# The influence of phoneme inventory on elicited speech errors in Arabic speakers of English

Faisal M. Aljasser[1], Keonya T. Jackson[2], Michael S. Vitevitch[2] and Joan A. Sereno[2]
[1] Qassim University | [2] University of Kansas

Previous studies have shown that nonnative phonemic contrasts pose perceptual difficulties for L2 learners, but less is known about how these contrasts affect speech production in L2 learners. In the present study, we elicited speech errors in a tongue twister task investigating L1 Arabic speakers producing L2 English words. Two sets of word productions were contrasted: words with phonemic contrasts existing in both L1 Arabic and L2 English (e.g. *tip* vs *dip, sing* vs *zing*) or words with phonemic contrasts existing in English alone (*pit* vs *bit, fat* vs *vat*). Results showed that phonemic contrasts that do not exist in Arabic induced significantly more speech errors in L2 Arabic speakers of English compared to native English speakers than did phonemic contrasts found in both languages. Implications of these findings for representations in L2 learners are discussed.

**Keywords:** speech production, speech errors, tongue twister, Arabic, phoneme inventory

Attaining native-like accuracy in the perception and production of a second language (L2) is the exception rather than the rule for adult L2 learners (e.g. Bongaerts et al 1997; Hyltenstam & Abrahamsson, 2003; Chan & Vitevitch, 2015). Ample cross-linguistic evidence suggests that nonnative *listeners* face difficulty in the recognition of L2 words that have non-native phonemic or phonetic contrasts that can lead to the activation of unintended words (*e.g.*, Broersma 2002; Cutler & Otake 2004; Pallier, Colome & Sebastian-Galles 2001; Sebastian-Galles, Echeverria & Bosch 2005; Weber & Cutler 2004). A classic example of this is the English /l/ – /r/ contrast being mapped to /ɾ/ by Japanese L1 speakers, resulting in misunderstanding of words such as *light* vs. *right* (see for example Aoyama et al. 2004; Cutler & Otake, 2004; Cutler, Weber & Otake 2006). Similarly, English minimal

https://doi.org/10.1075/ml.17014.alj
*The Mental Lexicon* 13:1 (2018), pp. 26–37. issn 1871-1340 | e-issn 1871-1375

pairs containing nonnative vowel contrasts (e.g. *cattle-kettle*) cause difficulties for L1 Dutch listeners (Broersma & Cutler, 2011; Weber & Cutler, 2004). (See also Pallier et al. (2001) for similar results from Spanish-Catalan bilinguals).

Late L2 learners are also typically less likely to *produce* certain L2 sounds in a native-like manner resulting in such speakers having an accent (See Swan & Smith, 2001 for examples of accented pronunciations in L2 English by different first language (L1) speakers; see also Wang et al., 2003). Flege and Port (1981:125) observed that: "…the most important and obvious aspect of foreign-accented speech is sound substitution." This is particularly obvious when producing L2 phonemic contrasts that are not available in the L1 (Flege, 1987). For example, the English /p/ and /v/ are difficult for L1 Arabic speakers because Arabic has /b/ but not /p/, and /f/ but not /v/. This is especially problematic when it results in an unintended word that only differs minimally along the L2 contrast (e.g. *bark* for *park*).

In the case of speech production, increasing evidence points to sub-lexical representations (i.e., linguistic units that are smaller than a word, such as phonemes or biphones) as the representations that cause some of the difficulties experienced by L2 speakers in accurately producing speech in the second language. For example, Gollan and Goldrick (2012) had monolinguals and Spanish-English and Mandarin-English bilinguals produce words and nonwords in a tongue-twister task. They found that the bilinguals made more speech errors than the monolingual speakers for nonword tongue twisters, suggesting that although lexical representations provide support to sub-lexical representations in the case of words, the lack of such lexical representations for nonwords results in more speech errors and makes it difficult to accurately produce sub-lexical representations. They further found that bilingual speakers made significantly more errors than the monolingual speakers for tongue twisters with similar phonemes (*e.g., dirt bus boot dose*) than dissimilar phonemes (*e.g., date fern foot den*), further suggesting that sub-lexical rather than lexical representations might underlie some production difficulties in L2 speakers.

