"Speakers who routinely use more than one language may not use any of their languages in ways which are exactly like that of a monolingual speaker." - Schmid, 2010: 1

Introduction

Learning an additional language is a challenging process as an adult. Indeed, it has been argued that it is impossible to change the established patterns of a native language (L1), or at least an undertaking of extraordinary difficulty (Herschensohn, 2000). Despite this, growing pressure in the globalised world of the 21st century pushes individuals, instructors, and researchers to discover ways to learn a new language (L2) or adjust one’s native tongue to an additional dialect (D2).

Adult language students anecdotally report that one of the most difficult parts of L2 learning is adjusting to novel phonetic contrasts in vowels and consonants. Consonants vary in world languages, often in terms of Voice Onset Time (VOT). First investigated in 1964 with a cross-linguistic acoustic study (Lisker and Abramson, 1964) VOT was defined as “the temporal relation between the onset of glottal pulsing and the release of the initial stop consonant” (Abramson, 1975: 17). This thesis aimed to investigate VOT in context of language learning, specifically addressing how adult learners might tackle learning a new consonant contrast.

VOT-based categorical boundaries are immediately apparent to native speakers of the language in question, despite the fact that the categorical differences between stop consonants are measured in milliseconds (Aslin et al., 1981). A categorical boundary is thus named because of the instant and non-continuous distinction between consonants based on (Serniclaes et al., 2001). Categorical VOT boundaries are phonemically contrastive, allowing speakers to both produce and perceive words with different meanings that otherwise do not vary phonetically.

These boundaries are also crucial to speech segmentation and the ability to both produce and perceive fluent speech in an L2 (Best, 1995; Kuhl & Iverson, 1995). Speakers become tuned to the boundary in their native language from infancy. The distinction between consonants is a very early, very persistent aspect of early childhood language acquisition (Aoyama et al., 2004). More than 40 years ago, it was demonstrated that infants can discriminate small differences in VOT that are the basis of phonemic distinctions (Eimas et al., 1971). For naive L2 learners, perceiving the difference between words that form minimal pairs on the basis of VOT alone (e.g. ‘bet’ vs. ‘pet’) is often the root of comprehension problems.

Oral stops in particular vary along a spectrum. The difference in duration of VOT is the cue that distinguishes a voiced consonant (b, d, g) from an unvoiced one (p, t, k). The boundary between such consonants can differ between languages considerably in placement and duration (Lisker and Abramson, 1964). VOT is dependent in complex ways on the context of the consonant, especially in spontaneous speech (Gottfried, 1984; Yao, 2009). Although the absolute VOT values may vary slightly in spontaneous speech, the categorical boundary is preserved by the overall difference in VOT between contrasting consonants (resulting in categorical perception of two distinct consonants). This experiment was designed to investigate what happens when adult speakers attempt to learn a novel contrast based on VOT (French /p/-/b/, which differ from English /p/ and /b/ in terms of VOT), and how they might adapt their production and perception to a contrast not found in their native language.

Literature Review

Is it possible for adults to learn a new consonant contrast? This question is central to the present study. To answer, it is necessary to investigate the implications of age and the possibility of a critical period for language acquisition. Previous research has shown that various non-linguistic factors are involved in L2 learning, with age being the most well-established (Herschensohn, 2000; Flege, 2005). Despite its long history as a notion in psychological and linguistic sciences, the critical period for language learning is not as airtight as it might have appeared in the past.

The critical period for language was conceived as a universal period in development during which humans appear primed for language acquisition. Usually, it is thought that this period ends at the onset of puberty or around thirteen years of age (Tikofsky, 1968; Komarova and Nowak, 2001). Critical periods for development are well-documented in biological and evolutionary science. These are defined as periods during which disruption can have negative and permanent developmental consequences for an organism (Rice and Barone Jr, 2000).

