City University of New York (CUNY)

CUNY Academic Works

Publications and Research

Kingsborough Community College

2016

Is There a Bilingual Advantage in Phonetic and Phonological Acquisition? The Initial Learning of Word-Final Coronal Stop Realization in a Novel Accent of English

Laura Spinu

CUNY Kingsborough Community College

Jiwon Hwang

SUNY Stony Brook

Renata Lohmann

Western University

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/kb_pubs/173

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

Is there a bilingual advantage in phonetic and phonological acquisition? The initial learning of word-final coronal stop realization in a novel accent of English

Abstract

Research question: We address the question of whether the cognitive advantage of the bilingual mind, already demonstrated in the case of auditory processing or novel word acquisition, also applies to other linguistic domains, specifically to phonetic and phonological learning.

Design: We compare the performance of 17 monolinguals and 25 bilinguals from Canada in a production experiment with two tasks: imitation and spontaneous reproduction of a novel foreign accent, specifically Sussex English.

Data and analysis: To eliminate potential sources of variability, our focus is on a sound already existing in the subjects' production (the glottal stop), but differently mapped to surface representations in the novel accent to which they were exposed (i.e. as an allophone of coronal stops in word-final position). We measured the glottal stop rates of our subjects in baseline, training, and post-training.

Results: The two groups behaved differently, with bilinguals showing a larger increase of their glottal stop rate post-training. Our results are thus consistent with a bilingual advantage in phonetic and phonological learning.

Originality: We interpret these findings in light of recent psycholinguistic work and conclude that echoic memory strategies, possibly underlain by stronger subcortical encoding of sound in bilinguals, may account for our results by facilitating the re-mapping between existing mental representations of sounds and existing articulatory command configurations.

Significance: Our study adds to the body of work showing that there is an advantage of bilingualism in second dialect learning in adulthood, and provides an explanation in terms of perceptual strategies in which echoic memory is involved. We also contribute to the recent body of research suggesting that imitation of an action can result in improved understanding of that action.

KEYWORDS: bilingualism, cognitive advantage, phonetic and phonological learning, echoic memory, accented speech, dialects of English, second dialect learning, speech perception

1. Introduction

Acquiring the phonetic and phonological subtleties of a second language in adulthood is one of the most difficult tasks in language learning, as is apparent to the ‘naked ear’.

Foreign accents are ubiquitous in many of today’s cultures, in schools, on TV, in professional environments and in academia. North America has even seen the rise of a foreign accent culture, with stereotypical forms, some stigmatized, some held in prestige. Generally speaking, a foreign accent is at times negatively perceived in some form or other by the native listener. This process is thought to be due to a reduction in cognitive fluency, which leads listeners to attribute more negative traits to accented speakers, all other things being equal (Lev-Ari & Keysar 2010). Often, foreign accents are stigmatized and stereotyped as illustrative of a “broken” form of the language (Hosoda et al. 2007, Lindemann 2005). Accent-related phenomena such as interlanguage or fossilization have been discussed at length (Adamson 2015, Gass 2013, Selinker & Rutherford 2013, Han & Odlin 2006). Even when a speaker possesses a high degree of fluency in terms of syntax and vocabulary, as well as strong native-like intuitions in language use (e.g., the ‘Joseph Conrad’¹ phenomenon’, Hu et al. 2013, Keys 2012, Bongaerts et al. 1995), emulating a native accent is still unattainable.

At birth, all humans are equipped with the ability to learn any possible human language, and they often do learn more than one language while growing up. With 193 countries and over 6,000 languages in the world today, most countries in the world are multilingual (Wei 2008), and both bilingualism and multilingualism are commonly encountered worldwide².

¹ The reference is to the Polish-born novelist Joseph Conrad whose native-like abilities in English grammar, vocabulary and style were in stark contrast with his strongly accented pronunciation (Jilka et al. 2007).

² Counting bilinguals is a notoriously difficult task and estimates vary largely, depending on one’s definition of bilingual. While it is estimated that more than half of the world’s population uses two or more languages

In terms of sensitivity to speech sounds, Eimas et al. (1971)'s seminal study has shown that infants can distinguish all possible phonemic contrasts, an ability which adults do not possess and which appears to be significantly reduced by the end of the first year of life (Werker et al. 1981). Some people, however, show the ability to learn a second language later in life more successfully, and at a faster rate than others. The role of individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition has been explored in numerous studies (see Robinson 2001 for a synthesis of the main findings). An important issue in studying learning ability in second language learners is the presence of a number of factors that prevent us from drawing very strong conclusions regarding the ways in which proficiency was achieved in the target language. This is because proficiency has been linked to a number of extraneous factors (Levi et al. 2007), such as motivation, attitude towards the target language, age at first exposure, duration and type of exposure (Masgoret & Gardner 2003). Particularly in the context of speech processing studies, such factors should ideally be removed or controlled for; however, since no two people share the exact same experience as second language learners, this appears to be an impossible goal.

Despite this, it is possible to study individual differences in language processing in the absence of such confounding factors, by focusing on the learning of novel accents of one's native language, either real accents or, for better control, artificially created ones (Adank et al. 2011), though the latter may raise the issue of (lack of) naturalness (Cristia et al. 2012). Accent is readily separable from other linguistic skills, enabling a glimpse into the initial stages of phonetic and phonological

(or dialects) in everyday life, some of the world's largest countries and regions are predominantly monolingual – e.g., the United States and Canada, large parts of China, the Arab-speaking world, and Latin America.

learning without the difficulties generally encountered in the study of second language acquisition, such as the lack of homogeneity in terms of motivation to learn the second language, the age of the first exposure to the language, the amount and type of input received, or differing learners' attitudes towards it (Masgoret & Gardner 2003). Studying speakers' initial learning after identical exposure to novel accents offers the possibility of uncovering individual differences in language processing in the absence of such confounding factors. Thus, the observed variability in the results obtained will not reflect differing proficiency stages of the speakers and may provide a more accurate picture of the mechanism underlying initial learning.

Over the years, learners' deviations from native speaker pronunciation norms have been attributed to an inability to hear the contrasts of the target language (perceptual reasons), an inability to implement novel articulatory commands (production reasons, either purely articulatory, on the basis of articulatory complexity, or at a higher level, due to an inability to couple articulatory commands and mental representations), a failure to convert the cue pattern in the non-native sound into a licit phonological representation (representational reasons), or combinations of the above.

1.1. The phonetic/phonological “disadvantage”.

1.1.1. Perception. We first turn to the idea that the difficulty in producing a sound is due to the learner's inability to identify it perceptually. In its most extreme form, this phenomenon is known as phonological deafening, referring to the gradual decline (by the age of 1) of the ability to distinguish non-native phonological contrasts (Aslin et al. 1998, Polka & Werker 1994, Werker & Lalonde 1989, Werker & Tees 1984). While this phenomenon has been widely documented, it is somewhat of a problematic notion, in that adults ultimately *can* hear and acquire

foreign sounds. If they were really “deaf” to foreign sound categories, they would never be able to learn a second language insofar as they would be never able to hear them. In light of this, phonological deafening should not be thought of as an inability to hear foreign phonemic contrasts, but only a weakening in the behavioural response to such differences (Calabrese 2012).

