

Research Article

THE PERCEPTION OF ARABIC VOWEL DURATION BY L1 HEBREW SPEAKERS

CAN A SHORT TRAINING REMOLD THE EFFECT OF THE NATIVE PHONOLOGICAL SYSTEM?

*Yael Zaltz**

Tel Aviv University

Osnat Segal 

Tel Aviv University

Abstract

The acquisition of a second language (L2) may be challenging in adulthood, as the phonological system of the native language (L1) can sometimes limit the perception of phonological contrasts in L2. The present study aimed to (a) examine the influence of an L1 (Hebrew) that lacks a phonemic contrast for vowel length on the ability to discriminate between short and long vowels in L2 (Arabic); and (b) assess the effect of a short training on the participants' discrimination performance. A total of 60 participants, 20 native Arabic speakers and 40 native Hebrew speakers, were tested using the ABX procedure in two sessions that were 10 days apart. A single training session was provided for half of the Hebrew speakers ($n = 20$) approximately 2–3 days after the first (pretraining) testing session. The results indicated that the trained Hebrew participants' discrimination levels (measured by accuracy and reaction times) were above chance level but were nevertheless lower in comparison to the Arabic speakers. However, a short training session was sufficient to yield a nativelike performance that generalized to untrained nonwords. These findings support the theoretical models that predict a reserved ability to acquire new phonetic/phonological cues in L2 and have important practical implications for the process of learning a new phonological system in adulthood.

The authors thank all the participants for agreeing to take part in the present study and the undergraduate students from the Department of Communication Disorders at Tel Aviv University for assisting in the data collection.

*Correspondence concerning this article should be addressed to Yael Zaltz, Department of Communication Disorders, The Stanley Steyer School of Health Professions, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. E-mail: yaelzaltz@gmail.com

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The ability to speak several languages is becoming ever more important in the modern global community. In many cases, adults are required to learn a second language after their native language has already molded their phonetic/phonological perceptual abilities. To acquire a new language (L2) in adulthood, one should learn to distinguish between sounds that are phonemes, that is, differentiate between words in the new phonological system. In many cases, there is some overlap between the phonological systems of the native language (L1) and L2. Thus, the L2-learner is required to focus on the “new” phonemic contrasts, adjusting their phonological representations according to the structure of the new phonological system. Research on the perception of nonnative contrasts suggests that the phonological system of the native language affects the nonnative perception of consonants (Escudero & Williams, 2011; Tyler et al., 2014). However, L1 may affect one’s perception of nonnative vowel contrasts differently, as vowels have different acoustic characteristics and less distinct categorical perception compared to consonants (Fry et al., 1962; Kronrod et al., 2016). The existing research has largely focused on native speakers of Romance (e.g., Italian) and German (e.g., English) languages, usually with stimuli in Japanese, German, or Finnish, to explore the effect of L1 on the discrimination of vowel length (e.g., Flege & Mackay, 2004; Hisagi et al., 2010, 2015; Nenonen et al., 2005). Furthermore, only scarce data exists on the influence of a short training on vowel discrimination in L2 (e.g., Wang & Munro, 1999). Thus, the purpose of the present study was to examine native Hebrew speakers’ discrimination of short and long vowel contrasts of spoken Arabic before and after a single training session. Their vowel contrast discrimination was compared to that of native Arabic speakers.

THE INFLUENCE OF L1 ON THE DISCRIMINATION OF CONSONANTS IN L2

During the first months of their lives, infants are able to discriminate phonetic contrasts regardless of whether these contrasts are phonemes or allophones in their native language (e.g., Eimas et al., 1971; Lasky et al., 1975; Segal et al., 2016; Trehub, 1976). However, toward the end of the first year of life, infants generally show reduced discrimination for many nonnative contrasts (see Werker & Tees, 1999 for a review). Studies have demonstrated that English-learning infants aged between 6 and 8 months were able to discriminate nonnative contrasts in Hindi, Salish (Thompson/Nthlakampx), and Spanish, whereas English adults could not discriminate the same nonnative contrasts. In addition, 10- to 12-month-old English infants already exhibited a decline in their perception compared to younger infants (Eilers et al., 1982; Werker et al., 1981; Werker & Lalonde, 1988; Werker & Tees, 1983, 1984). Similarly, 6–8-month-old Japanese-learning infants could discriminate the English /r/ and /l/ contrast, whereas Japanese adults could not. Furthermore, Japanese-learning infants showed a decline in discrimination toward the end of their first year of life (Kuhl et al., 2006; Tsushima et al., 1994).

However, individuals do not lose their ability to discriminate all nonnative contrasts as they age. For example, it has been demonstrated that English adult listeners can discriminate between Zulu clicks despite a lack of experience with such clicks (e.g., Best et al., 1988). Moreover, there is evidence that there is a variation in discrimination performance for different nonnative contrasts (e.g., Best et al., 2001; Tsao et al., 2006). For example,

adult English speakers were able to discriminate Zulu lateral fricatives (voiceless vs. voiced /ʃ/-/ʒ/) better than Zulu velar stops (voiceless aspirated vs. glottalized /k^h/-/k'/), and the latter were discriminated better than bilabial stops (plosive vs. implosive voiced bilabial stops /b/-/ɓ/) (Best et al., 2001).

The perceptual assimilation model (PAM) explains the variability in the discrimination of nonnative contrasts, especially consonants. According to the PAM, a decline in the discrimination of a particular nonnative contrast with age depends on whether or not the contrast is perceived as speech (Best et al., 1988, Best et al., 1995), whether it is perceived as similar to one or two categories of speech sounds in the native language (Best et al., 2001), and to what extent each of the two members of the contrast resemble the prototype of the native category (i.e., the level of “category goodness”) (Best et al., 1995). If the members of the nonnative contrast do not resemble any phonological native category or are not heard as speech, then the contrast is nonassailable and can be discriminated accurately even in adulthood. This is exemplified by native English speakers being able to discriminate Zulu clicks (Best et al., 1988). Similarly, when the two members of the contrast are assimilated to two different native phonological categories, which is also known as two-category assimilation, discrimination remains accurate. The discrimination of Zulu lateral fricatives /ʃ/-/ʒ/ by English adult speakers is an example of this condition (Best et al., 1988). However, discrimination is less accurate when both members of the contrast are assimilated to the same phonological category, which is referred to as a single-category (SC) assimilation. Two examples include the discrimination between Zulu bilabial stops by English speakers /b/-/ɓ/) (Best et al., 1988) and the discrimination of the English /r/ and /l/ contrast by Japanese adults (Kuhl et al., 2006; Tsushima et al., 1994).

Nevertheless, when there is assimilation to one category in the native language, there might be differences in category-goodness between the two members of the contrast, resulting in quite good discrimination. This is exemplified by the /k^h/-/k'/ discrimination by English adult speakers (Best et al., 2001). Other models, such as Flege’s speech learning model (SLM; Flege, 1995; cf. Guion et al., 2000) and Kuhl’s Native Language Magnet model (NLM; Grieser & Kuhl, 1989; Kuhl et al., 2008) also suggest that the discrimination of nonnative contrasts is influenced by previous knowledge in L1. However, the SLM primarily addresses the production acquisition of L2, and both the NLM and the SLM have no direct predictions about discrimination of nonnative contrasts.

While there has been a substantial amount of research conducted on the discrimination of nonnative consonants (e.g., Best et al., 2001; Best et al., 1988, 1995; Kuhl et al., 2006; Tsao et al., 2006), less is known about the discrimination of nonnative vowel contrasts (e.g., Flege et al., 1997; Polka & Bohn, 1996; Polka & Werker, 1994). There is also little research in regard to whether the discrimination of nonnative vowels follows the same predictions suggested by the PAM (Best & Tyler, 2007; Guion et al., 2000; Levy & Strange, 2008).