External factors (e.g. malnutrition, disease) can instigate biological changes in the brain and body which may last to adulthood (Glewwe and King, 2001; Klemp and Weisdorf, 2012). Critical periods must exist in human development, but how this may be applied to language acquisition and adult L2 learning is murkier.

Some insist that the human brain 'crystallises' at the end of the posited critical period. According to proponents of this idea, any linguistic change (such as learning a French /p/-/b/ contrast) should be impossible after such a 'crystallisation;' the plasticity necessary to acquire new phonetics, phonology, syntax, and morphology would be lacking (Pallier et al., 2003). Adult learners often reveal that they are not native speakers with subtle deviations from native speech, some of which may operate on subconscious levels. Arguably, VOT is one such feature. It can be very difficult for a learner to produce or perceive consonant contrasts in a native-like way in their target language despite years of training.

Nevertheless, some adult learners appear able to master a D2/L2, even though they are rare (Flege, 2005). Indeed, L2 acquisition seems to mirror that of L1 depending on individual ability and under the right circumstances. L2 learners may be able to access the same mechanisms and processes by which they acquired their first language throughout life (Flege, 2005).

Based on all this evidence, the reasonable claim would be that a critical language period does exist, but in a weaker formulation than some would assert. This is crucial to the success of this experiment, as all of the participants would be considered well beyond the critical period for language acquisition. Participants in the present study should have been able to learn the new contrast from French despite their age.

A particular study from which this experiment drew inspiration is very recent. Adult learners were able to acquire new allophones and generalise their limited knowledge to novel words in a second dialect (German, Carlson and Pierrehumbert, 2013). The authors investigated D2 acquisition under extreme time constraints (approx. 20 minutes). Specifically, they looked at what categorical changes on the phonological level imitation of another dialect might instigate (German, Carlson and Pierrehumbert, 2013). Using two allophones from Glasgow English (/t/ and /r/, which are very distinct in placement from the same in American English), the experimental stimuli were created. The 192 stimuli provided an equal distribution of the word-initial and word-medial conditions in which the D2 allophones occur (German, Carlson and Pierrehumbert, 2013: 233). All target words were in sentence-final position.

Undergraduates from Northwestern University participated in two sessions of accent imitation recordings. The experimental procedure was as follows; a baseline D1 recording was made of the participants speaking in their D1, followed by a listening/imitation task with Glasgow English (participants repeated sentences after a Glasgow speaker on a recording). The first training session consisted of two identical iterations of this procedure, and lasted 20 minutes or less (German, Carlson and Pierrehumbert, 2013). A generalisation task concluded the first week of the experiment. Participants were required to continue to imitate the accent they had heard while reading words aloud that were not in the training material (German, Carlson and Pierrehumbert, 2013).

After a short break of a week, the participants returned to complete the experiment. The second and third tasks above were repeated, with the addition of another generalisation task. Given the hypotheses about linguistic plasticity in adulthood, it would be a reasonable assumption that the rapid exposure to D2 sounds would not be sufficient to generate significant, persistent, or productive changes in the participants' production.

The results contradicted that expected outcome. In terms of the /t/ variant, almost all of the participants used the Glaswegian [t^h] properly in word-initial position, and in the other conditions used it correctly much of the time (German, Carlson and Pierrehumbert, 2013). The unfamiliar dialect's [ɾ] variant likewise showed statistically significant differences after the training, although it seemed the /r/ allophone proved more difficult to learn. The analysis revealed that the VOT, F3 minima, and closures for the consonants converged with Glaswegian values (German, Carlson and Pierrehumbert, 2013).

The differences were striking when compared with the baselines; participants jumped from nearly zero to almost 100% production of the [t^h] allophone from the target D2 between the baseline recording and the first training for both word-initial and medial /t/ (German, Carlson and Pierrehumbert, 2013: Fig. 4). This sudden change was consistent throughout the remainder of the study. Participants also changed their /r/ allophone in the very first training, ditching their native production for a 60% increase in use of [ɾ] in word-initial position, and a 40% increase in medial position (German, Carlson and Pierrehumbert, 2013: Fig. 5).