1.1.2. *Production.* It has also been claimed that the difficulty in producing a non-native sound is rooted in motor difficulties experienced by learners. Thus, older second language learners are thought to experience muscle memory interference when trying to produce novel segments accurately, as long-standing habitual configurations overcome new configurations (Flege 1988). This is not irreversible, as more recent work shows that adults can be as good as children in the production and perception of at least some foreign sounds (Baker et al. 2008, Flege & MacKay 2004), and adults’ performance in a second language can improve as their length of residence in the foreign language environment increases (Flege & Liu 2001).

In the case of second dialect learning, it is often the case that the sounds to be learned already exist in the learners’ inventory, and so articulatory difficulty alone cannot provide an adequate explanation for this phenomenon. Crucially, instances of dialectal variation involving vowel merging or unmerging, devoicing, spirantization/lenition do not require learners to acquire new articulatory commands, but to modify the mapping between these and phonemic categories (Kondratenko & Spinu 2014).

1.1.3. *Representation and Linking.* From a representational perspective, second language learners’ issues with producing a novel sound is thought to stem from the general difficulty in linking a certain surface form to a certain underlying form. In the model outlined in Calabrese (2012), the difficulty faced by listeners with unfamiliar (foreign) segments can be accounted for in

terms of an active constraint against them. The foreign configurations are blocked by active marking statements of the listeners' native language. As a result, when hearing an unfamiliar linguistic configuration such as a foreign sound or uncommon linking relationship, that configuration will be blocked by an active marking statement because of its status as grammatically illicit. The grammar will prescribe the application of a repair to adjust the configuration, in perception as well as production, which can lead to adaptation or phonetic illusions. The application of such repairs will generate a representation that is grammatically correct, but mistaken. A specific example is that of the English vowel [æ] perceived by native speakers of Italian: the Italian vocalic system does not contain the [+low, -back] vowel /æ/ of the English word /kæt/ and Italians interpret this vowel as either /ɛ/ or /a/. The illicit configuration [+low, -back] is rehabilitated either by replacing [+low] with [-low] or by replacing [-back] with [+back]. In the first repair, the illicit vowel [æ] is replaced with the vowel [ɛ] and in the second one, the illicit vowel [æ] is replaced by the vowel [a], such that the Italian rendition of /kæt/ is commonly either [kɛt] or [kat].

It is the case, however, that Italian speakers can eventually learn the new sound [æ] and produce native-like forms. This, according to Calabrese (2012), is largely due to the existence of two modes of speech perception: phonemic and phonetic. When subjects perceive stimuli according to native-language phonological categories, they are demonstrating 'phonemic' perception. When they instead show sensitivity to phonetic distinctions, they are using phonetically relevant perception. This perceptual mode is also necessary for the learning of allophonic variation in the native language, as well as in order to access sound contrasts in other languages, and to acquire foreign sound configurations. For this to become possible, the auditory or echoic memory component of sensory memory (Neisser 1967) is posited to play an important part (Calabrese

2012). Sensory representations of speech uttered by others can be stored (for up to 2-3 seconds) and checked against different mental representations, until the acoustic patterns are converted into licit phonological representations (e.g., when a second language learner is able to acquire a non-native sound) or ascribed to existing phonological representations (e.g., when one becomes able to parse correctly mispronunciations due to a speech disorder). Echoic memory, which tends to be reduced in late talkers (Grossheinrich et al. 2010), is located in the auditory cortex, and forms a part of the bottom-up system of perception for collection and storage of acoustic signal (Davis & Johnsruide 2007). Whereas familiar sounds activate phonetic traces in the auditory cortex, bypassing analysis by the top-down system and directly activating word exponents in the dictionary (Hickok & Poeppel 2007, 2004), unfamiliar sounds, sound configurations or mapping relations must be interpreted within the present grammar.

According to the analysis-by-synthesis model discussed in Calabrese (2012), the listener analyzes the acoustic input by deriving how it is produced by the speaker, synthesizes a virtual acoustic signal based on the output of this derivation, and matches the virtual to the actual signal. Given a sufficiently close match, the listener achieves a mental representation of the percept that corresponds to the invariant motor commands sent to the musculature underlying the vocal tract actions that produced the acoustic signal. The analysis-by-synthesis component is part of a top-down system, while echoic memory is part of a bottom-up system of perception dedicated to the collection and storage of the acoustic signal. To solve the problem of ‘non-phonological deafness’ (the fact that adult non-native speakers can and do ultimately acquire novel contrasts in their second language), one must assume that learners have the ability to detect the acoustic details – in term of cues and other acoustic patterns – of the non-native sounds, and preserve their “aural” representation in echoic memory before they are converted and interpreted in terms of

articulatory representations. This acoustic representation is the learning target that must be analyzed and properly converted into a correct featural representation through several cognitive adjustments. According to Calabrese, the importance of echoic memory and “aural” (i. e. linguistic uninterpreted) acoustic representation of sounds has not been well appreciated in the speech perception literature, with some exceptions such as the DIVA model of speech production/perception (Guenther 1995, 2006, Perkell et al. 2004a,b) in which acoustic targets play an important part.

1.2. The bilingual advantage. “Bilingualism profoundly affects the brain, yielding functional and structural changes in cortical regions dedicated to language processing and executive function” (Crinion et al. 2006, Kim et al. 1997 cf. Krizman et al. 2012, p. 7877). Such changes may underlie the positive effects of bilingualism on nonlinguistic cognition, e.g., multitasking, selective attention (e.g., Kovacs & Mehler 2009, Colzato et al. 2008, Costa et al. 2008), cognitive function (Bialystok et al. 2007, 2004, Bialystok 1999), and delayed onset and slower rate of decline in dementia as compared to monolinguals (Bialystok et al. 2007). Furthermore, bilingualism has been linked with improved function regarding certain aspects of linguistic processing, e.g., manipulating language in terms of discrete phonemic units (Bialystok et al. 2005; Bruck & Genesee 1995), novel word acquisition (Kaushanskaya & Marian 2009, Smith 2009), and vowel space re-structuring when learning novel accents (Kondratenko & Spinu 2014). A recent study in which subjects learned vocabularies that differentiated words using foreign phonetic contrasts (Antoniou et al. 2014) reports that bilinguals possess an advantage over monolinguals in terms of phonetic learning, which is however modulated by the universal difficulty of the specific phonetic contrast to be learned and by the phonetic similarity between

the target language and the learners' native language. Other recent experimental results with non-native contrasts (Tremblay and Sabourin 2012) suggest that multilinguals and bilinguals have enhanced speech perception abilities compared to monolinguals, even though the three groups do not differ in terms of their ability to discriminate a non-native contrast before any training is received. The authors speculate, however, that more sensitive method such as electroencephalography might yield different results, as studies have shown that a contrast can be discriminated at the neurophysiological level before or without being able to do so at the behavioral level.