THE INFLUENCE OF L1 ON THE DISCRIMINATION OF VOWEL CONTRASTS

Acoustically, vowels differ from the majority of consonants. Vowels have higher acoustic intensity and usually longer duration compared to consonants. Vowels are primarily distinguished from each other in the first three formant frequencies (Ladefoged, 2001). The acoustic differences between vowels and consonants influence how they are

perceived. Studies on classic categorical perception suggest that labeling functions are less steep for vowels than for consonants. Thus, the boundaries between categories may be less sharp and the within-category discrimination may be better compared to consonants (Altman et al., 2014; Fry et al., 1962; Kronrod et al. 2016). These different acoustic characteristics may raise some concerns in regard to how applicable the PAM is for understanding perceptions of nonnative vowel contrasts.

Several studies have reported the influence of the native language on the discrimination of nonnative vowel contrasts to emerge earlier in development compared to the influence of the native language on consonants (e.g., Kuhl et al., 1992). For example, English-learning infants showed a decline in discrimination for several nonnative vowel contrasts before reaching 8 months old, including the Norwegian /y/-/ʉ/ (Best et al., 1997) and the German /u-y/ and /U-Y/ (Polka & Werker, 1994). Additional research demonstrated that monolingual Spanish infants were no longer able to discriminate the Catalan-specific vowel contrast /e/ versus /E/ between 4 and 8 months of age (Bosch & Sebastián-Gallés, 2003).

The majority of studies on adults' nonnative vowel perceptions assessed how an individual discriminates between vowels by comparing their F1/F2 ratios. Such research has indicated that adults' discrimination performance does follow the PAM's suggested assimilation pattern. Specifically, the existing research has found evidence that when two members of a vowel contrast are assimilated based on F1/F2 values to two different native categories, the discrimination will be better than if they had been assimilated to an SC in the native language. There is evidence in support of this finding for English vowel contrasts discriminated by Italian speakers (Flege & Mackay, 2004), for Danish vowels discriminated by Australian English speakers (Faris et al., 2018), for French vowel contrasts discriminated by English speakers (Levy, 2009; Levy & Strange, 2008), and for Norwegian, Thai, and French vowel contrasts discriminated by American English speakers (Tyler et al., 2014).

However, existing studies have provided conflicting evidence in situations when the nonnative vowel contrasts were mainly based on duration. A number of studies have suggested that listeners mainly identify/categorize vowels based on psychoacoustic abilities, thus raising the possibility that nonnative speakers may efficiently perceive the vowel length contrast regardless of their native phonetic/phonological categories (e.g., Bohn, 1995; Cebrian, 2006; Flege et al., 1997; García Lecumberri & Cenoz Inagui, 1997). Researchers have used this notion to explain the findings that 18-month-old English speakers can discriminate short and long vowels even though vowel length is not phonemically contrastive in English (Mugitani et al., 2009). Altmann et al. (2012) also found support for this notion by demonstrating that Italian listeners with no knowledge of German were able to discriminate German vowel length contrasts in a same-different (AX) task as accurately as native listeners.

Other studies, however, have provided evidence in favor of the native language's influence on the discrimination of vowel length contrast, which supports the PAM-suggested two-category (e.g., McAllister et al., 2002) and SC assimilation patterns (e.g., Nenonen et al., 2005; Tsukada et al., 2014). In one study, Estonian, Spanish, and American English speakers who had lived in Sweden for at least 10 years were presented with Swedish vowel contrasts that differed in duration. The subjects were asked to listen to real and nonsense words, and their task was to decide if the production of the "word"

was a correct or an incorrect production of the real word. The participants' discrimination performances were in accordance with the duration cues in their L1s. However, the Estonian speakers performed similarly to the control group, which can be explained by their L1's contrastive vowel length (as predicted by the two-category assimilation of the PAM) (Best et al., 1995; McAllister et al., 2002). Tsukada's (2011) study also provided evidence in support of the two-category assimilation. In this study, Arabic and Japanese listeners (whose native languages include duration categories for vowels) were more accurate in discriminating vowel length contrasts than Persian listeners (whose native language does not include length categories) (Tsukada, 2011).

A number of studies have found that L1 has a negative influence on the discrimination of vowel length, supporting the SC assimilation suggested by the PAM (e.g., Best et al., 1995). In one study, native adult speakers of Thai, Italian, and American English with minimal experience with Japanese performed better than random chance but did not perform as well as the Japanese speakers. However, native speakers of Thai, Italian, and English have some previous experience with the perception of duration changes in vowels. English includes tense and lax vowels that differ in duration (Leung et al., 2016), Italian includes duration changes in stressed vowels (D'Imperio & Rosenthal, 1999), and Thai includes different phonemic categories for vowel length. This may explain the previously mentioned chance performance with Japanese vowels. The performance of the Thai, Italian, and English speakers was nevertheless poorer than that of the Japanese native speakers, reflecting the influence of L1 on discrimination (Tsukada et al., 2014). Evidence for the influence of the native language on the perception of nonnative vowel contrasts was also found for Russian native listeners, as they had reduced discrimination between long and short vowels in L2 (Finnish) when the vowels resembled one phonemic category in Russian (Nenonen et al., 2005). This was found although duration differences exist in Russian vowels in the context of lexical stress (Chrabaszcz et al., 2014).

A series of electrophysiological and behavioral studies support not only SC assimilation but also discrimination based on category goodness. In these studies, American English speakers discriminated the Japanese contrast (e.g., *tado* vs. *taado*) well above chance but performed poorly in comparison to Japanese listeners (Hisagi et al., 2010, 2015; Hisagi et al., 2014). These findings suggest that nonsimilar vowel exemplars of the same L1 category can be discriminated based on differences in their similarity to the common exemplar of the category, supporting a within category discrimination (Hisagi et al., 2010, 2015)

TRAINING OF NONNATIVE VOWEL CONTRASTS

According to the Automatic Selective Perception (ASP) model, speech processing in L1 is based on automatic selective perceptual routines that do not require special attention. However, speech processing of nonnative contrasts may require the allocation of attentional resources (e.g., Strange & Shafer, 2008). In this context, one interesting question is whether and how a native language's influence on vowel discrimination can be changed through an auditory training that promotes attention to specific speech contrasts (e.g., Kabakoff et al., 2020). Studies that assessed the influence of training on L2 learners' perception of speech contrasts were conducted mainly with consonants. Evidence

suggests that intensive training using consonants with allophonic variation significantly improved L2 learners' perceptions and generalization of nonnative consonants (e.g., Bradlow & Bent, 2008; Hazan et al., 2005; Iverson et al., 2005; Jamieson & Morosan, 1986; Pruitt et al., 2006; Wang et al., 1999).

A few studies have also shown that intensive training can improve the perception of nonnative vowel contrasts (e.g., Kabakoff et al., 2020; Lengeris & Hazan, 2010; Rato, 2014), including length (e.g., a generalization of American English vowels that differ in length by both native Japanese and Korean speakers; Nishi & Kewley-Port, 2007, 2008). To the best of our knowledge, there is only one existing study that has examined the effect of a short training on the perception of duration contrast (Wang & Munro, 1999). This study showed that it is possible for Japanese listeners to correctly identify 95% of English vowels that differed in duration after a single training session. It was not clear, however, whether this improvement resulted in a lasting change in perception or whether a short training may similarly facilitate vowel length discrimination in languages with different phonological system such as Semitic languages.

THE PRESENT STUDY

The existing data on L1's influence on the perception of vowel length and its susceptibility to a short training is limited and inconsistent, which motivated us to explore how Hebrew (as a native language) influences the discrimination of long versus short vowels in spoken Arabic. In Arabic, the duration differences between vowels are phonemic, whereas duration in Hebrew does not distinguish between vowels (or consonants) and has no phonemic value (Amir et al., 2014). There are only five vowels in Hebrew (/a, i, e, o u/) that lack a phonemic distinction between short and long vowels. The duration of vowels can vary between 70 and 100 milliseconds depending on whether or not the syllable is stressed and whether the vowel is in a bisyllable or one-syllable word (Most et al., 2000; Silber-Varod et al., 2016). In contrast, the Arabic dialect spoken in the rural parts of central Israel (also known as the "Triangular" [Muthallath] region) includes 10 vowels, that is, the short and long vowels (/a, e, i, o, u/) (Amir et al., 2014). The Arabic short and long vowels have durations of 58–74 milliseconds and 114–136 milliseconds, respectively (Amir et al., 2014). Thus, the duration of Arabic vowels is shorter or longer in comparison to the duration of Hebrew's prototypical vowels. Although long and short Arabic vowels may assimilate to one phonemic vowel category in Hebrew, because their duration is not similar to that of the prototypical Hebrew vowels, within category discrimination may be possible based on acoustic representations (Best et al., 2001; Hisagi et al., 2010, 2015; Hisagi et al., 2014).