Again, these changes persisted throughout the experiment. Both allophones were remembered with nearly the same accuracy as in the first week in the later generalisation tasks, even after time had passed. In summary, German, Carlson and Pierrehumbert concluded that the participants in this study had used language acquisition mechanisms to rapidly reorganise their phonetic and phonological systems. A very rapid change (within one training) was both productive and persistent, because they were able to generalise the knowledge to new words without additional D2 input.

That study is not alone in demonstrating that late-stage learners can change their production and perception in a D2 or L2. Evidence from research on second dialect acquisition suggests that dialect and accent variables within a native language can be affected by social pressures in speakers who move from one region of England to another (e.g., Evans and Iverson, 2007). Additionally, Flege and Eefting's work from the 1980s asserted that late-stage learners of English adjusted their categorical boundaries for consonants in both English and their native Dutch, forming an intermediate VOT-based category (Flege and Eefting, 1987). Using synthetic speech, they found that participants were consciously aware of the acoustic differences between Dutch and English /t/, and that the most proficient among them moved their Dutch VOT toward a more ‘English’ pattern by shortening it significantly (Flege and Eefting, 1987).

All of this evidence suggests that the production and perception of consonants are changeable aspects of language for adult learners both in L1 and late-learned L2/D2. But if adult speakers are not by definition ‘crystallised’ into their native languages, what could be happening with the brain and the systems involved in language? Late learners may be able to alter their subtle brain structure or their language processing through learning a new language, preliminary work suggests (Mayberry et al., 2011; Morgan-Short et al., 2012).

A particularly interesting study on adult brain plasticity used an artificial language, taught in an explicit classroom setting and an implicit, ‘immersion’ environment to different groups of participants. The participants were all over the age of the posited critical period, and once they had achieved reasonable proficiency in the artificial language underwent an Event-Related Potential (ERP) brain imaging session (Morgan-Short et al., 2012). ERPs reflect real-time scalp-recorded electrophysiological brain activity, and in this case showed that those with the implicit training had patterns of activity that mimicked those of the participants in their native languages (Morgan-Short et al., 2012). This evidence would suggest that linguistic plasticity may continue well beyond the critical period.

If brain plasticity persists into adulthood, it is likely that the participants in this experiment would have been able to acquire new phonemes. A more interesting question may be whether they should have been able to acquire fine-grained detail of an L2 such as VOT, especially in such a short training period and with so little exposure to the target language. The trainings for this experiment could be described as more classroom-like and not very immersive, and therefore it would be surprising if fine-grained changes were observed. A further question would be *how* the participants do this, and what underlying mechanisms play a role. Several models for language acquisition compete in the literature. The most salient for this study are Motor theory, the Speech Learning Model, and the Native Language Magnet model.

Motor theory was initially proposed in the 1950s by a team working on technology for the blind, and sought to link production and perception inextricably (Liberman, 1996). The essential claim of this model for speech perception was that people perceive language in terms of the articulatory gestures that produce sound, not through interpreting the sounds themselves (Liberman and Mattingly, 1985). After transitioning from associationist attitudes to a cognitivist paradigm once it was shown that infants can perceive phonetic contrasts before they are able to speak, the model asserted that speech is ‘a special code’ (Liberman, 1996). Encoding occurred below the level of the neuromotor commands to the articulatory muscles (Liberman et al., 1967: 431). A ‘decoder’ would be necessary to match the articulatory gestures to the corresponding phonemes.

Although some of the initial claims of Motor Theory (MT) have since been abandoned, the central claims remain. According to the revised version from recent years, the first of these remains that the objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands that instruct the phonetic articulators (Liberman and Mattingly, 1985; Liberman and Whalen, 2000).