Using a language-switching task Goldrick, Runnqvist and Costa (2014) found additional evidence that sub-lexical representations influence aspects of speech production (see also Balukas & Koops, 2015 and Olson, 2013). They had Spanish-English bilinguals name a picture after being cued to name it in either Spanish or English. When the speakers previously named a picture in Spanish, and then were cued to name the current picture in English, Goldrick et al. found that the English production had a "Spanish accent" (as measured by the voice-onset time of the initial phonological segment of the words). They suggested that in an interactive model of speech production (Dell, 1986) several similar phonological segments will be partially activated and compete to be selected and produced. In the case of bilingual speakers, those competing phonological segments may be the same phoneme as typically produced in the two different languages (i.e., the different voice onset times for /d/ in English and Spanish), which might account for the observed results. Although phonetic analyses show that L2 learners *can* produce non-native phonemic contrasts in a way that approximates the phonetic norms of the L2 (see Flege, 1980, for example, for evidence in the pronunciation of the English /p/-/b/ by L1 Arabic speakers), clearly there are some conditions, such as language switching, that make it difficult to produce such contrasts accurately. Another situation that may make it difficult to produce such contrasts accurately is the tongue twister task, where time pressure and competing speech plans make it difficult to correctly produce words, even in native speakers of a language.

In the current study we wanted to test how sub-lexical representations of L2 phonemic contrasts not available in the L1 affect speech production in L2 learners of English. To do so, we used a tongue twister task (Shattuck-Hufnagel, 1992) to elicit phonological speech errors in English words with phonemic contrasts that are either present or absent in the L1 in order to examine the sub-lexical phonological representations of second language Arabic speakers of English. As a comparison, we also elicited speech errors from native speakers of English.

In the tongue twister task (Shattuck-Hufnagel, 1992) competing speech plans are activated by asking participants to rapidly repeat lists of words that have alternating onsets (e.g. *sing zing zip sip*). Controlling a number of other factors that are known to influence lexical retrieval in speech production, such as neighborhood density (Vitevitch, 2002) and phonotactic probability (Vitevitch, Armbruster & Chu, 2004), an analysis of speech error patterns will provide insight about the quality of the sub-lexical representations that are involved in speech production (*e.g.*, Goldrick et al., 2014; Gollan & Goldrick, 2012). That is, despite being able to accurately produce the L2 contrasts under ideal situations (e.g., Flege, 1980), the weakness of phonological representations of segments that are present in the L2, but are not found in the L1, will be revealed in the tongue twister task. We predicted that Arabic learners of English would make more speech errors in English tongue twisters with a phonemic contrast that is absent in Arabic (*e.g., pit-bit, fan-van*) than in tongue twisters where the phonemic contrast is present in Arabic (*e.g., sing-zing, tip-dip*).

## Method

### Participants

21 native English speakers (16 male and 5 female, mean age 18 years), who were students at the University of Kansas enrolled in an introductory course

in Linguistics participated for course credit. A group of 24 native speakers of Arabic, all male (mean age = 27 years) who were students and visitors at the University of Kansas, volunteered to participate in this experiment without compensation. They were recruited by word of mouth through the Saudi Students Association at the university, and all spoke what is commonly referred to as Gulf Arabic. The native Arabic speakers all studied English for six years at school starting at the age of 12 prior to arriving in the US. On average, they had lived in the US for 3 years. According to their self-reported English proficiency, 17 participants were "advanced" and 6 were "intermediate". All native English and native Arabic participants had normal or corrected-to-normal vision and no history of speech or hearing problems as determined by self-report.

## Materials

The stimuli were 92 English words, consisting of 46 minimal pairs. The initial phonemes in the minimal pairs contained English phonemic contrasts that only differed in voicing: /p/-/b/, /f/-/v/, /t/-/d/, and /s/-/z/. Critically, two of the phonemic contrasts (/p/-/b/ and /f/-/v/) are found in English but not in Arabic. The other two contrasts (/t/-/d/, /s/-/z/) exist in both English and Arabic.

Two lists of 23 four-word tongue twister sets were created from the 46 minimal pairs. In each list, the tongue twister set contained two minimal pairs that differed minimally along the same phonemic contrast (e.g., *pit bit bin pin*). In each tongue twister set (4 words), the initial phoneme of the first word and last word were the same (e.g., /p/), and the initial phoneme of the second word and the third word were the same (e.g., /b/) as in the example *pig big bill pill*.