In this context, the present study's purpose was twofold: (a) to assess the effect of a native language (Hebrew) that has no phonemic duration contrast on the ability to discriminate between long and short vowels in a nonnative (L2) language (Arabic); and (b) to assess the influence of a short training on one's ability to improve their discrimination of nonnative vowel length. We assumed that the Hebrew speakers, for whom the long and short vowels assimilate to one phonemic category would show reduced performance in discrimination of vowel length compared to the Arabic speakers, for whom the two vowels represent two phonemic categories. We also assumed, however, that their performance would be above chance because they are exposed to duration differences in

Hebrew, between stressed and weak vowels, and in English, which they learn as a second language at school. We further hypothesized that the short training provided will induce fast improvements in the accuracy and speed of discrimination. Specifically, we assumed that the adult participants would learn to leverage the acoustical duration differences that exist in Hebrew in the context of stress differences to improve their discrimination between two vowels that differ in length, although they resemble one category in their native language.

MATERIAL AND METHODS

PARTICIPANTS

Sixty young adults (all female, aged 20–35 years) participated in this study. The participants were divided to three groups: (a) the “Hebrew training” group, which included 20 native Hebrew speakers ($M = 26.6 \pm 2.96$ years); (b) the “Hebrew control” group, which comprised a separate set of 20 native Hebrew speakers ($M = 26.15 \pm 2.37$ years); and (c) the “Arabic control” group, which included 20 native Arabic speakers ($M = 23.8 \pm 1.46$ years). The Hebrew speakers had minimal exposure to Arabic in elementary school; specifically, they studied Modern Standard Arabic for up to 2 years. However, they had no exposure to spoken Arabic. All participants were undergraduate students at Tel Aviv University and had advanced knowledge in English as a second language. All participants met the following prerequisite criteria: (a) pure tone air-conduction thresholds ≤ 20 dB hearing level bilaterally at octave frequencies of 500–4,000 Hz (ANSI 2004), (b) no history of language or learning disorders, (c) no known attention deficit disorders, and (d) no previous experience in psychoacoustic testing. We used a detailed, self-reporting questionnaire to collect the previously mentioned background information (criteria 2–4).

STIMULI

The stimuli consisted of 60 consonant–vowel–consonant (CVC) nonwords (Appendix). We chose single syllable words because they are common in Hebrew and Arabic and were used in previous studies that assessed discrimination of vowel contrasts (e.g., Faris et al., 2018; Flege & MacKay, 2004; McAllister et al., 2002; Nenonen et al., 2005; Tsukada, 2011; Tsukada et al., 2014; Tyler et al., 2014). In Arabic CVC words can have long or short vowels. For example, /bir/ “charity” contrasts with /bi:r/ “well, fountain” (Tsukada, 2011). CVCVC words were not used because in Arabic closed final syllables in bisyllabic words tend to receive long vowels (Hayes, 1995; Watson, 2011), which might have provided an additional cue for the Arabic speakers. In addition, in Arabic vowels tend to be neutralized, that is, short, in open syllables (Thelwall & Sa’Adeddin, 1990) therefore CVCV words were not used as well.

There were 30 nonwords that included short Arabic vowels [a/o], and 30 nonwords with similar consonants included the long Arabic vowels [a:/o:] (e.g., [foj, fo:j]). The consonants of the nonwords varied in terms of place and manner of articulation to avoid possible influences on discrimination. Forty nonwords were used for training (“trained” nonwords) and 20 were used to test the generalization of the learning-gains to new words

TABLE 1. Acoustic characteristics of the long and short Arabic vowels produced by two speakers.

Speakers	Vowels/acoustic characteristics	Short /o/	Long /o/	<i>p</i> value	Short /a/	Long /a/	<i>p</i> value
Speaker 1	Mean duration (sec)	0.1 (0.01)	0.14 (0.02)	$p < 0.001$	0.1 (0.02)	0.14 (0.02)	$p < 0.001$
	Mean F0 (Hz)	243.88 (7.08)	244.26 (6.03)	$p = 0.844$	246.9 (7.57)	237.62 (6.91)	$p < 0.001$
	Mean F1 (Hz)	668.11 (84.17)	638.56 (76.71)	$p = 0.248$	815.69 (120.53)	851.26 (70.25)	$p = 0.838$
	Mean F2 (Hz)	1,420.14 (129.60)	1,321.17 (109.39)	$p = 0.039$	1,643.29 (100.12)	1,606.4 (69.83)	$p = 0.434$
Speaker 2	Mean duration (sec)	0.1 (0.02)	0.15 (0.02)	$p < 0.001$	0.1 (0.02)	0.15 (0.03)	$p < 0.001$
	Mean F0 (Hz)	235.14 (9.34)	222.68 (7.04)	$p < 0.001$	236.41 (5.64)	222.75 (5.03)	$p < 0.001$
	Mean F1 (Hz)	721.12 (56.86)	665.45 (62.17)	$p = 0.027$	846.44 (61.42)	862.79 (43.07)	$p = 0.507$
	Mean F2 (Hz)	1,510.36 (135.45)	1,351.04 (93.69)	$p = 0.239$	1,673.53 (97.05)	1625.05 (85.31)	$p = 0.292$

Note: Standard deviation is shown in parentheses. Sec = seconds, Hz = Hertz, F0 = pitch, F1 and F2 (in Hz) = first and second formants, respectively. The *p* values for the multiple pairwise comparisons between the long and short vowel measures are also shown.

(“generalization” nonwords). Two female native Arabic speakers who were native speakers of the dialect spoken in the center of Israel (Muthallath dialect, in Arabic) produced the stimuli (30 words each). The purpose of using two speakers was to exclude the possibility that the manner in which vowels were uttered by a particular speaker would influence the results. The speakers recorded the CVC nonwords within sentences to maintain natural production. The stimuli were digitally recorded in a soundproof room by means of a AT-892-TH microphone using Sound-Forge software (version 7.0) and stereo channels at a sampling rate of 44,100 Hz and 16-bit quantization level. The amplitudes were normalized (-16dB RMS^1) to prevent intensity differences between the nonwords. The stimuli were delivered from an IBM-compatible personal computer, by means of a GSI-61 audiometer and were presented binaurally using THD-50 headphones at a comfortable level. We calculated several acoustic characteristics for each vowel using the speech analyzing software PRAAT (Boersma & Weenink, 2018). These included vowel length, fundamental frequency (F0), and first and second formant frequencies.

Table 1 displays the mean results of the acoustic characteristics (separated by vowel and speaker). A General Linear Model analysis was conducted with Vowel Length (short, long), Vowel (/a/, /o/), and Speaker (speaker 1, speaker 2) as the between-subject factors and F0, F1, F2, and duration as the dependent variables. The multivariate analysis indicated that Vowel Length had a significant effect on duration ($F[1,223] = 360.24$ $p < 0.001$, $\eta^2 = 0.63$), F0 ($F[1,223] = 88.61$ $p < 0.001$, $\eta^2 = 0.29$), and F2 ($F[1,223] = 6.54$ $p = 0.011$, $\eta^2 = 0.03$). Further, there were significant Vowel Length \times Vowel interactions for F0 ($F[1,223] = 8.49$ $p = 0.004$, $\eta^2 = 0.04$) and F1 ($F[1,223] = 4.51$ $p = 0.035$,

$\eta^2 = 0.02$). In addition, there was a significant Vowel Length \times Speaker interaction for duration ($F[1,223] = 8.19$ $p = 0.005$, $\eta^2 = 0.04$) and F0 ($F[1,223] = 21.41$ $p < 0.001$, $\eta^2 = 0.09$), and the Vowel Length \times Vowel \times Speaker interaction for F0 was significant ($F[1,223] = 5.17$ $p = 0.024$, $\eta^2 = 0.02$). There were no other significant ($p < 0.05$) interactions with Vowel Length.