The second claim is that the perception and production of speech share the same gestures, in an innate specification that requires epigenetic activation and development during the critical period (Liberman and Mattingly, 1985; Galantucci, Fowler and Turvey, 2006). The work of many others has professed to support these central claims including work on ‘mirror neurons.’ Advocates of MT argue that mirror neuron activity provides direct empirical evidence for motor system involvement in speech perception, because the areas full of activity during speech perception appear to correspond directly to those involved in speech production (Rizzolatti, Fogassi, & Gallese, 2001; Rizzolatti & Craighero, 2004).

In contrast to MT's focus on the mechanical aspects of speech perception, the Speech Learning Model (SLM) is firmly rooted in L2 acquisition research. In observing immigrant children, many researchers in the 1960s, 1970s and 1980s framed the process in terms of a complex relationship between L1 and L2 (the Contrastive Analysis hypothesis, see especially Upshur, 1962; Eckman, 1977). According to the CAH, the more closely the phonemes of a language resembled those of the native one, the 'easier' it should be to learn for a speaker (Flege, 1987; Ellis, 1994). The hypothesis claimed that interference from L1 can significantly hamper L2 learning efforts.

Work by Flege and colleagues however, has challenged this paradigm. In repeated studies with immigrant communities, it emerged that late learners were able to approximate the acoustic values of their target language along something of a spectrum of accuracy (Flege and Port, 1981; Piske, MacKay and Flege, 2001; Flege et al., 2006). Those who began learning at a later age produced significantly different phonemes from L1 monolinguals. Crucially, these phonemes were also different from those of L2 native speakers (Flege, MacKay and Meador, 1999; Flege, Schirru, and MacKay, 2003). Some of the participants in these studies achieved native-like production and perception despite beginning late.

Building on the results of this and many other such studies on late learners, Flege and others developed the SLM. Essentially, they advocate a paradigm under which adult L2 learners can perceive and produce the phonetic properties of an L2 (Flege, Schirru, and MacKay, 2003). The central claim of the model is that the processes and mechanisms of successful L1 acquisition remain intact and accessible throughout life (Flege, 1999). Essentially, language learning relates to organising the same phonetic space in different ways to convey meaning through contrasts (Flege, 1999). SLM therefore proposes several hypotheses about language learning:

i. The greater the perceived dissimilarity of an L2 sound from the closest L1 sound, the more likely there will be a new category formed.

ii. Category formation in an L2 becomes less likely as childhood progresses, while representations for ‘nearby’ L1 sounds develop.

iii. When a category is not formed for an L2 sound, this is because it is too similar to an L1 counterpart (‘merging’).

iv. When a new category is established for an L2 sound, it may dissimilate from ‘nearby L1 *and* L2 sounds, to preserve phonetic contrasts.

Flege, 2005

Based on these principles, the SLM is very distinct from MT, although the targets of their proponents may be distinct. MT tends to focus on L1 acquisition and speech perception, and from the very beginning the SLM was based on L2 research. There have been few attempts to combine the two into a gestalt, given their distinct foci.

The newer Native Language Magnet model is another possibility for speech perception and L2 learning. This model of perception and production is much more recent than the SLM or MT. Born out of a sense that the computational models of infant language acquisition were missing factors that may influence success in a native language, and how these condition the brain, the Native Language Magnet (NLM) model was first proposed in 1992 (Kuhl, 1992).

In early studies, proponents of this model saw a ‘perceptual magnet’ effect when adults and infants were exposed to variants of speech sounds (Kuhl, 1991). Human adults and children were drawn to sounds typical of their L1. Specifically, infants' speech perception in research conditions demonstrated a dual transition near the end of the first year of life (Kuhl et al., 2008). Non-native speech perception declines and native language speech perception skills simultaneously show improvement. After these changes, it appears more difficult to change production or perception, which proponents of the NLM suggest is evidence for neural commitments made in the brain that favour the native language over all others (the native language neural commitment, e.g. Kuhl, 2000).