The two lists (List A and List B) were counterbalanced. They were set up where half of each block of List A was in the voiced-first order (/b/-/p/-/p/-/b/, /v/-/f/-/f/-/v/, /z/-/s/-/s/-/z/, /d/-/t/-/t/-/d/) and the other half was in the voiceless-first order (/p/-/b/-/b/-/p/, /f/-/v/-/v/-/f/, /s/-/z/-/z/-/s/, /t/-/d/-/d/-/t/) The word set that was voiced-first order on List A was voiceless-first order on List B, and vice versa for every set (e.g., *ten den dense tense* in List A would be *den ten tense dense* in List B). Finally, tongue twisters were randomly ordered in each list. Words with a phoneme contrast that exists in the inventory of Arabic were compared to words with a phoneme contrast that does not exist in the inventory of Arabic. The stimuli were equivalent (all $ps > .05$) in their familiarity (Nusbaum, Pisoni & Davis, 1984), word frequency (Kučera & Francis, 1967 and Brysbaert & New, 2009), neighborhood density, and neighborhood frequency (Luce & Pisoni, 1998) (see Table 1). It is important to note that because of the limited number of real words that are minimal pairs that not only contrast in the desired phonemes, but are also comparable on the variables we controlled (as described above), we were not able to create equal numbers of tongue twisters across conditions. Rather than simply reduce the overall number of stimuli and exclude the /*s-z*/ contrast (which is less often examined in studies of Arabic and English perhaps for the reasons we just described) to balance the number of stimuli, we decided to instead include the words that we could, but use percentages in our scoring to adjust for the different number of actual stimuli across the conditions.

**Table 1.** Means (SDs in parentheses) for the stimuli used in the experiment

| Variable name | Distinction present in Arabic ($n = 36$) | Distinction absent in Arabic ($n = 56$) |
|---|---|---|
| Familiarity (7-point scale) | 6.41 (1.29) | 6.76 (.42) |
| Word frequency$_{KF}$ ($\log_{10}$) | 1.19 (1.10) | 1.23 (.86) |
| Word frequency$_{SUBTLEXus}$ ($\log_{10}$) | 2.76 (1.04) | 2.94 (.90) |
| Neighborhood density | 22.83 (7.99) | 23.11 (8.08) |
| Neighborhood frequency ($\log_{10}$) | 1.06 (.29) | 1.10 (.26) |

## Procedure

Subjects were seated at a comfortable distance from a computer screen in a quiet room. For the Arabic speakers, it was necessary to establish, prior to the presentation of the experimental lists, that they could make a distinction in their pronunciation between the phonemic contrasts that are not available in Arabic (i.e. *p-b* and *f-v*). Therefore, all of the stimuli were first presented visually on the computer screen to the Arabic speakers one word at a time. The participant was instructed to read each word aloud. The experimenter listened to the participant's pronunciation to verify that a distinction was made between /p/-/b/ and /f/-/v/. Only one Arabic speaker failed to consistently make a distinction in pronouncing the words in isolation. This participant was excluded from the study.

Each participant was then assigned to one of the lists (*i.e.*, List A or List B). To minimize the possibility that memory demands might affect the productions, each tongue twister set was presented to participants visually on the computer screen in black Calibri 72-point with lower case characters on a white background using Microsoft PowerPoint and remained visible on the computer screen until all of the required productions had been made. The participants were instructed to repeat each set of four words five times as quickly as possible regardless of whether they made any errors or not. The first slide contained a practice set of words that did not contain any of the critical contrasts (*meet neat nice mice*) just to familiarize the participants with the task. After the practice set, participants were asked if they had any questions. They were then instructed to use the "Enter" key to start the first experimental set and to use the same key to advance to the next set of words. Responses were audio-recorded for later analysis.

## Results

The recorded responses were examined for speech errors by two raters (two of the authors). Intra-rater reliability was high (91%).

In the present study, we used the conventions established in Baars et al. (1975) and also used in Vitevitch (2002) to score responses: because errors on words that occur later in the tongue twister are conditioned by errors that occur earlier, and because a few participants stop producing the remaining words in the tongue twister once they detected they made an error, we did not count the total number of errors that occurred in each repetition. Rather, we simply counted whether the four words in the tongue twister were produced correctly or if one or more speech error occurred during the repetition. Furthermore, although participants were instructed to repeat each set of words five times, they occasionally would lose count and produce more than or fewer than five repetitions. When participants produced more than five repetitions, we only scored the first five repetitions. To accommodate participants producing fewer than five repetitions (2.1% of the data), we computed proportions of speech errors to repetitions actually produced.