The pairwise comparisons for these interactions are also detailed in Table 1. The results indicate that there is a consistent difference in duration between the short and long vowels, as expected. Additional significant differences were found for the F0 and formant frequencies, including the following: (a) higher F2 for the short vowel as compared to the long vowel for the /o/ vowels produced by speaker 1, (b) higher F0, F1, and F2 for the short vowel compared to the long vowel for the /o/ vowels produced by speaker 2, and (c) higher F0 for the short vowel compared to the long vowel for the /a/ produced by both speaker 1 and 2. Note that the duration differences for the vowels /o-o:/ ($M = 47.88$ milliseconds) and /a-a:/ ($M = 50.30$ milliseconds) are similar to the duration differences reported by Amir et al. (2014) for the “Muthallath” (Triangular) Dialect ($M = 53$ milliseconds for /o/ and /o:/, respectively and $M = 56$ milliseconds for /a/ and /a:/, respectively).

PROCEDURE

Discrimination Testing

Four triplets were constructed for each nonword (60 nonwords \times 4 triplets = 240 triplets; see Appendix) using ABX—SuperLab 4.5 software. Each triplet was produced by a single speaker and consisted of two identical nonwords and one nonword that differed only in vowel length. The four triplets were the following: (a) short-long-long, (b) short-long-short, (3) long-short-short, and (d) long-short-long. The interstimulus interval was 500 milliseconds. The participants were instructed to “decide whether the last non-word in a triplet resembles the first or second non-word and to press ‘1’ or ‘2’ on the keyboard, accordingly, as fast and accurately as possible.” We did not provide any feedback to the participants.

STUDY DESIGN

The Hebrew training group participated in three sessions: pretraining testing, training, and posttraining testing. There were approximately 2–3 days between the first and second sessions and 1 week between the second and last sessions. The first (pretraining) and last (posttraining) sessions were identical and included 60 nonwords, one triplet for each nonword, which were presented in a random order. These sessions lasted approximately 10 minutes. The second (training) session included 40 of the 60 nonwords, with each nonword presented four times. Thus, overall training consisted of 40 nonwords \times 4 triplets = 160 triplets. The 40 nonwords that were used for training were termed “trained” nonwords. The remaining 20 nonwords were termed “generalization” nonwords. The training session lasted approximately 20 minutes (Figure 1). No feedback was provided during training.

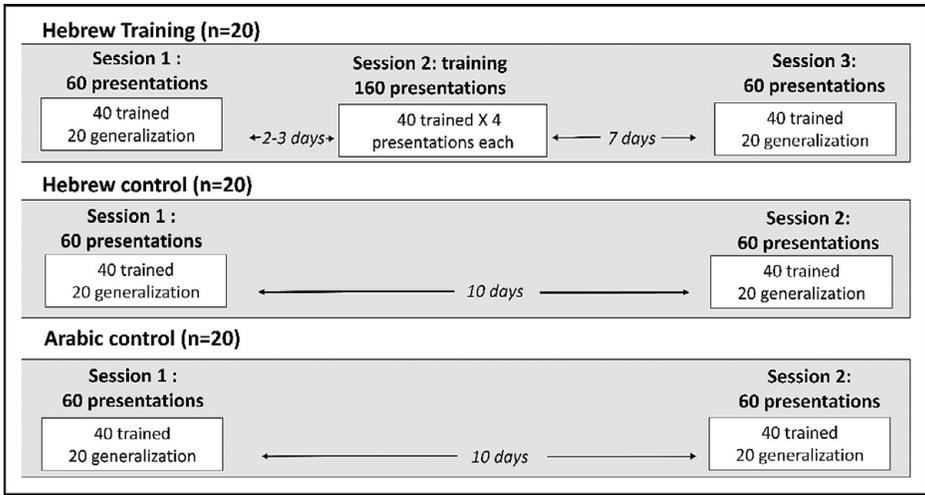


FIGURE 1. Study design. The study included three groups: a Hebrew training group, a Hebrew control, and an Arabic control group. The Hebrew training group participated in three sessions: pre- and post-training sessions that included 60 nonwords (40 “trained” and 20 “generalization”) each, and a training session that included the 40 “trained” nonwords \times 4 presentations each (i.e., overall 160 presentations). The Hebrew and Arabic control groups participated only in two sessions that included 60 “nonwords” (40 “trained” and 20 “generalization”).

The Hebrew and Arabic control groups participated in two sessions that were spaced approximately 10 days apart. These sessions were identical to the training group’s pre- and posttraining ones, with each including 60 nonwords, and lasted approximately 10 minutes. Note that the three groups of participants were tested with both the “trained” and “generalization” words in the first and last sessions to confirm that the training effects are demonstrated only in the trained group.

All participants provided us with informed consent. The study was approved by the Institutional Review Board of Tel-Aviv University.

DATA ANALYSIS

We assessed the participants’ perceptions of vowel length using two variables: accuracy of discrimination and reaction time (RT). We calculated accuracy of discrimination as the percent of correct responses (percent correct), and we measured RTs by the interval between the end of the third nonword and the onset of the button press (in milliseconds). It was hypothesized that the native Hebrew speakers would show reduced performance compared to the Arabic speakers in measures of correct responses and RT. Training was hypothesized to improve accuracy and RTs, plausibly by improving vowel length perception in the context of phonemic distinction. To compare performance across groups, several 2 Session \times 3 Group repeated measures analyses of variance (ANOVAs) were conducted with Session (first, last) as the within-subject variable and Group (Hebrew training, Hebrew control, Arabic control) as the between-subject variable. Specifically, two ANOVAs were conducted for the trained nonwords (one for

percent correct and one for RT), and two ANOVAs were conducted for the generalization nonwords (one for percent correct and one for RT). Before conducting the ANOVAs, residuals of all the dependent measures were tested for normality using one-sample Kolmogorov–Smirnov test. Seven of the eight tests failed to reach significance (all p 's > 0.20), and one provided marginally significant result ($p = 0.038$). This suggested that overall, the data met the assumptions of the planned ANOVA tests. All the ANOVAs included contrast analyses with Bonferroni corrections.

RESULTS

GROUP COMPARISONS

Accuracy data: Figure 2(a) suggests that, in the first session, the native Hebrew speakers' discrimination scores in regard to vowel length were above chance level (>50%). A one-sample t -test confirmed that the mean percent of correct responses ($M = 67.30$, $SD = 8.03$) was found to be significantly higher than chance (50%) [$t(39) = 13.63$, $p < 0.001$]. Nevertheless, performance of the Hebrew speakers was lower in comparison to the native Arabic speakers ($M = 81.8$ $SD = 4.5$). Following a single training session, however, the trained group's ability to discriminate ($M = 81.3$ $SD = 1.4$) improved to the level of native Arabic speakers ($M = 83.5$ $SD = 5.3$). The data analysis of the percent correctly identified confirm this observation, showing significant effects of Group ($F[2,57] = 27.40$ $p < 0.001$, $\eta^2 = 0.49$) and Session ($F[1,57] = 66.52$ $p < 0.001$, $\eta^2 = 0.54$) and a significant Group \times Session interaction ($F[2,57] = 28.44$ $p < 0.001$, $\eta^2 = 0.5$). The pairwise comparisons of this interaction revealed that the Arabic control group performed better in the first session in comparison to the two native Hebrew speakers' groups ($p < 0.001$), with no significant difference between the two latter groups ($p > 0.99$). In the last session, both the Arabic control and the Hebrew training groups had higher scores compared to the Hebrew control group ($p < 0.001$), with no significant difference between the first two groups ($p = 0.856$). Additional pairwise comparisons revealed that only the Hebrew training group showed a