For adults, these neural commitments would make them more attuned to the acoustic properties of their native language and less able to produce to perceive non-native values. Several studies have appeared to support this view; L1 interference seems to affect adult learners in L2 under some circumstances (Flege 1995; McCandliss et al. 2002; Iverson et al. 2003). Of particular interest are studies that provide evidence for adult Japanese speakers who have learned English. The /r–l/ contrast is intensely difficult for Japanese speakers in English.

Work with this contrast using speakers of English, Japanese and German appears to suggest that speakers attend to different dimensions of the same stimulus depending on the patterns of their native language (Iverson et al. 2003). Kuhl and others have therefore argued that these results suggest that human organisation of phonetic categories is based around prototypic members from their native language. Furthermore, this organisation is early and species-specific (Kuhl, 2004). Kuhl et al. (2008) proposed five general principles that were drawn from the original NLM.

MT, the SLM, and the NLM all provide important insights into the mechanisms underlying human speech perception, but they are not the only considerations that need to be addressed for this experiment. Overarching claims about human learning could influence the results. Learning in humans is likely to be domain-specific, meaning that there are many independent and highly-specialised knowledge groupings that we can access at any given time (Brown, 1990; Hirschfeld and Gelman, 1994). However, these domains may not influence one another. In terms of this experiment, domain specific learning would predict that training for production will not influence perception of a novel contrast. The claim would be that production and perception fall into independent domains, a claim supported by recent L2-learning studies on adult Japanese speakers who improved pronunciation, but not perception, of /r/ and /l/ in English after phonetic training (Hattori, 2010).

Questions of domain-specificity aside, language learning is part of a gestalt of social and psychological factors in each individual. Social attitudes appear to influence ability to learn an L2. Anecdotally, it appears that having a certain openness to new experiences seems to go along with increased ability to learn a language. In order to investigate these claims, this experiment used a Likert-type scale of tolerance-intolerance of ambiguity, developed for use in personality tests (Budner, 1962).

This questionnaire requires that participants indicate how strongly they agree or disagree with statements about ambiguous material (new experiences, social interaction, and personal preferences). Intolerance of ambiguity was defined as 'the tendency to perceive (i.e. interpret) ambiguous situations as sources of threat,' while tolerance of ambiguity was seen as 'the tendency to perceive ambiguous situations as desirable.' (Budner, 1962). Additional research described those with a low tolerance for ambiguity may be dogmatic, authoritarian, and inflexible in their attitudes (Bochner, 1965). Low tolerance for ambiguity is also correlated with mental health problems such as depression and anxiety (Andersen and Schwartz, 1992; Leyro, Zvolensky and Bernstein, 2010).

Several studies have linked tolerance of ambiguity with L2 learning, showing that a higher TOA score correlates with higher proficiency in the target language (Chapelle and Roberts, 1986; Furnham and Ribchester, 1995; Lee, 1998). Humans are not one-dimensional, and even in the context of a laboratory experiment participants will necessarily be influenced by psychological and social factors such as ambiguity tolerance. In fact, the novel experimental experience may be better suited to those with higher ambiguity tolerance in the first place. If low tolerance of ambiguity is correlated with resistance to change in a general sense, the specific changes needed to learn a new consonant contrast may be influenced.

The degree to which this influences adult learners can be tied directly to the models of speech perception. SLM and NLM-e both explicitly refer to social factors as major influences on the learning of languages. In the case of the SLM studies, immersion in a novel linguistic and cultural environment through immigration may reveal that intense social pressure can drive linguistic changes in adults (Flege, 2005). The NLM-e directly links psychological and social factors to language learning, although it focuses on the social environment of infant language acquisition (Kuhl et al., 2008). MT does not make much comment on the possibility of influence from social cues, possibly due to the modular paradigm in which the theory operates.