A recent study on auditory processing (Krizman et al. 2012) showed that subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Specifically, the subcortical response to the speech sound [da] was different in bilinguals as compared to monolinguals. Bilinguals showed a significantly larger auditory brainstem response relative to monolinguals in both a quiet and multitalker babble condition. Furthermore, while for monolinguals there was a decrease in the amplitude of the fundamental frequency (F0, 100 Hz) when the stimulus was presented in multitalker babble relative to when it was presented in quiet, no change was noted in bilinguals in F0 amplitude between the two conditions.

Despite such advantages, certain disadvantages of bilingualism related to language processing have also been reported, for instance slower lexical retrieval and reduced fluency as compared to monolinguals (Michael & Gollan, 2005).

2. Current study

To gain insight into the mechanism underlying phonetic and phonological acquisition, we examine the properties of phonetic/phonological learning upon initial exposure to a novel accent

of English by monolingual English speakers and bilingual speakers of French-English, or English and another language. The phenomenon of interest is not a foreign contrast, but the subtly different use in the novel accent of allophones already existing in the subjects' inventory.

2.1. Experiment. In a large-scale experiment conducted at Concordia University in Montreal, Canada, we aimed to explore the mechanism underlying the skill of being good with accents. English monolinguals, French-English bilinguals, and bilingual speakers of English and another language were trained on three novel accents of English, specifically Russian English, Sussex English, and South African English. While the larger goal is to become able to quantify “accentedness” acoustically (based on our measurements of production data of monolinguals and bilinguals) and perceptually (based on future perception experiments in which native speakers rate the recordings collected on their degree of accentedness), the amount of data collected and the wide array of aspects investigated (vowel space, consonant properties, intonation, etc.) lend themselves to the presentation of this work in ‘installments’. The current study focuses on a single aspect of the Sussex English accent, also known more generally as a dialect of Estuary English. The phenomenon of interest is not the learning of new contrastive categories (as is customary in studies of second language acquisition), but rather the learning of subtler cues; that is, the realization of word-final voiceless coronal stops, which is exclusive as a glottal stop in this regional dialect of English, but not in that of our Canadian subjects.

2.2. Hypothesis. Given the already demonstrated existence of a bilingual advantage in language learning, we predict that bilinguals will exhibit more effective learning of Sussex English than monolinguals. Specifically, bilinguals will show a larger (as compared to their baseline) increase of their glottalized stop rate post-exposure to the Sussex accent.

2.3. Stimuli. This study reports on the Sussex English results, and consequently our description focuses on this subset of the data exclusively. It should be added, however, that the entire procedure was repeated three times (to include the other two accents investigated, Russian and South African English) and that native speakers recorded the stimuli for all three accents. Our recorder of Sussex English was a male graduate student at Concordia University. The speaker was monolingual and raised in SE England (Sussex area), having moved to Canada (Montreal) at the beginning of his graduate studies. He was recorded in a quiet room on a laptop with an external microphone. The recordings were done using the Praat software (Boersma & Weenink, 2012) with a sampling frequency of 44,100 Hz. As we were interested in both prosodic and segmental properties, our stimuli were constructed such that they included 15 different vowels in a bVt or pVt context (all real words of English). These words were embedded in short sentences which were read with various intonation patterns, e.g., declarative, yes/no question, tag question, exclamative, as well as an introductory sentence at the beginning of each word series, as shown in (1). This resulted in a total number of 75 sentences (15 words x 5 intonation patterns).

(1) **Bit** is a noun.

This **bit** is mine.

This **bit** is mine, isn't it?

Is this **bit** mine?

This **bit** is mine!

One of the distinguishing characteristics of the Sussex dialect, the glottal stop allophone of word-final stops, was also consistently present in our speaker's production, with the voiceless word-final coronal stops being realized as glottal stops 100% of the time (Hughes & Trudgill 1997, Wells

1997), e.g., *beat* [biʔ], *bit* [biʔ], *bait* [beʔ].

2.4. Participants. The participants were 50 undergraduates at Concordia University (mean age 27). Four groups were identified post-experimentally:

- (i) Canadian English monolinguals:** 13 females, 4 males, n=17; of these, 7 speakers had been raised in Quebec, 6 in Ontario, 3 in British Columbia, and 1 in Saskatchewan.
- (ii) Quebec French-English (FE) early bilinguals:** 7 females, 5 males, n=12; all speakers were from Quebec (10 from Montreal).
- (iii) English-‘Other’ (EO) bilinguals:** 12 females, 1 male, n=13. These were speakers born in a country where they learned a different language (i.e. Ukrainian, Romanian, Hebrew, Greek, Italian, Lebanese Arabic, Punjabi, Swahili, Mauritian Creole) who moved to North America with their families in their childhood, typically between 3-14 years of age and also acquired English.
- (iv) Non-English monolinguals:** 2 females, 4 males, n=6, i.e. speakers who acquired a first language different from English and started learning English past their teens.

Two speakers were excluded from the analysis, one due to not having completed the experiment and the other due to having provided inconsistent data with regard to her language background. For this study, we analyzed the data from the first three groups, comparing the monolinguals to the larger bilingual group made up of speakers from (ii) and (iii). Because the task required subjects to learn pronunciation aspects of an English dialect, we did not include participants from group (iv), who already spoke English with non-native accents. As compensation for taking part in the experiment, all participants received extra credit in

linguistics classes they were enrolled in at the time.

2.5. Procedure. The experimental procedure was divided into several stages, as follows:

(1) *Instructions and administration of a demographic questionnaire*, asking for the subjects' age, the place where they grew up, languages spoken (and proficiency level), foreign accent exposure, and whether they had ever been diagnosed with hearing/speech disorders of any kind.

(2) *Baseline*. This stage consisted of having the subjects record 6 sentences in their native accents, prior to exposure to the experimental accents. These sentences were similar to the ones recorded by the native speakers, except that the vowels of interest were in a /hVt/ context (see the Appendix).

(3) *Training*. This stage of the experiment consisted of two parts, a receptive one and a productive (imitative) one. For the first part, each subject listened to the stimuli continuously for 5 minutes, and for the second part, the subjects were instructed to imitate each sentence one by one, immediately after listening to it. To eliminate any doubts with respect to phonemic representation, the sentences to be read out loud were also presented on the screen in orthographic transcription. The presentation of the stimuli and associated recordings was done using the ModelTalker software (Yarrington et al. 2008). Acceptable ranges for Pitch (F0) and Loudness (amplitude) were set based on microphone calibration. The Pitch range was set as $\pm 20\%$ of the average pitch of practice sentences recorded during calibration, and the Loudness range was -12 to -1 dB. A silence threshold of -45 dB was also used to detect silence at the bounds of each recording. If a recording was produced outside of the acceptable ranges for Pitch or Loudness, or if silence could not be detected at either

end, then the recording software prompted the speaker to repeat the sentence. Silence in excess of 200 ms was then trimmed, but no further manipulation of the recordings took place. When a sentence was produced appropriately, the software automatically moved on to the following sentence. Before beginning to record, the subjects had a practice session with 10 items to familiarize themselves with the procedure.

(4) *Post-Test*. At this stage, each subject re-recorded the 6 sentences from the Baseline phase, trying to reproduce to the best of his or her ability the accent s/he had just been trained on.