The repetition of a tongue twister was scored as a speech error if the utterance contained either a complete or an incomplete reversal of the initial phonemes of the words in the set. Consider the tongue twister *fan van vase face*. The following examples would be counted as a speech error: complete reversals such as *van fan vase face*, or *fan van face vase*, or *van fan face vase*; or incomplete reversals such as *fan fan vase face*, or *fan fan face vase*, or *fan fan face face*. Notice that the total number of errors made in each of the sample utterances differs. However, in accord with the conventions described above (Baars et al., 1975; Vitevitch, 2002), we simply scored the tongue twister as being produced correctly or not (i.e., we did not count the total number of errors made in each repetition of the tongue twister). Thus, a participant could make anywhere from 0 to 5 "errors" for a given tongue twister. Again, we used proportions to accommodate participants who made fewer than 5 repetitions of a tongue twister and to accommodate the different number of tongue twisters across conditions.

Responses that were not correct but not counted as speech errors included errors not involving initial consonants, producing a word not in the word list, failures to repeat any of the words in the word list, and errors in which participants misread or mispronounced a word (e.g., saying *tone* instead of *ton*). These types of errors accounted for 7% of all of the responses, and were not analyzed any further since they were not of interest.

A mixed ANOVA was used to analyze the proportion of speech errors made with Speaker (English vs. Arabic speaker) as a between participants variable, and Phoneme contrast (present vs. absent in Arabic) as a within participants variable. There was no difference between lists (A vs. B), so we collapsed across this variable for subsequent analyses.

There was a significant main effect for Speaker, with the native Arabic speakers making more speech errors in their L2 of English ($mean = 32\%$) than the native speakers of English ($mean = 11\%$; $F(1, 43) = 38.37$, $p < .001$). There was also a significant main effect for Phoneme Contrast (i.e., between words with a phoneme contrast either found in Arabic or not found in Arabic) ($F(1, 43) = 56.42$, $p < .001$), such that more errors were made on words with phoneme contrasts not found in Arabic ($mean = 29\%$) than on words with phoneme contrasts that are found in Arabic ($mean = 13.0\%$). Although there were more tongue twisters with phoneme contrasts not found in Arabic than with phoneme contrasts found in Arabic, recall that we are reporting percentages (not overall numbers of errors), thereby accommodating the different number of stimuli in each condition. Furthermore, having more tongue twisters with phoneme contrasts not found in Arabic could have increased the phonotactic probabilities of those phoneme segments in the context of the experiment (e.g., Dell et al., 2000), which should have resulted in fewer speech errors in those higher-probability segments (Vitevitch, Armbrüster & Chu, 2004), not more as was observed.

These main effects should be considered in the context of a significant interaction between Speaker and Phoneme Contrast ($F(1, 43) = 37.83$, $p < .001$), such that Arabic speakers made many more errors on words with a phoneme distinction not found in Arabic ($mean = 46\%$) compared to native English speakers on the same types of words (12%) (Tukey's test $t(43) = 6.89$, $p < .001$). In the case of words with a phoneme distinction found in Arabic, Arabic speakers made only slightly more errors ($mean = 17\%$) compared to native English speakers ($mean = 9\%$), however this difference was still statistically significant (Tukey's test $t(43) = 2.81$, $p < .01$).

## Discussion

The results of the present tongue twister experiment showed a main effect of Speaker. That is, native Arabic speakers learning English as a second language made more speech errors on English words than did the native English speakers. This finding replicates Gollan and Goldrick (2012) who reported that Spanish-English and Mandarin-English bilinguals make more speech errors than the English monolinguals. Although the native Arabic speakers in this study were all quite proficient in English, it is perhaps to be expected that they would make more speech errors in their L2 compared to native speakers of the same language especially given the speech errors observed by Gollan and Goldrick (2012) in participants that had acquired English earlier in life and were more proficient than the participants in the present study.