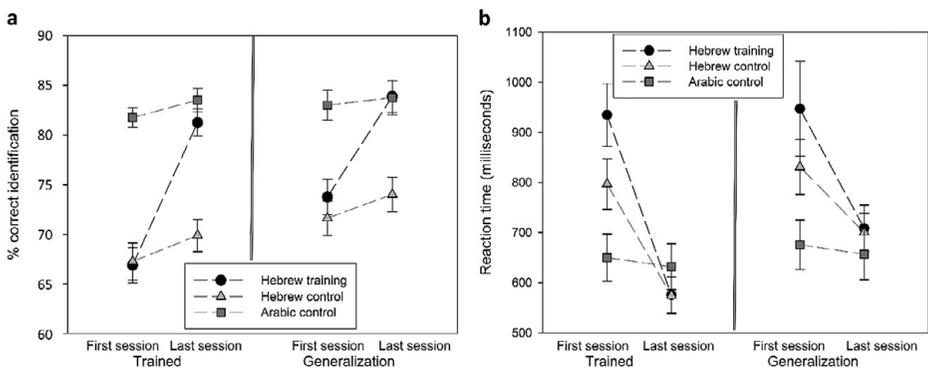


FIGURE 2. Mean (± 1 standard error) performance of the native Arabic speakers (Arabic control), trained native Hebrew speakers (Hebrew training), and controlled native Hebrew speakers (Hebrew control) for the trained words (left-hand side) and the untrained words (right-hand side [generalization]). Figure 2(a) displays the percent correctly identified and Figure 2(b) displays the reaction times.

significant improvement between the first and last (i.e., pre- and posttraining) sessions ($p < 0.001$).

RT data: Figure 2(b) further suggests that both groups of Hebrew speakers had longer RTs ($M = 934.71$ $SD = 62.40$, $M = 796.78$ $SD = 50.35$, for the trained and control groups, respectively) compared to the native Arabic speakers ($M = 649.90$ $SD = 211.53$) in the first session. The RT data analysis indicated that Session had a significant effect ($F[1,57] = 59.62$ $p < 0.001$, $\eta^2 = 0.51$), that Group did not have a significant effect ($F[2,57] = 1.91$ $p = 0.16$), and that the Group \times Session interaction was significant ($F[2,57] = 14.67$ $p < 0.001$, $\eta^2 = 0.34$). Pairwise comparisons revealed that, in the first session, the Arabic control group had shorter RTs compared to the Hebrew training group ($p = 0.001$), and that there were no significant differences between the Arabic and Hebrew control groups ($p = 0.175$) or between the Hebrew training and control groups ($p = 0.225$). In the last session, there were no significant differences across groups ($p > 0.05$).

INDIVIDUAL ANALYSIS

The individual-level analysis of the percent correctly identified (Figure 3) showed that only two (of 40) native Hebrew speakers reached an “Arabic-like” performance (calculated as the native Arabic speakers’ mean ± 1 SD) in the first session. However, in the last session, 16 (of 20) native Hebrew speakers from the Hebrew training group and three (of 20) from the Hebrew control group achieved an “Arabic-like” performance. We conducted a Fisher exact test to compare the Hebrew training and

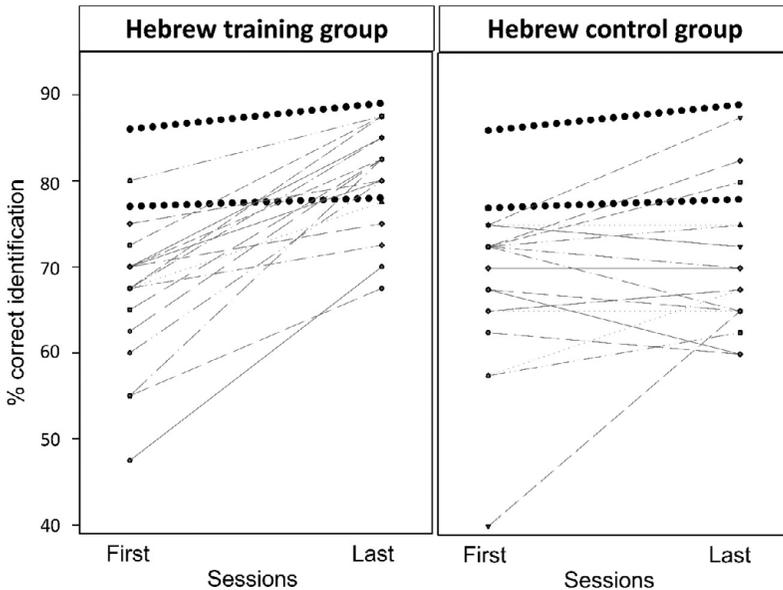


FIGURE 3. Individual-level performances of the Hebrew training and control groups in the first and last testing sessions with the trained nonwords. The “Arabic-like” performance range (calculated as the Arabic control group’s mean ± 1 standard deviation) is shown as dotted lines for the purpose of comparison.

control groups in regard to the group members who achieved this “Arabic-like” performance in the last session. The results revealed a significant difference between the groups (Fisher exact test: $p < 0.001$), further strengthening the training’s positive effect on an individual’s performance with the trained words.

GENERALIZATION TO UNTRAINED NONWORDS

We identified a similar performance pattern for the trained and untrained nonwords for both the percent correctly identified (Figure 2[a]) and RTs (Figure 2[b]), suggesting a large generalization of the training-induced gains. Specifically, the data analysis for the percent correctly identified showed that Group ($F[2,57] = 14.08$ $p < 0.001$, $\eta^2 = 0.33$), Session [$F(1,57) = 17.32$ $p < 0.001$, $\eta^2 = 2.33$], and the Session \times Group interaction ($F[2,57] = 7.4$ $p = 0.001$, $\eta^2 = 0.21$) all had significant effects. A pairwise comparison revealed that the Arabic control group performed better in the first session in comparison to the two native Hebrew speaker groups ($p < 0.001$). In addition, there were no significant differences between the latter groups ($p > 0.99$). For the last session, both the Arabic control group and the Hebrew training group showed better discrimination compared to the Hebrew control group ($p < 0.001$), with no significant difference between the first two groups ($p > 0.99$). Additional pairwise comparisons revealed that only the Hebrew training group displayed a significant improvement between the first and last sessions ($p < 0.001$),

In the analysis of the participants’ RTs, Session had a significant effect ($F[1,57] = 23.46$ $p < 0.001$, $\eta^2 = 0.29$), Group did not have a significant effect ($F[2,57] = 2.29$ $p = 0.111$), and the Group \times Session interaction had a significant effect ($F[2,57] = 5.63$ $p = 0.006$, $\eta^2 = 0.17$). The pairwise analysis revealed that, in the first session, the Arabic control group had shorter RTs than the Hebrew training group ($p = 0.023$), but that there were no significant differences between the Arabic and Hebrew control groups ($p = 0.360$) or between the two native Hebrew groups ($p = 0.723$). In the last session, there were no significant differences across the groups ($p > 0.99$).

DISCUSSION

The present experiment produced two main findings. First, Hebrew speakers had reduced discrimination performance of vowel length in comparison to the native Arabic speakers, supporting the PAM prediction on the influence of the native phonetic categories on discrimination of nonnative contrasts (Best et al., 1995). Second, a single training session significantly improved the discrimination between nonnative vowel stimuli that belong to the same duration category in L1, possibly because adult Hebrew speakers already have experience in using duration differences of vowels for the perception of lexical stress. Thus, it is possible that previous experience in perceiving durational suprasegmental cues of lexical stress was adapted for phonemic discrimination of vowels.

The finding that the Hebrew speakers were slower and less accurate than the native Arabic speakers in discriminating between short and long vowels in Arabic suggests that listening experience in L1 alters the listener’s ability to discriminate vowel length in L2. There are some disagreements in the literature concerning the influence of the native language on discrimination of nonnative contrasts of vowel length (Hisagi et al., 2016;

McAllister et al., 2002; Nenonen et al., 2005; Tsukada et al., 2014). While some studies suggest that listeners can efficiently perceive the vowel length contrast regardless of their native phonological categories (e.g., Altmann et al. 2012; Bohn, 1995; Cebrian, 2006; Flege et al., 1997; García Lecumberri & Cenoz Inagui, 1997), others argue that the native language has some influence on vowel length perception, (e.g., Hisagi et al., 2010, 2015; McAllister et al., 2002; Tsukada et al., 2014; Nenonen et al., 2005), as is suggested by the PAM (Best et al., 1995). The PAM predicts reduced discrimination when two members of a phonetic-phonemic contrast in a nonnative language are assimilated to one phonetic-phonological category in the native language, that is, an SC assimilation (Best et al., 1995; Kuhl et al., 2006). The finding of reduced discrimination for the Hebrew speakers compared to the Arabic speakers in the current study therefore supports this prediction.