(5) *Final briefing and administration of a post-experimental questionnaire*, asking if any of the accents sounded familiar, what type of accent they thought each of them was, which was subjectively most difficult to reproduce, and if they had any other comments.

Stages (3) and (4) were repeated three times in all, for each of the accents of interest. The Sussex English accent was the last accent presented to each subject. Each subject was tested individually in a soundproof booth and the entire experiment (including the instructions, practice, training and testing of the three accents, and final briefing) lasted about 90 minutes (approximately 25 minutes per accent).

2.6. Analysis. Following, in part, Sumner and Samuel (2005), we manually inspected and classified the realization of word-final coronal stops in the Baseline, Post-Test and Training (Imitation) conditions as (a) canonically released (indicated by a visible release burst and aspiration-like noise after the stop closure), (b) unreleased (indicated by absence of a burst or other evidence of a release after the closure) (c) glottalized (characterized by irregular, widely spaced pulses before or after the stop closure begins or by a period of creakiness without a stop closure) and (d) tapped (signaled by an extra short closure with decreased intensity). We referred to

existing literature for these descriptions (Thomas 2011; Docherty & Foulkes 2005, 1999; Ladefoged & Maddieson 1996). A series of ANOVAs were then run to investigate the effects of Group (monolinguals and bilinguals) and Phase (Baseline, Imitation, Post-Test) on the various types of Coda Realization (released, unreleased, glottalized, tapped).

2.7. Results. Descriptive statistics with means and standard errors are provided in Table 1.

Table 1. Descriptive statistics.

		Baseline	Imitation	Post-Test
Monolinguals	Canonical release	18.1 (4.2)	45 (1.9)	26.0 (6.1)
	Unreleased	27.1 (4.5)	15.5 (2.7)	15.7 (5.3)
	Glottalized	36.8 (5.1)	71.7 (4.0)	49.4 (6.6)
	Tapped	17.0 (3.0)	7.8 (2.6)	9.3 (1.6)
Bilinguals	Canonical release	39.7 (6.1)	8.2 (2.0)	15.8 (3.7)
	Unreleased	27.2 (4.0)	19.3 (2.3)	23.7 (4.1)
	Glottalized	15.5 (3.8)	65.7 (4.2)	52.1 (5.2)
	Tapped	16.0 (3.0)	6.8 (1.7)	7.6 (1.2)

Note: means and standard deviations (in parentheses) for the two groups.

While the behaviour of the two bilingual groups was sufficiently similar for them to be collapsed into one single group in the statistical analyses, we also performed separate comparisons and present their results separately in our figures. This highlights the fact that, while group (iii), made up of English-‘Other’ bilinguals, was more heterogeneous (different language backgrounds and ages of exposure to English), the behaviour of its members did not essentially differ from that of the bilinguals coming from a more homogeneous background, specifically the French-English bilinguals from group (ii). Figure 1 presents the proportion of different coda realizations by group (Monolingual, FE Bilingual, EO Bilingual) and condition (Baseline, Imitation, Post).

Figure 1. Coda realization proportions by group and condition.

A repeated measures ANOVA was run on production rates of each coda realization (released, unreleased, glottalized, tapped) with Phase (Baseline vs. Post-Test) as a within-subject factor and Group (monolingual vs. bilingual) as a between-subject factor. For the release rates, we found no main effect of Phase or Group but found a significant interaction between Phase and Group $F(1, 40)=14.16, p=.001$, indicating that bilingual speakers' release rates decreased in the Post-Test (39.7% $\rightarrow$ 15.8%) while monolingual speakers' release rates increased (18.0% $\rightarrow$ 26.0%). For the unreleased rates, we found a main effect of Phase, $F(1, 40)=6.56, p=.014$. Both monolingual and bilingual groups decreased their unreleased rates in the Post-Test. For the glottalization rates, we found a main effect of Phase, $F(1, 40) = 42.69, p < .001$ but we also found

an interaction of Phase with Group, $F(1,40) = 10.17$, $p = .003$, indicating that bilingual speakers significantly increased the glottalization rates in the Post-test (15.5% → 52.1%) while monolingual speakers did not (36.8% → 49.4%). Lastly, for the rates of tap, we only found a main effect Phase, $F(1,40) = 13.0$, $p = .001$. Both groups decreased the tap rates in the Post-Test. Next, we compared the distribution of final coronal stop realizations during the Imitation phase between the two groups in order to confirm that glottalization was the most predominant realization of final stop during the Imitation phase. A repeated measures ANOVA was run on production rates with Coda Realization (released, unreleased, glottalized, tapped) as a within-subject factor and Group (monolingual vs. bilingual) as a between-subject factor. A main effect of Coda Realization was found, $F(3, 120) = 158.9$, $p < .001$. Pairwise comparisons showed that glottalization rates were significantly higher than released stops ($p < .001$), unreleased stops ($p < .001$) and tap rates ($p < .001$). Thus, during the training, both groups produced glottalization significantly more frequently than other production types, indicating that the majority of speakers successfully imitated the Sussex English speaker's glottalization at this stage. There was neither a main effect of Group nor a significant interaction of Group with Coda Realization, suggesting that each group's production of final [t] during the training was not significantly different from the other group's.

Turning to the individual results, Figure 2 presents the percent glottalization by individual subject by group and phase. As indicated by the significant interaction (between Baseline and Post) we found in the repeated measures ANOVA, the separation of the baseline from the imitation and post condition is larger in the case of the bilingual groups (36.6% average distance between Baseline and Post), while most of the monolinguals displayed a tendency to return to their baseline after training (12.6% average distance between Baseline and Post). Of the

monolinguals, subjects 19 and 21, and to a lesser extent 30 and 32, show the largest separation between Baseline on the one hand, and Imitation and Post-training, on the other hand. These results suggest that people who are ‘good with accents’ are present in both groups; however, they are more numerous in the bilingual group.

Figure 2. Glottalization percentage (out of overall word-final voiceless coronal stop production) by individual subject and condition.

3. Discussion

To summarize, our results show a difference in behaviour between the Baseline (baseline accent)

and Post-Test (spontaneous production of the Sussex accent after training) for both groups, as well as a marked difference between the monolingual and the bilingual group. In the Post phase, monolinguals exhibited a small increase in glottalization as well as stop release rates, and a small decrease in unreleased and tap rates. This pattern, however, was only present in 4 of the 17 monolinguals examined. The behaviour of bilinguals was more consistent in this respect, with most of the speakers having increased their glottalization rates post-training. Not only did the bilinguals exhibit significantly higher glottalization rates in Post versus Baseline, they also maintained the glottalization rates from Imitation in their Post phase. These results may be interpreted as demonstrating more effective learning of the stop realization pattern in Sussex English on the part of the bilinguals, thus supporting our hypothesis. Both groups performed equally well in the Training phase, producing word-final voiceless coronal stops as glottals over 65% of the time.