We also found a main effect of Phoneme Contrast such that fewer speech errors were made on English words when they contained a phoneme contrast that was found in both Arabic and English compared to English words with a phoneme contrast that was found in English but not in Arabic. This finding is also perhaps to be expected given that the phonemes common to English and Arabic (*e.g.* coronal stops /t/ and /d/ or fricatives /s/ and /z/) are also frequently found cross-linguistically (Greenberg, 1966) and therefore are considered less marked (Jakobson, 1971), whereas the phonemes found in English but not Arabic (*e.g.*, labials /p/ and /v/) are less frequently found cross-linguistically and therefore are considered more marked (see place of articulation markedness hierarchy (Lombardi, 2002)).

What is more interesting is the significant interaction between Speaker and Phoneme Contrast. The native Arabic speakers, who were intermediate or advanced speakers of the second language (English), and who were able to accurately produce the phoneme distinctions when there was no time-pressure or competing speech plans, nevertheless, when producing these phonemic distinctions in the tongue twister task, made almost 4 times as many speech errors on English words that contained phoneme contrasts not present in Arabic compared to the native English speakers in the same phoneme condition. For the words containing phoneme contrasts present in both Arabic and English, the native Arabic speakers made only about twice as many speech errors on those English words compared to the native English speakers in the same phoneme condition.

The present findings provide additional evidence of how sub-lexical representations influence speech production, further extending the findings from Gollan & Goldrick (2012), Goldrick et al., (2014) and Li, Goldrick & Gollan (2017) who previously explored the role of sub-lexical representations in speech production. Li et al. (2017) suggested that the increased number of speech errors observed in their bilingual speakers was in part due to a frequency-lag in the sub-lexical representations. That is, the relative frequency that L2 or bilingual speakers experience certain sub-lexical representations compared to other sub-lexical representations may partially underlie their difficulty in accessing those representations.

We further suggest that the frequency-lag of phonemes not found in the native language may result in representations that are strong enough to enable intermediate or advanced speakers of a second language to perceive and produce these L2 distinctions under most conditions (e.g., Flege, 1980). However, when placed under pressure to speak quickly and to select among competing speech plans (as elicited by the tongue twister task), the weakness of the phonemic representations not found in the native language can be observed even in advanced speakers of an L2.

The present findings suggest that some of the problems L2 learners experience in perception and production may not only be due to misperceptions, but may also be due to still nascent representations of phonemes not found in the native language, despite many years of practice with the L2. Such nascent representations could be modeled as significantly fewer exemplars of certain sub-lexical representations stored in the lexicon (*cf.*, Goldinger, 1996 and Pallier et al., 2001). In an abstractionist account of the lexicon, such as a connectionist model of speech production (e.g., Dell, 1986), such nascent representations might be represented as weaker connection weights between nodes. Those connection weights could become stronger with increased experience or exposure over time (e.g., Warker & Dell 2006). Additional studies of speech errors in bilingual and L2 speakers might be useful in discriminating between the exemplar and abstractionist accounts of the mental lexicon (e.g., Pallier et al., 2001).

## Acknowledgements

We wish to thank the Council for International Exchange of Scholars for funding FMA through the Fulbright Scholars Program while he was at the University of Kansas.

## References

Aoyama, K., Flege, J., Guion, S., Akahane-Yamada, R., and Yamada, T., (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233–50. https://doi.org/10.1016/S0095-4470(03)00036-6

Baars, B. J., Motley, M. T. & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 14, 382–391. https://doi.org/10.1016/S0022-5371(75)80017-X

Balukas, C. & Koops, C. (2015). Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*, 19, 423–443. https://doi.org/10.1177/1367006913516035

Bongaerts, T., van Summeren, C., Planken, B., and Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465. https://doi.org/10.1017/S0272263197004026

Broersma, M. (2002). Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. In: Hansen, J. and Pellom, B. L. (eds) *Seventh International Conference on Spoken Language Processing*. Denver, CO: Center for Spoken Language.

Broersma, M., and Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology*, 64, 74–95. https://doi.org/10.1080/17470218.2010.499174

Brysbaert, M. & New, B. (2009) Moving beyond Kucera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behavior Research Methods*, 41, 977–990 https://doi.org/10.3758/BRM.41.4.977

Chan, K. Y. and Vitevitch, M. S. (2015). The influence of neighborhood density on the recognition of Spanish-accented words. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 69–85.