The Hebrew speakers have nevertheless demonstrated some ability to perceive the vowel length contrast in Arabic, as was demonstrated by their above chance discrimination performance in the first session. This ability might be related to their previous experience with duration differences of vowels in Hebrew that characterizes stressed versus weak syllables (Silber-Varod et al., 2016). It can also be the result of their advanced knowledge in English, a language with duration differences between the tense and lax vowels (Leung et al., 2016).

It should be noted that the native Arabic speakers in the present study did not reach maximum discrimination scores ($M = 82\%$), despite the fact that vowel length is a robust phonemic contrast in their native language. Furthermore, their mean score was lower than reported in a previous study that examined native Arabic speakers' perceptions of vowel length (with mean correct identification higher than 94% for different vowels; Tsukada, 2011). This discrepancy can be explained by the fact that Tsukada (2011) tested discrimination using real words, allowing the participants to rely on their lexical knowledge in Arabic. In the present study, however, we chose to use nonwords to avoid lexical knowledge as an intervening factor when comparing Arabic and Hebrew speakers, as the latter have no lexical knowledge in Arabic.

As mentioned in the preceding text, the present study's second main finding is that a short training significantly improved the discrimination of vowel length in L2, as evidenced by the trained Hebrew group's increase in correct responses. These improvements resulted in most participants achieving native-level performances by the end of training. A significant training-induced gain in one's ability to discriminate L2 vowel contrasts is in agreement with the SLM (Flege, 1995, 2003). The SLM argues that the ability to discriminate and categorize speech remains intact throughout life. Thus, even adults can learn to perceive new contrasts in L2 and adjust the weight given to different phonetic/acoustic cues if they have enough attentional resources and speech input (Flege & MacKay, 2004). This explanation is also in line with the ASP model (Strange & Shafer, 2008). The ASP characterizes first-language speech processing in adults as reflecting automatic selective perceptual routines for the detection and discrimination of language-specific, phonologically relevant phonetic contrasts. However, nonnative listeners may require attentional resources to facilitate the perception of novel contrasts (e.g., Hisagi et al., 2010). Accordingly, whereas the Arabic speakers showed good performance for contrasts in their native language, the Hebrew speakers needed training to improve discrimination. Their improvement, however, does not necessarily mean that they acquired automatic selective perceptual routines similarly to Arabic speakers.

It was recently demonstrated that listening exposure to L2 alone might not be sufficient for facilitating discrimination of vowel length contrasts by American English speakers who are learning Japanese as an L2 (Hisagi et al., 2016). The authors suggested that targeted *training* of L2 phonology may be necessary to allow for changes in one's processing of L2 speech contrasts at an early, automatic level (Hisagi et al., 2016). The findings of the present study add to this line of research by showing that a short-targeted training is efficient in enhancing attention and consequently changing discrimination of a vowel length contrast. Furthermore, the fact that the participants displayed an improvement in their discrimination of vowel length approximately seven days after the training session suggests that the training-induced changes that consolidated into their long-term memory (e.g., Hauptmann & Karni, 2002; Hauptmann et al., 2005; Molloy et al., 2012; Roth et al., 2005). During consolidation, the neural representations of the trained contrast may have been strengthened. Further studies with additional identification and categorization tasks are advisable to support a profound and sustained change in selective perceptual routines following training. Also, tasks that incorporate every-day listening conditions, such as background noise, may serve as a test of whether the perception of vowel length phonological contrast can become nativelike. In addition, comparing the influence of training on vowels versus consonants (e.g., the vowel contrast /a-a:/ in Hebrew speakers vs. the /b-p/ distinction in Arabic speakers) may help to shed light on whether some contrasts are more difficult to acquire than others, broadening the implications of the current study.

It is important to note that the quick improvement of Hebrew speakers in discrimination after one training session might have also been related to the fact that duration contrast is a salient psychoacoustic cue (Burnham, 1986) that serves as a suprasegmental cue for lexical stress in Hebrew (Silber-Varod et al., 2016). Thus, one short training might have been sufficient to emphasize this duration cue and make it a legitimate vowel contrast in the context of CVC syllables for the Hebrew speakers. Further studies that explore native speakers of stress (e.g., Hebrew) and nonstress languages (e.g., Chinese) may enhance our understanding of whether suprasegmental duration cues in L1 support one's ability to learn vowel length contrasts in L2 (e.g., Arabic).

Along with the increase in correct responses following one training session, there was a decrease in RTs for the trained Hebrew speakers. This faster discrimination performance may suggest that the participants became more efficient in differentiating between vowel lengths either during or after the training session. However, we observed a similar decrease in RTs for the Hebrew control group who received no training between their two testing sessions (and did not show significant improvements in accuracy). It is possible, therefore, that the short exposure to the task that was provided to both Hebrew speaking groups in the first testing session (60 nonwords) was enough to initiate tuning and adaptation processes, as previously suggested for nonspeech auditory stimuli (Hauptmann & Karni, 2002; Hauptmann et al., 2005). These procedural processes may have assisted the Hebrew speakers in adapting to the specific experimental settings, and forming an efficient task solution routine that improved the RTs for both the trained and control groups (e.g., Ahissar & Hochstein, 1997; Vakil et al., 2014). However, only training yielded an actual perceptual change that was reflected in improved discrimination performance.

Importantly, the training session used in the present study improved performance not only with the target nonwords but also with a new set of nonwords that the participants did

not encounter during training. This result suggests a generalization (transfer) of the learning. Specifically, it may suggest that the short exposure to a narrow set of nonword stimuli during training facilitated learning that is not context-dependent and thus can generalize to various stimulus features (in terms of the surrounding consonants). Such learning may be attributed to neural changes in high order brain areas that allow processing of input from untrained local networks (Ahissar et al., 2009; Censor & Sagi, 2009; Censor et al., 2012; Harris et al., 2012; Karni, 1996; Zaltz et al., 2020). It remains to be tested whether the training-induced changes for vowel length can generalize across different vowels as well. Future studies may want to explore this question by testing generalization to untrained vowels.

LIMITATIONS

The present study has several limitations. One limitation is related to the fact that although duration was constantly longer for both vowels and speakers, there was also a difference in pitch and F1, especially for one of the two speakers (see Table 1). Amir et al. (2014) previously reported similar differences in formant frequencies for short versus long vowels in spoken Arabic. One can argue that the difference in formant frequencies may have influenced the present study's discrimination results. However, previous studies have reported that the effect of pitch on the production and perception of vowel length is negligible for native speakers of Japanese (Beckman, 1986; Minagawa-Kawai et al., 2002). Burnham (1986) also argued that duration is a more prominent acoustic cue compared to spectral cues. Thus, it is plausible to assume that our participants based their discrimination mainly (if not entirely) on the duration cues. Another possible limitation of the present study may be that the Hebrew speakers had advanced knowledge in English as an L2, so they might have had some experience with duration differences in English vowels even though vowel distinction is not phonemic in English. Finally, one cannot rule out the possibility that Hebrew speakers would have better discriminated between vowel length contrasts had they presented with bisyllable words. Future studies may test this possibility.

CONCLUSIONS

The primary focus of the current study was to test whether one's listening experience with a native language (i.e., Hebrew) that has no phonemic contrast of vowel length influences their ability to discriminate between long and short vowels in a nonnative (L2) language (i.e., Arabic). The results showed that although the native Hebrew speakers reached above chance discrimination, they were unable to discriminate as well as native Arabic speakers. This result reflects the influence of L1 on vowel contrast perception. Our second goal was to assess whether the native language's influence can be reformed, retained, and generalized following a short training. Our findings suggest that a short training can improve the perception of vowels using the recruitment of top-down abilities such as attention resources. Further, our findings indicate that these changes can generalize to untrained stimuli. Additional studies are needed to examine whether a short perceptual training paradigm can also improve the production of nonnative vowel contrasts. Overall, the present findings support the models that predict

that even adults can learn phonetic/acoustic cues due to neural plasticity (e.g., Flege, 1995, 2003). The present findings have important practical implications for the clinicians, teachers, and educators who teach and train individuals to learn an L2.