With respect to the baseline differences between the two groups, specifically to the higher rates of glottalization observed with monolinguals, these may be partially underlain by a certain degree of language transfer in the case of bilinguals. Half of our bilinguals had French as one of their languages, and the occurrence of the glottal stop in French is restricted in general, before a word starting with aspirated /h/, particularly when it is preceded by a consonant-final word, and also in vowel-initial words (Tranel 1981). Two other uses of the glottal stop in French are described in Malécot (1974): (a) in utterance-initial and -medial position serving as a stress marker, and (b) medially and in utterance-final position, serving to call attention to a preceding or following element or to abort an unwanted utterance - both environments being relatively restricted and not limited to word-final voiceless coronal stops. On the other hand, it is very common to have glottalization in both British and American English (Selkirk 1980).

Another baseline difference consisted of not only a lower glottalization rate, but also a higher rate of stop release for the bilingual EO group. While it is difficult to speculate on the reasons underlying this difference because of this group's heterogeneity, language transfer is likely a factor. Few of the other languages spoken by our subjects include glottal stops in their inventory, while the majority release their stops in all positions, unreleased stops being generally considered more marked, because the release burst of word-final stops was found to be the most salient cue to place of articulation (Malécot 1958 cf. Gundel et al. 1988).

To sum up, the baseline differences in coda realization may be motivated by the fact that some of the Canadian English speakers behave like American English speakers in this respect, whereas bilingual speakers may generally exhibit language transfer from French (as well as the other languages) in their production of English.

Regardless of whether the Baseline differences are intrinsic to the dialect spoken, the possibility arises that our results might be interpreted not as demonstrating a bilingual advantage in accent learning, but rather as a bilingual ability to adapt and compensate for an initial disadvantage (i.e. at Baseline). This is a plausible interpretation, the validity of which should be clarified in future work. As far as our specific results are concerned, we believe several aspects may argue against it. First, monolinguals did significantly increase their glottalization rates in Imitation, showing that despite higher glottalization in Baseline there was still room for learning, but returned to their Baseline pattern after training. Second, monolinguals also increased their release rates in Post, which is unexpected in light of the Sussex accent displaying 100% glottalization in this context. Both of these cast some doubt on monolinguals' initial learning effectiveness. And finally, even though a significant difference was found between the two groups' baseline glottalization rates, inspection of the individual patterns (see Figure 2) reveals that while three

of the monolingual subjects produced high rates of glottalization in baseline (60-73%), the majority of monolinguals were not very different from the bilingual group at this stage. Despite the initially higher distribution of glottalization in monolinguals, it was conceptually still possible for them to further increase it in order to reproduce the Sussex accent successfully - in practice this was achieved by a couple of the monolingual subjects. The overall group, however, was not as successful at departing from baseline as the bilingual group. It is a distinct possibility that starting out with a lower rate of glottalization may have allowed bilinguals more 'room to converge', while monolinguals with high rates of glottalization may have considered their speech as similar to Sussex with respect to the realization of word-final coronal stops. This, however, does not explain why the monolinguals who produced lower rates of glottalization in Baseline did not converge after Training. If we consider the individual results in Figure 2, we find that one of the most successful monolingual learners (Subject 19) had very little glottalization at baseline, which appears to support this view. However, Subject 25, who also started out with reduced glottalization, returned to baseline in the post-test, and Subject 21, another successful learner, started out with a high glottalization rate at baseline. This was also the case for successful bilingual learners such as Subject 47 in the FE group and Subjects 13 and 17 in the EO group. These findings make it difficult to draw a strong conclusion as to how baseline differences may have affected learning, an issue that should definitely be addressed in future research.

One of the main questions arising in light of our results is to what factors we may ascribe monolinguals' behavior, specifically their return to baseline after training, at least as far as this aspect of Sussex English is concerned. We started out with three possible explanations, based on articulatory, perceptual, or representational grounds. The articulatory difficulty of the segment is likely not a factor here, as the glottal stop is present as an allophone of coronal stops (most often

phonemically voiceless) in North American English (Selkirk 1980), and moreover, it was present in the monolingual group's baseline condition 36.8% of the time. Representational issues are also not likely to have played a major part at the phonemic level, because our speakers had access to the underlying representations (via orthographic transcriptions). As for a possible perceptual explanation, the monolinguals' behaviour during the training phase and their ability to imitate the increased glottalization rates of the Sussex accent suggest that this cannot be the main reason behind the monolinguals' deviation from the target accent either. The only viable explanation appears to be in terms of linking. Specifically, a subtler mechanism could have played a part, that is, the subjects' were not able to modify the mapping between an existing representation and existing articulatory instructions. From the perspective of Calabrese's model of speech perception (2012), we may interpret this as monolinguals being more prone than bilinguals to by-passing phonological (re)analysis and directly activating word exponents in the dictionary.

This provides the starting point for our examination of what might account for bilinguals' larger increase in glottalization rates as a result of training. If, as compared to monolinguals, bilinguals have an increased ability to modify the mapping between existing representations and existing articulatory instructions, it is not implausible for this ability to be underpinned by differing memory strategies/capabilities. According to Calabrese (2012), the conversion of a new sound pattern into a correct featural representation is underlain by echoic memory processes. The results of Krizman et al. (2012) suggest that there is a bilingual advantage in the subcortical encoding of sound. While to our knowledge this specific connection remains unexplored, more precise subcortical timing was linked to better reading abilities (Banai et al. 2009, Hornickel et al. 2009, cf. Kraus & Slater 2016) and good readers were shown to have superior echoic memory (Sipe & Engle 1986). Furthermore, the left-hemispheric specialization supporting rapid acoustic

processing was linked to the precision of subcortical encoding of speech sounds (Kraus & Slater 2016). Thus, there is a possibility that the subcortical encoding of sound is linked to short-term memory processes. Specifically, stronger subcortical encoding may result in longer availability of acoustic information in echoic memory, activating the phonetic mode of speech perception and leading to more efficient sound processing. A direct result of this would be enhanced encoding strategies, e.g., more time for phonological (re)analysis. This would enable bilinguals to successfully learn new mapping configurations, such as the Sussex word-final /t/ being realized as a glottal stop. Bilinguals have been shown to outperform monolinguals in tasks involving episodic memory recall (Ljungberg et al. 2013), but the question of whether their sensory memory, particularly phonologically related (echoic memory), presents advantages remains open. Our results indicate that this may be the case. While no causal relationships between subcortical encoding of sound, echoic memory, and enhanced phonetic/phonological learning may be established on the basis of the findings reported here, our work – by bringing them to the foreground – represents a step further in our understanding of speech processing in general and the bilingual advantage in particular.

Because approximately half of our bilingual group consisted of French-English bilinguals, the question arises whether being a speaker of French might have provided an advantage that would explain the bilinguals' success with the new mapping pattern between underlying and surface representations. At least two points would argue against this. First, as we have seen, the EO bilingual group behaved similarly to the FE group. Second, as discussed previously, glottal stops are relatively restricted in French in general. In Canadian French specifically, glottal stops can occur between vowels (Walker 1984), but not as allophones of consonants. We tentatively conclude that it is not the specific other language spoken by bilinguals that played a part here but

the very fact of being a bilingual, a conclusion already supported by the fact that both our bilingual groups behaved similarly, and other types of bilinguals (i.e. Mandarin-English and Korean-English bilinguals) were also shown to have an advantage in phonetic learning (Antoniou et al. 2014).

