) در محرک‌های صوتی رشته‌های آوایی مدنظر برای شنونده‌ها به معنای نبود واکه و دیرش بلند واکه (وجود حداقل شش تناوب واکه‌ای در آغاز و شش تناوب واکه‌ای در انتها) در این محرک‌ها برای آنها به معنای حضور واکه بود. منحنی خوانش واکه‌ای در نواحی میانی پیوستار تناوب واکه‌ای (با دیرش متوسط واکه‌های /u/، /æ/، /e/ و /a/ روندی تدریجی با سطح اطمینانی بین ۵۰ تا ۸۰ درصد داشت که سطح اطمینان پاسخ‌ها به‌صورت تابعی از دیرش واکه‌ها افزایش می‌یافت. این واقعیت نشان می‌دهد الگوی توزیع پاسخ‌های شنیداری برای C1eC2V2C3، C1æC2V2C3، C1uC2V2C3 و C1aC2V2C3 با آنچه در مطالعه درک واکه‌های خیالی به دست آمده است، مطابقت ندارد؛ زیرا مطابق با فرضیه درک واکه‌های خیالی، درک واکه در خوشه‌های همخوانی وابسته به تعداد تناوب‌های واکه‌ای در محرک‌های صوتی نیست و حتی نبود تناوب واکه‌ای در محرک‌های صوتی مربوطه مانع از درک واکه خیالی نمی‌شود.

بر این اساس، شواهد به‌دست آمده در پژوهش حاضر با فرضیه درک واکه‌های خیالی به‌طور کامل مطابقت ندارد؛ زیرا درک واکه‌های خیالی به‌طور نظام‌مند برای تمامی توالی‌های C1C2 و تمامی واکه‌های فارسی روی نمی‌دهد؛ بنابراین، با توجه به نتایج به‌دست آمده در این پژوهش، مشخص شد فرضیه مطرح‌شده مبنی بر آنکه «فارسی‌زبانان هجاهای حاوی خوشه‌های آغازی C1C2 را در رشته‌های آوایی بی‌معنی با توجه به محدودیت‌های نظام آوایی زبان فارسی، به‌صورت توالی‌های آوایی C1VC2 درک می‌کنند» (یعنی رشته‌های آوایی غیرمجاز C1C2 را با درک یک واکه خیالی ترمیم می‌کنند)، فرضیه معتبری نیست.

کتابنامه

رحمانی، حامد، محمود بی‌جن‌خان، و محمد قاجارگر (۱۳۹۱). ناشنوایی تکیه در گویشوران فارسی. *شنوایی‌سنجی* (۳): ۱۰۲-۱۰۷.

Berent, I., T. Lennertz, J. Jun, M. Moreno, & P. Smolensky (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences* (105), 5321– 5325.

Boersma, P., & D. Weenink (2018). Praat: doing phonetics by computer [Computer program]. Version 4.3.01, Retrieved from <http://www.praat.org/>.

- Davidson, L. (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* 24 (2), 261–286.
- Davidson, L. & J. A. Shaw (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics* 40 (2), 234–248. ISSN 0095-4470. Doi: <http://dx.doi.org/10.1016/j.wocn.2011.11.005>.
- Dupoux, E., K. Kakehi, Y. Hirose, Ch. Pallier, & J. Mehler (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25 (6), 1568–1578. ISSN 1939 1277(ELECTRONIC);0096-1523(PRINT). Doi: 10.1037/0096-1523.25.6.1568.
- Dupoux, E., E. Parlato, S. Frota, Y. Hirose, & Sh. Peperkamp (2011). Where do illusory vowels come from? *Journal of Memory and Language* 64 (3), 199–210.
- Durvasula, K. & J. Kahng (2015). Illusory vowels in perceptual epenthesis: the role of phonological alternations. *Phonology* 32 (03), 385–416. ISSN 1469-8188. Doi: 10.1017/S0952675715000263.
- Durvasula, K. & J. Kahng (2016). The role of phrasal phonology in speech perception: What perceptual epenthesis shows us. *Journal of Phonetics* 54, 15–34. ISSN 0095-4470.
- Durvasula, K., H. H. Huang, S. Uehara, Q. Luo, & Y. H. Lin (2018). Phonology modulates the illusory vowels in perceptual illusions: Evidence from Mandarin and English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9 (1), 1-27.
- Fant, G., & K. T. Högskolan (1957). *Modern instruments and methods for acoustic studies of speech*. Stockholm: Royal Institute of Technology, Division of Telegraphy-Telephony.
- Guevara-Rukoz, A., I. Lin, M. Morii, Y. Minagawa, E. Dupoux, & Sh. Peperkamp (2017). Which epenthetic vowel? Phonetic categories versus acoustic detail in perceptual vowel epenthesis. *The Journal of the Acoustical Society of America* 142 (2), 35-59. EL211–EL217.
- Halle, M. (1998). The stress of English words. *Linguistic Inquiry* 29(4): 539-568.
- Jusczyk, P. W., A. D. Friederici, J. M. Wessels, V. Y. Svenkerud, & A. M. Jusczyk (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of memory and language* 32(3), 402-420.
- Jusczyk, P. W., P. A. Luce, & J. Charles-Luce (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33(5), 630-645.
- Kabak, B. & W. J. Idsardi (2007). Perceptual distortions in the adaptation of English consonant clusters: syllable structure or consonantal contact constraints? *Language and Speech* 50 (1), 23–52.
- Lewis, E., & M. Tatham (2001). Automatic Segmentation of recorded speech into syllables for speech synthesis. *Proceedings of Eurospeech* 1: 1703-1707, edited by J. Hirschberg & T. M. Mitchell. Danmark: Aalborg University.
- Mattingley, W., E. Hume, & K. C. Hall (2015). The influence of preceding consonant on perceptual epenthesis in Japanese. Paper presented at the 18th *International Congress of Phonetic Sciences*. 29-33. Glasgow: University of Edinburgh.
- Monahan, Ph. J., E. Takahashi, Ch. Nakao, & W. Idsardi (2009). Not All Epenthetic Contexts are Equal: Differential Effects in Japanese Illusory. In *Proceedings of the 17th*

- Annual Japanese / Korean Linguistics Conference* edited by I. Shoichi, H. Hoji, P. M. Clancy, & S. O. Sohn. 391–405. Tokyo, Japan: University of Tokyo.
- Moreton, E. (2002). Structural constraints in the perception of English stop–sonorant clusters. *Cognition* 84, 55–71.
- Nearey, T. M. (1992). Context effect in a double-weak theory of speech Perception. *Language and speech* 35 (2): 153-171.
- Nearey, T. M. (1997). Speech Perception as pattern recognition. *Journal of Acoustic society of America* 101 (6), 3241 -3254.
- Peperkamp S., E. Dupoux (2002). A typological study of stress “deafness”. In *Laboratory Phonology 7* edited by C. Gussenhoven, & N. Warner. 40-203. Berlin: Mouton de Gruyter.
- Rahmani, H., T. Rietveld, & C. Gussenhoven (2015). Stress “Deafness” reveals absence of lexical marking of stress or tone in the adult grammar. *PLoS ONE* 10(12), 3-46.
- Stevens, K. N. (1989). On the Quantal Nature of Speech. *Journal of Phonetics* 17:3-46.
- Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). Cambridge, MA: MIT press.
- Wilson, C., & L. Davidson (2013). Bayesian analysis of non-native cluster production. In *Proceedings of NELS 40* (2): 265–278, edited by N. LaCara, Y. Park & Y. Fainleib. Cambridge, MA: MIT Press.