This experiment was designed to investigate the extent to which naive speakers can learn a novel VOT boundary for consonants in an L2 and how the non-linguistic factor of tolerance of ambiguity may play a role in L2 learning. Building on the research from the speech perception models above, English speakers were trained in French /p/ and /b/, which differ significantly in VOT duration from English VOT values. They heard French words that began with /p/ or /b/ and sentences that contained /p/ and /b/ initial words, and were instructed to repeat them as accurately as possible. English /p/ is generally aspirated, and has a long VOT (In French, the /p/ VOT is considerably shorter and often leads to confusion in native English speakers, who mistake it for a /b/. French /b/ may be prevoiced, or have an even shorter VOT than its English counterpart.

The aim was to investigate whether or not English speakers, with no little or previous experience of French, were able to adjust their production of French /p. and /b/ to better match that of French speakers and whether or not any changes in production would improve their perception of these consonants. In addition, the experiment investigated the role of psychological and social factors that may influence L2 learning in the form of a Tolerance of Ambiguity questionnaire. The experiment focused on four main research questions:

1. What happens to production of L2 consonants for English speakers with concentrated and rapid production training in French?
2. What happens to the perception of these consonants in French?
3. Does a change in production in French result in a change in production in English?
4. Do participants with a higher tolerance of ambiguity achieve better results with learning a new consonants?

This experiment explored novel territory. Firstly, it used strictly production training in the attempt to influence both production and perception. This investigates the link between the two and how they may interact. Secondly, the experiment focused on rapid phonetic training in an L2. The phenomena described here have been investigated in D2 acquisition (e.g. German, Carlson and Pierrehumbert, 2013), but work with short-term, rapid second language phonetic training is sparser. Finally, the investigation of the relationship between tolerance of ambiguity and improvement in L2 rapid acquisition has the potential to reveal more about the non-linguistic factors that influence successful L2 learning.

The models for speech perception outlined above can be used to make specific and competing predictions about the results of this experiment. If the MT model is correct, then the production training in French for this experiment should provide the necessary stimulus for both production and perception. Changes in the motor patterns for production would necessarily instruct the perception ‘decoder’ in the novel contrast. The results would show that perception and production are intimately linked, and that significant changes in one of them should lead directly to significant changes in the other.

In addition, the fact that participants in this study were asked to imitate the French speakers as closely as possible may influence the outcome.Humans tend to be excellent imitators. This may function as a means to ease social interaction among humans (Adank, Hagoort and Bekkering, 2010). This process of imitation is often referred to as convergence or accommodation to the speaker, which has been shown to occur in conversation and also in exposure to speech in laboratory settings (Goldinger, 1998; Pardo, 2006; Delvaux and Soquet, 2007). According to some accent training research, imitation may increase ability to comprehend an unfamiliar accent (Adank, Hagoort and Bekkering, 2010).

Under MT, the imitation training could lead to new motor commands, and new speech perception patterns. This would imply that both production and perception would improve if the participants changed their speech significantly to imitate the French speakers. If only production were to change and the training had no effect on the perception of the participants, the results would suggest that the intimate link proposed by MT is unlikely. Furthermore, a high or low tolerance of ambiguity should not correlate with any changes due to the separate modular nature of linguistic and non-linguistic processing.

Non-linguistic factors come into play with the remaining speech perception models. Under the Speech Learning Model, the results of this experiment should be considerably different from those predicted by Motor Theory. Given sufficient time and training, the participants should be able to adjust their production and perception of /p/ and /b/ to reflect the differences between English and French versions. They should have been able to reorganise their phonetic space such that they can create a new category, as long as the perceived dissimilarity between the English and French sounds is large enough (Flege, 2005). It could be more likely that they would make a new /p/ category, as that is perhaps the more dissimilar consonant from the English pattern. If they did create a new category for perception or production, then it could be intermediate between the French or the English phonemes in terms of VOT.

Finally, the NLM-e model would make additional predictions. If correct (as revised, NLM-e, Kuhl et al., 2008), then the neural commitments that the participants made as infants to their native language should prevent them from being able to fully form a new category for /p/ and /b/ in French. The individual TOA score should correlate with the language learning abilities in French. Specifically, the high-tolerance participants should also be those who change perception/production the most.