Another question following from our results concerns the development of the mechanism through which bilinguals become effective learners of novel accents, exhibiting increased echoic memory and linking strategies. At an early point in language development, bilinguals are faced with the task of learning multiple sound-UR mapping patterns for different languages (e.g., French-English bilinguals learn that short positive VOT corresponds to phonemically voiced stops in English but voiceless in French³). To do so, they must develop a strategy of some sort, related to executive function (that is, involving the mechanisms underlying attention, memory, and encoding). Presumably, the early/simultaneous acquisition of two languages, when the same sound can function differently in each language, engages more frequent use of the phonetic mode of speech perception, thus providing them with a better ‘workout’ of their echoic memory and short-term to long-term conversion strategies, all of which act as a (more or less) effective ‘anti-phonological deafening’ tool. Recent work by Liu & Kager (2015) shows the existence of perceptual turbulence for bilingual infants at 8-9 months, but stabilized perception for both monolingual and bilingual infants from 11 months onwards. Smith (2009) notes that adults who

³ According to Flege’s Speech Learning Model (Flege 1995) there is an expectation that corresponding members of the sound inventories of bilingual speakers (such as voiceless stops in French and English) are identified with one another at the cognitive level. This hypothesis was supported by the results of Fowler et al. (2008), who found that French-English bilinguals, while having distinct voiceless stops in English and French, nevertheless produce VOT patterns that differ from those of monolingual English and French speakers, failing to replicate Sundara et al.’s (2006) study in which no differences were found between the [t] of French bilinguals and monolinguals. While Fowler et al.’s (2008) findings support the existence of cross language phonetic influences in the speech of bilingual speakers, they may not be sufficient for drawing conclusions regarding perceptual or representational matters.

experience learning an L2 in early childhood approach new language learning more similarly to children than adults with later L2 learning experience: “In a sense, through their experience learning an L2, people learn *how* to learn languages” (p. 87). By contrast, monolinguals do not have to develop or exercise these mechanisms to the same extent. We have seen, however, that the new pattern was learned successfully by at least two of the monolingual subjects, though it remains unclear to what extent they may have used the same strategies as bilinguals. The question of whether these resources can be tapped into later in life or whether these skills can be learned in different ways (for instance, through various types of discrimination exercises or by certain types of music training⁴) should be addressed in future studies. Several successful speaker strategies have been examined in the case of foreign accented speech, including the reorganization or maximization of vowel space, the use of multiple cues to distinguish certain segmental contrasts, and being able to employ prosody efficiently (Skoruppa & Peperkamp 2010, Bent et al. 2008, Kluge et al. 2008, Gavac 2002). Future production and perception studies should determine the extent to which some of these strategies are preferred and selectively used by bilingual learners in their second dialect/second language learning and whether they lead to more effective or more native-like communication. Identifying such strategies may result in practical applications such as the development of more effective pronunciation teaching tools for second language learners.

One direct follow-up to our study would be to compare echoic memory in monolinguals and bilinguals, as reliable experimental paradigms already exist and could be tested using tasks such as delayed tone matching (Javitt et al. 1997) or examining suffix effects (Watkins & Todres

⁴ Neural and cognitive advantages in the processing of auditory stimuli were also reported for musicians, similar to bilinguals (Kraus & Chandrasekaran 2010, Wong et al. 2007).

1980). Several other methods for echoic memory testing are described and compared in Crowder (2014). A further direction would be to pursue the connection between bilingualism and reading skills, as experimental studies have shown superior echoic memory processes in good versus poor readers: whereas good and poor readers recalled the test items equally if tested immediately after presentation, the poor readers showed a faster decline in recall of the items as retention interval increased (Sipe & Engle 1986).

As a potential limitation to our study, the inclusion of an imitation condition as the basis for training may be criticized. It has been suggested in the literature that direct imitation does not reflect linguistic skill, as it produces behaviour which temporarily exceeds actual competence, with no “carry-over into the post-imitative tasks” (Barry et al. 1989), and this appears to accurately describe the behaviour of the monolinguals in our study. More recently, however, Kuhl & Meltzoff (1995) showed that imitation can result in nonidentical, but functionally equivalent reproductions of the modeled behaviour, against the view that direct imitation bypasses all levels of linguistic processing. Recent studies have further strengthened this point by demonstrating that imitation improves language comprehension (Adank et al. 2010), and can lead to fast modification of both phonetic and phonological intonation representations (D’Imperio & German 2015, D’Imperio et al. 2014). Ultimately, as pointed out by Markham (1997), acquisition itself is a strongly imitative phenomenon.

Our study adds to the body of work using direct imitation in language tests (D’Imperio et al. 2014, Delvaux et al. 2011, Markham 1997, Neufeld 1987), and suggests that imitation may play different roles in the initial learning of phonetic/phonological patterns, depending on a speaker’s status as either monolingual or bilingual. This conclusion must nevertheless be interpreted with caution because our subjects also received passive (listening) training. Future studies should aim

to tease apart the effects of the two training methods more rigorously.

Another potential drawback of our study is intrinsic to the experimental design. Our subjects were not asked to focus on any particular cue, including word-final glottalization of coronal stops, and it is a distinct possibility that while some subjects attended to certain cues, a number of different cues were more salient to other subjects. Perceptually, prosody and intonation were shown to influence listeners' impressions of foreign accents (Jilka 2000) but it has been suggested that the effects of segmental and syllable structure information on the perception of accentedness relative to prosody may be dependent on native language (Anderson-Hsieh et al. 1992). Indeed, recent work with Korean-accented English has shown that while segmental and prosodic information have different weights in the perception of foreign-accented speech (Serenio et al. 2016), segmental information contributes substantially more to the perception of foreign accentedness. Depending on our subjects' native language, different cues may have been consciously perceived as more salient and deliberately pursued in accent imitation and regeneration. The complete picture – and a clear, bilingual advantage – will only emerge once additional cues to 'nativeness' for the Sussex dialect have been analyzed. Preliminary results with our data suggest that, while very minimal, monolinguals may display an advantage in spontaneously recreating the Sussex speaker's intonation patterns.

4. Conclusion

Our main result was that, following brief initial exposure and training on a novel accent of English, bilinguals made more progress than monolinguals concerning at least one phonetic (and phonological) aspect, which could be a result of cognitive advantages related to phonetic and phonological processing. This conclusion needs to be taken with a grain of salt, however, given

the existence of baseline differences between the two groups. Even though these differences may be intrinsic to the dialects spoken by our subjects, they have affected the strength of the statistical results and thus our conclusion can only be tentative. Nevertheless, the big picture painted by our study suggests that bilinguals are indeed more effective learners of the word-final glottalization patterns characteristic of Sussex English. This assertion is based on the fact that (1) a significant difference between Baseline and Post-training was found in glottalization rates for bilinguals but not monolinguals, (2) monolinguals increased their stop release rates in the Post condition, and (3) individual behavior (Figure 2) shows that most of the monolinguals returned to their baseline after training, whereas bilinguals maintained the glottalization rates produced during the training session. To validate this conclusion, further perceptual experiments with the collected data are planned in which native speakers of the accents tested are going to rate our subjects in terms of how native they sound. A second result was that imitation facilitates short-term phonetic/phonological learning (Adank et al. 2010). Our findings add empirical evidence to the body of work on the bilingual advantage, and also on the mechanism underlying phonetic and phonological learning. We have further speculated that the bilingual advantage is due to more efficient coupling of sensory and cognitive functions, and suggested that echoic memory may play an important role by facilitating the re-mapping between existing mental representations of sounds and existing articulatory command configurations, an idea which may also be supported by independent findings of stronger subcortical encoding of sound by the bilingual brain (Krizman et al. 2012).