Cutler, A., Weber, A., and Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, 34, 269–284. https://doi.org/10.1016/j.wocn.2005.06.002

Cutler, A., and Otake, T. (2004). Pseudo-homophony in non-native listening. *Journal of the Acoustical Society of America*, 115, 2392. https://doi.org/10.1121/1.4780547

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Reed, K. D., Adams, D. R., and Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1355–1367.

Flege, J. E. (1980). Phonetic Approximation in Second Language Acquisition. *Language Learning*, 30, 117–134. https://doi.org/10.1111/j.1467-1770.1980.tb00154.x

Flege, James. E. (1987). The production of 'New' and 'Similar' phones in a foreign language: Evidence for the effect of equivalence classification." *Journal of Phonetics*, 15, 47–65.

Flege, J. E., and Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24, 125–146.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166–1183.

Goldrick, M., Runnqvist, E. and Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science*, 25, 1031–1036. https://doi.org/10.1177/0956797613520014

Gollan, T. H. and Goldrick, M. (2012). Does bilingualism twist your tongue? *Cognition*, 125, 491–497. https://doi.org/10.1016/j.cognition.2012.08.002

Greenberg, J. (1966). *Language Universals*. The Hague: Mouton.

Hyltenstam, K., and Abrahamsson, N. (2003). Maturational constraints in SLA. In Doughty and Long (Eds.), *The handbook of second language acquisition*. Rowley, MA: Blackwell. https://doi.org/10.1002/9780470756492.ch17
Jakobson, R. (1971). *Selected Writings Volume 1: Phonological Studies*. The Hague: Mouton.
Kučera, H., and Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
Li, C., Goldrick, M. & Gollan, T. H. (2017). Bilinguals' Twisted Tongues: Frequency Lag or Interference? *Memory & Cognition*, 45, 600–610.
Lombardi, L. (2002). Coronal epenthesis and markedness. *Phonology*, 19, 219–251. https://doi.org/10.1017/S0952675702004323
Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1–36. https://doi.org/10.1097/00003446-199802000-00001
Nusbaum, H. C., Pisoni, D. B., and Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report*, 10, 357–376.
Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41, 407–420. https://doi.org/10.1016/j.wocn.2013.07.005
Pallier, C., Colome´, A., and Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, 12, 445–449. https://doi.org/10.1111/1467-9280.00383
Sebastián-Gallés, N., Echeverría, S., and Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, 52, 240–255. https://doi.org/10.1016/j.jml.2004.11.001
Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition*, 42, 213–259. https://doi.org/10.1016/0010-0277(92)90044-I
Swan, M., and Smith, B. (2001). Learner English. *A teacher's guide to interference and other problems*. https://doi.org/10.1017/CBO9780511667121
Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 735.
Vitevitch, M. S., Armbrüster, J., and Chu, S. (2004). Sublexical and lexical representations in speech production: effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 514.
Wang, Y., Jongman, A., and Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, 113, 1033–1044. https://doi.org/10.1121/1.1531176
Warker, J. A. & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning Memory and Cognition*, 32, 387–398.
Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50, 1–25. https://doi.org/10.1016/S0749-596X(03)00105-0

## Appendix

| Phoneme distinction present in Arabic | Phoneme distinction absent in Arabic |
|---|---|
| /s/ – /z/ | /f/ – /v/ |
| seal zeal zed said | fan van vase face |
| sing zing zip sip | fast vast vat fat |
| | file vile vine fine |
| | fail vail vein fain |
| | fear veer veal feel |
| | fend vend vault fault |
| /t/ – /d/ | /p/ – /b/ |
| tub dub duck tuck | pack back bad pad |
| ten den dense tense | pounce bounce bound pound |
| town down doubt tout | pig big bill pill |
| tusk dusk done ton | pile bile bike pike |
| tip dip din tin | pun bun bus pus |
| tine dine die tie | pet bet best pest |
| two dew dune tune | pan ban bass pass |
| | pit bit bin pin |

## Address for correspondence

Faisal M. Aljasser
Department of Psychology
University of Kansas
1415 Jayhawk Blvd.
Lawrence, KS 66045
USA

faisalmj24@hotmail.com

## Co-author information

Keonya T. Jackson
Department of Psychology
University of Kansas
kejackson94@gmail.com

Michael S. Vitevitch
Department of Psychology
University of Kansas
mvitevit@ku.edu

Joan A. Sereno
Department of Linguistics
University of Kansas
sereno@ku.edu