NOTE

¹RMS: Root Mean Square.

REFERENCES

- Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–406.
- Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*, 285–299.
- Altmann, C. F., Uesaki, M., Ono, K., Matsuhashi, M., Mima, T., & Fukuyama, H. (2014). Categorical speech perception during active discrimination of consonants and vowels. *Neuropsychologia*, *64*, 13–23 <https://doi.org/10.1016/j.neuropsychologia.2014.09.006>.
- Altmann, H., Berger, I., & Braun, B. (2012). Asymmetries in the perception of non-native consonantal and vocalic length contrasts. *Second Language Research*, *28*, 387–413. <https://doi.org/10.1177/0267658312456544>
- Amir, N., Amir, O., & Rosenhouse, J. (2014). Colloquial Arabic vowels in Israel: A comparative acoustic study of two dialects. *Journal of the Acoustical Society of America*, *136*, 1895–1907. <https://doi.org/10.1121/1.4894725>
- Beckman, M. E. (1986). *Stress and non-stress accent*. Foris Publications.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). American listeners' perception of non-native consonant contrasts varying in perceptual assimilation to English phonology. *Journal of the Acoustical Society of America*, *109*, 775–794. <https://doi.org/10.1121/1.1332378>
- Best, C. T., McRoberts, G. W., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior & Development*, *18*, 339–350. [https://doi.org/10.1016/0163-6383\(95\)90022-5](https://doi.org/10.1016/0163-6383(95)90022-5).
- Best, C. T., McRoberts, G. W., & Sithole, N. N. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 345–360. <https://doi.org/10.1037/0096-1523.14.3.345>
- Best, C. T., Singh, L., Bouchard, J., Connelly, G., Cook, A., & Faber, A. (1997, April). Developmental changes in infants' discrimination of non-native vowels that adults assimilate to two native categories. In *Poster presented at the Meeting of Society of Research in Child Development*, (pp. 3–6).
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro, & O.-S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege*, (pp. 13–34). John Benjamins Publishing Company.
- Boersma, P. & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program]. Version 6.0.37. <http://www.praat.org/>
- Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research*, (pp. 279–304). York Press.
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*, 217–243. <https://doi.org/10.1177/00238309030460020801>
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729. <https://doi.org/10.1016/j.cognition.2007.04.005>
- Burnham, D. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistic*, *7*, 207–239.
- Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, *34*, 372–387. <https://doi.org/10.1016/j.wocn.2005.08.003>.
- Censor, N., & Sagi, D. (2009). Global resistance to local perceptual adaptation in texture discrimination. *Vision Research*, *49*, 2550–2556.

- Censor, N., Sagi, D., & Cohen, L. G. (2012). Common mechanisms of human perceptual and motor learning. *Nature Reviews Neuroscience*, *13*, 658–664.
- Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research*, *57*, 1468–1479. https://doi.org/10.1044/2014_JSLHR-L-13-0279.
- D'Imperio, M., & Rosenthal, S. (1999). Phonetics and phonology of main stress in Italian. *Phonology*, *16*, 1–28. <https://doi.org/stable/4420141>
- Eilers, R. E., Gavin, W. J., & Oller, D. K. (1982). Cross linguistic perception in infancy: The role of linguistic experience. *Journal of Child Language*, *9*, 289–302. <https://doi.org/10.1017/S0305000900004724>.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306. <https://doi.org/10.1126/science.171.3968.303>
- Escudero, P., & Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *Journal of the Acoustical Society of America*, *129*, EL1–EL7. <https://doi.org/10.1121/1.3525042>
- Faris, M. M., Best, C. T., & Tyler, M. D. (2018). Discrimination of uncategorized non-native vowel contrasts is modulated by perceived overlap with native phonological categories. *Journal of Phonetics*, *70*, 1–19. <https://doi.org/10.1016/j.wocn.2018.05.003>.
- Fllege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research*, (pp. 233–277). York Press.
- Fllege, J. E., Bohn, O.-S., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics*, *25*, 437–470. <https://doi.org/10.1006/jpho.1997.0052>.
- Fllege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, *26*, 1–34. <https://doi.org/10.1017/S0272263104026117>
- Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, *5*, 171–189. <https://doi.org/10.1177/002383096200500401>
- García Lecumberri, M. L., & Cenoz Inagui, J. (1997). L2 perception of English vowels: Testing the validity of Kuhl's prototypes. *Revista Alicantina de Estudios Ingleses*, *10*, 55–68. <https://doi.org/10.14198/raei.1997.10.06>
- Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, *25*, 577–588. <https://doi.org/10.1037/0012-1649.25.4.577>
- Guion, S. G., Fllege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, *107*, 2711–2724. <https://doi.org/10.1121/1.428657>
- Harris, H., Glikberg, M., & Sagi, D. (2012). Generalized perceptual learning in the absence of sensory adaptation. *Current Biology*, *22*, 1813–1817.
- Hauptmann, B., & Karni, A. (2002). From primed to learn: The saturation of repetition priming and the induction of long-term memory. *Cognitive Brain Research*, *13*, 313–322.
- Hauptmann, B., Reinhart, E., Brandt, S. A., & Karni, A. (2005). The predictive value of the leveling off of within session performance for procedural memory consolidation. *Cognitive Brain Research*, *24*, 181–189. <https://doi.org/10.1016/j.cogbrainres.2005.01.012>
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.
- Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*, 360–378. <https://doi.org/10.1016/j.specom.2005.04.007>
- Hisagi, M., Miyagawa, S., Shafer, V. L., Kotek, H., Sugawara, A., & Pantazis, D. (2014). Perception of Japanese vowel duration contrasts by L1 and L2 learners of Japanese: An EEG study. In *MITWPL #73 Proceedings of FAJL 7*, (pp. 45–55).
- Hisagi, M., Shafer, V. L., Miyagawa, S., Kotek, H., Sugawara, A., & Pantazis, D. (2016). Second-language learning effects on automaticity of speech processing of Japanese phonetic contrasts: An MEG study. *Brain Research*, *1652*, 111–118. <https://doi.org/10.1016/j.brainres.2016.10.004>
- Hisagi, M., Shafer, V. L., Strange, W., & Sussman, E. (2010). Perception of a Japanese vowel length contrast by Japanese and American English listeners: Behavioral and electrophysiological measures. *Brain Research*, *1360*, 89–105. <https://doi.org/10.1016/j.brainres.2010.08.092>
- Hisagi, M., Shafer, V. L., Strange, W., & Sussman, E. (2015). Perception of Japanese consonant length contrasts by Japanese and American English listeners: Behavioral and electrophysiological measures. *Brain Research*, *1626*, 218–231. <https://doi.org/10.1016/j.brainres.2015.06.001>.

- Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America*, 118, 3267–3278. <https://doi.org/10.1121/1.2062307>.
- Jamieson, D., & Morosan, D. (1986). Training non-native speech contrasts in adults: Acquisition of the English /d/-/θ/ contrast by francophones. *Perception and Psychophysics*, 40, 205–215. <https://doi.org/10.3758/bf03211500>
- Kabakoff, H., Go, G., & Levi, S. V. (2020). Training a non-native vowel contrast with a distributional learning paradigm results in improved perception and production. *Journal of Phonetics*, 78, 1–17. <https://doi.org/10.1016/j.wocn.2019.100940>
- Karni, A. (1996). The acquisition of perceptual and motor skills: A memory system in the adult human cortex. *Cognitive Brain Research*, 5, 39–48.
- Kronrod, Y., Coppess, E., & Feldman, N. H. (2016). A unified account of categorical effects in phonetic perception. *Psychonomic Bulletin & Review*, 23, 1681–1712. <https://doi.org/10.3758/s13423-016-1049-y>.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, 363, 979–1000. <https://doi.org/10.1098/rstb.2007.2154>.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, 13–21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x>.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom B. (1992). Linguistic experiences alters phonetic perception in infants by 6 months of age. *Science*, 255, 606–608. <https://doi.org/10.1126/science.1736364>
- Ladefoged, P. (2001). *Vowels and consonants*. Blackwell.
- Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, 20, 213–225. [https://doi.org/10.1016/0022-0965\(75\)90099-5](https://doi.org/10.1016/0022-0965(75)90099-5).
- Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *Journal of the Acoustic Society of America*, 128, 3757–3768. <https://doi.org/10.1121/1.3506351>.
- Leung, K., Jongman, A., Wang, J., & Sereno, Z. (2016). Acoustic characteristics of clearly spoken English tense and lax vowels. *The Journal of the Acoustical Society of America*, 140, 45–58. <https://doi.org/10.1121/1.4954737>.
- Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *Journal of Acoustical Society of America*, 126, 2670–2682. <https://doi.org/10.1121/1.3224715>
- Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, 36, 141–157. <https://doi.org/10.1016/j.wocn.2007.03.001>
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30, 229–258. <https://doi.org/10.1006/jpho.2002.0174>
- Minagawa-Kawai, Y., Maekawa, K., & Kiritani, S. (2002). Effects of pitch accent and syllable position in identifying Japanese long and short vowels: Comparison of English and Korean speakers. *Journal of the Phonetic Society of Japan*, 6, 88–97.
- Molloy, K., Moore, D. R., Sohoglu, E., & Amitay, S. (2012). Less is more: Latent learning is maximized by shorter training sessions in auditory perceptual learning. *PLoS One*, 7, e36929.
- Most, T., Amir, O., & Tobin, Y. (2000). The Hebrew vowel system: Raw and normalized acoustic data. *Language and Speech*, 43, 295–30. <https://doi.org/10.1177/00238309000430030401>
- Mugitani, R., Pons, F., Fais, L., Dietrich, C., Werker, J. F., & Amano, S. (2009). Perception of vowel length by Japanese- and English-learning infants. *Developmental Psychology*, 45, 236–247. <https://doi.org/10.1037/a0014043>.
- Nenonen, S., Shestakova, A., Huottilainen, M., & Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and Language*, 92, 26–32. <https://doi.org/10.1016/j.bandl.2004.05.005>
- Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech Language and Hearing Research*, 50, 1496–1509. <https://doi.org/10.1044/1092-4388>

- Nishi, K., & Kewley-Port, D. (2008). Nonnative speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech Language and Hearing Research*, *51*, 1480–1493. <https://doi.org/10.1044/1092-4388>
- Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of Acoustical Society of America*, *100*, 577–592. <https://doi.org/10.1121/1.415884>
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 421–435. <https://doi.org/10.1037//0096-1523.20.2.421>
- Pruitt, J. S., Jenkins, J. J., & Strange, W. (2006). Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *Journal of the Acoustical Society of America*, *119*, 1684–1696. <https://doi.org/10.1121/1.2161427>
- Rato, A. (2014). Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. *Concordia Working Papers in Applied Linguistics*, *5*, 529–546.
- Roth, D. A., Kishon-Rabin, L., Hildesheimer, M., & Karni, A. (2005). A latent consolidation phase in auditory identification learning: Time in the awake state is sufficient. *Learning & Memory*, *12*, 159–164. <https://doi.org/10.1101/87505>
- Segal, O., Hejli-Assi, S., & Kishon-Rabin, L. (2016). The effect of listening experience on the discrimination of /ba/ and /pa/ in Hebrew-learning and Arabic-learning infants. *Infant Behavior and Development*, *42*, 86–99. <https://doi.org/10.1016/j.infbeh.2015.10.002>
- Silber-Varod, V., Sagi, H., & Amir, N. (2016). The acoustic correlates of lexical stress in Israeli Hebrew. *Journal of Phonetics*, *56*, 1–14. <https://doi.org/10.1016/j.wocn.2016.01.003>
- Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153–191). John Benjamins Publishing Company.
- Thelwall, R., & Sa' Adeddin, M. A. (1990). Arabic. *Journal of the International Phonetic Association*, *20*, 37–39. <https://doi.org/10.1017/S0025100300004266>
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by adults and infants. *Child Development*, *47*, 466–472. <https://doi.org/10.2307/1128803>
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America*, *120*, 2285–2294. <https://doi.org/10.1121/1.2338290>
- Tsukada, K. (2011). The perception of Arabic and Japanese short and long vowels by native speakers of Arabic, Japanese, and Persian. *Journal of the Acoustical Society of America*, *129*, 989–998. <https://doi.org/10.1121/1.3531801>
- Tsukada, K., Hirata, Y., & Roengpitya, R. (2014). Cross-language perception of Japanese vowel length contrasts: Comparison of listeners from different first language backgrounds. *Journal of Speech, Language, and Hearing Research*, *57*, 805–814. https://doi.org/10.1044/2014_JSLHR-S-12-0416
- Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994). Discrimination of English /r-/l/ and /w-/y/ by Japanese infants at 6–12 months: Language-specific developmental changes in speech perception abilities. In *1994 International Conference on Spoken Language Processing* (pp. 1695–1698). Yokohama, Japan.
- Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*, 4–21. <https://doi.org/10.1159/000356237>
- Vakil, E., Hassin-Baer, S., & Karni, A. (2014). A deficit in optimizing task solution but robust and well-retained speed and accuracy gains in complex skill acquisition in Parkinson's disease: Multi-session training on the Tower of Hanoi puzzle. *Neuropsychologia*, *57*, 12–19. <https://doi.org/10.1016/j.neuropsychologia.2014.02.005>
- Wang, X., & Munro, M. J. (1999). The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention to temporal and spectral cues. In *Proceedings of the 14th International Congress of Phonetic Sciences* (Vol. 3, pp. 125–128).
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tone. *Journal of the Acoustical Society of America*, *106*, 3649–3658. <https://doi.org/10.1121/1.428217>
- Watson, J. C. E. (2011). Word stress in Arabic. In M. Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (Vol. 5, pp. 2990–3018). Blackwell.

- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349–355. <https://doi.org/10.2307/1129249>
- Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24, 672–683. <https://doi.org/10.1037/0012-1649.24.5.672>.
- Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, 37, 278–286. <https://doi.org/10.1037/h0080725>.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63. [https://doi.org/10.1016/S0163-6383\(02\)00093-0](https://doi.org/10.1016/S0163-6383(02)00093-0).
- Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50, 509–535. <https://doi.org/10.1146/annurev.psych.50.1.509>
- Zaltz, Y., Kishon-Rabin L, Karni A, Ari-Even Roth, D. (2020). Practice makes transfer imperfect: Evidence from auditory learning. *Ear and Hearing*, 41, 1470–1482. <https://doi.org/10.1097/AUD.0000000000000860>

APPENDIX

TRAINED WORDS

/a/

tad—ta:d, nad—na:d, lad—la:d, bash—ba:sh, nam—na:m, zal—za:l, nat—na:t, lab—la:b, mad—ma:d, zak—za:k, zam—za:m, xab—xa:b, bak—ba:k, zan—za:n, lash—la:sh, nab—na:b, mab—ma:b, san—sa:n, tak—ta:k, nal—na:l

/o/

nol—no:l, hok—ho:k, kon—ko:n, nosh—no:sh, sol—so:l, nod—no:d, mod—mo:d, ron—ro:n, ton—to:n, yol—yo:l, son—so:n, lob—lo:b, zof—zo:f, sob—so:b, lom—lo:m, zod—zo:d, mob—mo:b, hod—ho:d, dom—do:m, tod—to:d

GENERALIZATION WORDS

/a/

san—sa:n, han—ha:n, bam—ba:m, fad—fa:d, mash—ma:sh, naf—na:f, sat—sa:t, kat—ka:t, tab—ta:b, had—ha:d, las—la:s

/o/

doy—do:y, bon—bo:n, shon—sho:n, fob—fo:b, bof—bo:f, mon—mo:n, foj—fo:j, bok—bo:k, hod—ho:d