Appendix

The following sentences were used in the Baseline (baseline) and Post-test (experimental) conditions.

The words in bold end in voiceless coronal stops, which were analyzed for the purposes of this study.

- (1) I **hate** your **coat** and that **hat** on your head.
- (2) He **got hit** and badly **hurt** in the hood.
- (3) I don't know how you can stand the **heat** in your **hut**.
- (4) Tell me who thinks he is too **hot** to say hi.
- (5) They had run **out** of **wheat** and **oat** at the farm.
- (6) He **ought** to have **caught** the **goat** instead of being coy and letting it go.

References

- Adamson, H. D. (2015). Interlanguage: forty years later. In ZhaoHang Han and Elaine Tarone (Eds.), *Studies in Second Language Acquisition* 37.01. Amsterdam, the Netherlands: Benjamins: 167–168.
- Adank, P., Noordzij, M. L. & Hagoort, P. (2012), The role of planum temporale in processing accent variation in spoken language comprehension. *Human Brain Mapping*, 33: 360–372. doi:10.1002/hbm.21218
- Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science* 21(12). 1903–1909.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42, 529–555.
- Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, 18(4), 683–695.
- Aslin, R. N., Jusczyk, P. W. & Pisoni, D. B. (1998). Speech and auditory processing during infancy: constraints on and precursors to language. In *The Handbook of child psychology: cognition, perception, and language*, ed. by D. Kuhn & R. Siegler, 147–254. New York: Wiley.
- Baker, W., Trofimovich, P., Flege, J.E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language & Speech*, 51, 316–341.
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S. & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex* 19: 2699–707.

- Barry, W. J., Heoquist, C. E. & Nolan, F. J. (1989). An approach to the problem of regional accent in automatic speech recognition. *Computer Speech and Language* 3, 355–366.
- Bent, T., Bradlow, A. R. & Smith, B. L. (2008). Production and Perception of Temporal Contrasts in Foreign-Accented English. *Phonetica* 65 (3), 131–147.
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development* 70. 636–644.
- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology & Aging* 19. 290–303.
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading* 9. 43–61.
- Boersma, P. & Weenink, D. (2012). Praat: Doing phonetics by computer [Computer program]. Version 5.3.19, retrieved 25 June 2012 from <http://www.praat.org>
- Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis. *The age factor in second language acquisition*, 30–50.
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy* 2. 29–49.
- Bruck, M., & F. Genesee. (1995). Phonological awareness in young second language learners. *Journal of Child Language* 22. 307–324.
- Calabrese, A. 2012. Auditory representations and phonological illusions: A linguist's perspective on the neuropsychological bases of speech perception. *Journal of Neurolinguistics* 25: 355–381.
- Colzato L.S., Bajo, M.T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W.,

- & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory and Cognition* 34. 302–312.
- Costa, A., Hernández, M. & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition* 106. 59–86.
- Crinion J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green, D. W. & Price, C. J. (2006). Language control in the bilingual brain. *Science* 9. 1537–1540.
- Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic Processing of Accented Speech Across the Lifespan. *Frontiers in Psychology*, 3, 479. <http://doi.org/10.3389/fpsyg.2012.00479>
- Crowder, R. G. (2014.) Echoic memory and the study of aging memory systems. In *New Directions in Memory and Aging (PLE: Memory): Proceedings of the George A. Talland Memorial Conference*. Psychology Press, 181–204.
- Davis, M. H. & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research* 229: 132–147.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly* 39. 379–397.
- D’Imperio, M. & German, J. S. (2015). Phonetic detail and the role of exposure in dialect imitation. *Proceedings of ICPhS 2015*, Glasgow, 10-14 August.
- D’Imperio, M., Cavone, R. & Petrone, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in Psychology*, pages, doi: <http://dx.doi.org/10.3389/fpsyg.2014.01226>.

- Dogil, G., & S. Reiterer. (2009). Language Talent and Brain Activity. *Trends in Applied Linguistics* 1. Mouton de Gruyter: Berlin-New York.
- Docherty, G. J., & Foulkes, P. (2005). Glottal variants of /t/ in the Tyneside variety of English. In W. J. Hardcastle, & J. Mackenzie Beck (Eds.), *A figure of speech: A festschrift for John Laver* (pp. 173–197). London: Lawrence Erlbaum.
- Docherty, G. J., & Foulkes, P. (1999). Newcastle upon Tyne and Derby: Instrumental phonetics and variationist studies. In P. Foulkes, & G. J. Docherty (Eds.), *Urban voices* (pp. 47–71). London: Arnold.
- Flege, J. E. (1995). Second-language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Timonium, MD: York Press, 233–273.
- Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders: a review*. Norwood, NJ: Ablex, 224–401.
- Flege, J. & MacKay, I. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1–34.
- Flege, J. & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23, 527–552.
- Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics*, 36(4), 649–663. <http://doi.org/10.1016/j.wocn.2008.04.001>
- Gass, S. M. (2013). *Second language acquisition: An introductory course*. Routledge.
- Gavac, N. (2002). *Acoustic Measurement of Canadian English-accented French*.

B.A.Thesis, Simon Fraser University.

Grossheinrich, N., Kademmann, S., Bruder, J., Bartling, J. & Von Suchodoletz, W. (2010).

Auditory sensory memory and language abilities in former late talkers: A mismatch negativity study. *Psychophysiology*, 47: 822–830. doi: 10.1111/j.1469-8986.2010.00996.x

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102, 594–621.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds, *Journal of Communication Disorders* 39: 350–365.

Gundel, J. K., Houlihan, K. & Sanders, G. (1988). On the functions of marked and unmarked terms. In M. Hammond, E. A. Moravcsik and J. R. Wirth (eds.) *Studies in syntactic typology*, Amsterdam: John Benjamins, 285–301.

Han, Z., & Odlin, T. (Eds.). (2006). *Studies of fossilization in second language acquisition* (Vol. 14). Multilingual Matters.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67–99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech perception. *Nature Reviews Neuroscience*, 8, 393–402.

Hornickel, J., Skoe, E., Nicol, T., Zecker, S. & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *PNAS* 106: 13022–27.

Hosoda, M., Stone-Romero, E. F. & Walter, J. N. (2007). Listeners' cognitive and affective reactions to English speakers with standard American English and Asian accents. *Perceptual*

- and Motor Skills*, 104, 307–326.
- Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: behavioural predictors and neural substrates. *Brain and language*, 127(3), 366–376.
- Hughes, A. & Trudgill, P. (1997). *English Accents and Dialects. An Introduction to Social and Regional Varieties of English in the British Isles*. London/New York: Arnold. 3rd ed.
- Javitt, D. C., Strous, R. D., Grochowski, S., Ritter, W. & Cowan, N. (1997). Impaired precision, but normal retention, of auditory sensory (“echoic”) memory information in schizophrenia. *Journal of Abnormal Psychology*, Vol 106(2), 315–324.
- Jilka, M. (2000). The contribution of intonation to the perception of foreign accent. Doctoral dissertation, University of Stuttgart.
- Jilka, M., H. Baumotte, N. Lewandowski, S. Reiterer, & G. Rota. (2007). Introducing a Comprehensive Approach to Assessing Pronunciation Talent. In *Proc. 16th ICPhS Saarbrücken*, 1737–1740.
- Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin and Review* 16(4). 705–710.
- Keys, K. J. (2012). Interlanguage phonology. Theoretical questions and empirical data. *Revista Linguagem & Ensino*, 5(1), 75–91.
- Kim, K. H. S., Relkin, N.R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature* 388. 171–174.
- Kluge, D. C., Reis, M. S., Nobre-Oliveira, D., & Rauber, A. S. (2008). Intelligibility of accented speech: The perception of word-final nasals by Dutch and Brazilians. *Proceedings of V Jornadas en Tecnologia del Habla*, Bilbao: Universidad del Pais Vasco,

199–202.

Kondratenko, Y. & Spinu, L. (2014). Being “better” with accents: evidence from bilinguals.

Proceedings of the Forth-eighth Annual Meeting of the Chicago Linguistic Society, eds.

Beltrama, A., Chatzikonstantinou, T., Lee, J.L., Pham, M., & Rak, D. 387–400.

Kovacs, A., & J. Mehler. (2009). Cognitive gains in 7-month-old bilingual infants In *Proc. The*

National Academy of Sciences 106 (16). 6556-6560. Kraus N., & Chandrasekaran

B. (2010). Music training for the development of auditory skills. *Nat Rev Neurosci* 11: 599–605.

Kraus N., & Slater, J. Beyond Words: How Humans Communicate Through Sound.

Annual Review of Psychology : 83–103. doi: 10.1146/annurev-psych-122414-033318.

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of

sound is enhanced in bilinguals and relates to executive function advantages. In

Proceedings of the National Academy of Sciences 109. 7877–7881.

Kuhl, P., & Melzoff, A. (1995). Vocal learning in infants: Development of perceptual-motor links

for speech. In *Proceedings of the international conference of Phonetic Sciences*, Stockholm,

146–149.

Ladefoged, P. & Maddieson, I. (1996). The sounds of the world’s languages. Cambridge, MA:

Blackwell.

Lev-Ari, S. & Keysar, B. (2010). Why don’t we believe non-native speakers? The influence of

accent on credibility. *Journal of Experimental Social Psychology*, 46, 1093–1096.

Levi, S. V., Winters, S. J. & Pisoni, D. (2007). Speaker-independent factors affecting the

perception of foreign accent in a second language. *Journal of the Acoustical Society of*

America 121 (4).

- Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America* 29. 117–123.
- Lindemann, S. (2005). Who speaks ‘broken English’? US undergraduates’ perception of non-native English. *International Journal of Applied Linguistics*, 15, 187–212.
- Liu, L. & Kager, R. (2015). Bilingual exposure influences infant VOT perception. *Infant Behaviour & Development* 38, 27–36.
- Ljungberg, J.K., Hansson, P., Andrés, P., Josefsson, M., & Nilsson, L.G. (2013.) A Longitudinal Study of Memory Advantages in Bilinguals, *PLoS ONE* 8(9): e73029. doi:10.1371/journal.pone.0073029
- Malécot, A. (1974). The Glottal Stop in French. *Phonetica* 31(1): 51–63.
- Markham, D. (1997). Phonetic imitation, accent and the learner. PhD Dissertation. *Travaux de l’Institut de Linguistique de Lund* 33. Lund University Press.
- Masgoret, A.-M. & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning* 53, 167–210.
- McAllister, R., Flege, J. E. & Piske, T. (2002). Influence of L1 on acquisition of Swedish quantity in Spanish, English and Estonian. *Journal of Phonetics* 30. 229–258.
- Moyer, A. (1999). Ultimate attainment in L2 phonology. The critical factors of age, motivation and instruction. *Studies in Second Language Acquisition* 21. 81–108.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., et al. (2004a). Cross-subject correlations between measures of vowel production and perception. *Journal of the Acoustical Society of America*, 116(4 Pt. 1), 2338–2344.

- Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., et al. (2004b). The distinctness of speakers' /s-sh/contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research*, 47, 1259–1269.
- Polka, L., & Werker, J.F. (1994). Developmental changes in the perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20. 421–435.
- Sipe, S., & Engle, R. W. (1986). Echoic memory processes in good and poor readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Jul;12(3): 402–412.
- Selinker, L. & Rutherford, W. E. (2013). *Rediscovering interlanguage*. Routledge.
- Selkirk, E. O. (1980). *The phrase phonology of English and French*. New York: Garland.
- Sereno, J.A., Lammers, L., & Jongman, A. (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics*, 37, 303–322. DOI:10.1017/S0142716414000575
- Skoruppa, K. & Peperkamp, S. (2011). Adaptation to novel accents: feature-based learning of context-sensitive phonological regularities. *Cognitive Science*, 35, 348–366.
- Smith, A.H. (2009). *Language Learning in Adulthood: Why some have more trouble than others*. Doctoral Dissertation, Stanford University.
- Thomas, E. R. (2011). *Sociophonetics: An introduction*. Basingstoke: Palgrave Macmillan.
- Tranel, B. (1981). *Concreteness in Generative Phonology: Evidence from French*. University of California Press.
- Tremblay, M.-C., & Sabourin, L. (2012). Comparing behavioral discrimination and learning

- abilities in monolinguals, bilinguals and multilinguals. *Journal of the Acoustical Society of America*, 132(5), 3465–3474.
- Walker, D. C. (1984). *The Pronunciation of Canadian French*. University of Ottawa Press.
- Watkins, M. J., & Todres, A. K. (1980). Suffix Effects Manifest and Concealed: Further Evidence for a 20-Second Echo. *Journal of verbal learning and verbal behaviour* 19.
- Wei, L. (2008). Research Perspectives on Bilingualism and Multilingualism, in *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism* (eds L. Wei and M. G. Moyer), Blackwell Publishing Ltd., Oxford, UK. doi: 10.1002/9781444301120.ch1
- Wells, J. (1997). What is Estuary English? *English teaching professional* issue 3, pp. 46–47.
- Werker, J.F., & Lalonde, C.E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental Psychobiology* 24 (5). 672–683.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development* 7. 49–63.
- Wong P. C. M, Skoe E., Russo N. M., Dees T., & Kraus N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience* 10: 420–422.
- Yarrington, D., Gray, J., Pennington, C., Bunnell, H. T., Cornaglia, A., Lilley, J., Nagao, K., & Polikoff, J. B. (2008). ModelTalker Voice Recorder: An interface system for recording a corpus of speech for synthesis. *Proceedings of the ACL-08: HLT Demo Session*(pp. 28–31). Columbus, OH.