Methodology

Participants

A total of 10 participants took part in the experiment, six female and four male. They ranged in age from 19 to 59 years. Nine indicated that they had no professionally-diagnosed disability that might affect the results of the study such as hearing problems, learning difficulties, or problems with reading and writing. One indicated that she experienced mild reading and spelling difficulties.

All participants were native, monolingual speakers of British English and their parents were also native speakers of British English. All were living in Greater London at the time of testing and had been born and raised in England. All participants had little experience of or fluency in the French language and had not spent any appreciable length of time in a Francophone country. Although seven had studied French at school, none indicated any level of fluency and none had ever used French outside a classroom setting.

The participants were asked to fill in a questionnaire about the circumstances of their native language acquisition, their educational attainment, the areas of the world in which they have lived, and the levels they had achieved in learning additional languages. The highest level of education for a participant was a completed master's degree, and the lowest was completion of two years of secondary school.

Stimuli

The stimuli for this experiment were recorded by four native French speakers. All of them were female and grew up in the North of France. All recordings were carried out in sound-attenuated recording rooms at Chandler House, UCL, using a high quality rode nt1-a microphone connected to a PC via a fireface UC sound card. Recordings were made with a sampling rate of 44.5 kHz, 16-bit resolution.

Three sets of stimuli were recorded. First, words with the target consonants in focused, word-initial position in isolation and in simple sentences. Second, a set of more complex sentences. These stimuli were used for training. The final set of stimuli were minimal pairs containing the target contrast. These stimuli were used to test whether or not training had led to improvements in perception of the French /p/-/b/

contrast. The formulation of these stimuli is consistent with those in the German, Carlson and Pierrehumbert accent-training experiment, which also used target sounds in a sentence-final position (German, Carlson and Pierrehumbert, 2013).

The first set of experimental stimuli were recorded by two of the talkers. They read aloud a list of 24 words with word-initial focused target consonants (/p/ and /b/, see Table 1.1, Appendix 1). These words were chosen for their simplicity and for the possibility of finding an image to accompany the stimulus. As such, they are generally nominal. These words were recorded in isolation and in simple carrier sentences ('Elle/il a vu un ________ .').

Additional stimuli recorded with the first two volunteers came from a set of translated standardised sentences. They mimic the Bamford-Kowal-Bench (BKB) sentences that are commonly used in phonetic and phonological research. These sentences were translated into French by a native speaker and designed for use in English to test speech perception in children with hearing loss (Bench, Kowal and Bamford, 1979; see Table 1.2, Appendix 1). The sentences were manually sorted for sentence-final, /p/ and /b/ initial words. These stimuli are also imageable.

The final set of stimuli were recorded by two additional volunteers, and consist of French /p/-/b/ minimal pairs. The words were drawn from dictionaries and checked by native French speakers for accuracy. Although some of the words are rare or old-fashioned, none are nonsense words or false (see Table 1.3, Appendix 1). Additionally, these words did not occur any other part of the experiment. After recording, words and sentences were saved into individual wav files in Praat (Boersma & Weenink, 2014).

Training stimuli (isolated words and sentences) were presented in quiet. Pilot testing with a native monolingual English speaker with little experience of and no formal learning of French, showed that the French /p/-/b/ words were easily identifiable in quiet. Consequently, these words were presented at various five different noise levels; Quiet, +3 dB, 0 dB , -3 dB and -6 dB SPL. The noise conditions were created by adding speech-shaped noise (S.Rosen, UCL). Speech-shaped noise rather than multi-talker babble was used as previous research has shown that the accent of the babble can affect speech processing (Van Heukelem & Bradlow, 2005). To create the speech-in-noise conditions, the RMS amplitude of the sentence and noise were determined and scaled to fit the SNR condition. They were then combined through addition at the four SNRs; +3 dB, 0 dB , -3 dB and -6 dB using an automated script in Praat. Lastly, all files were equalized for intensity at 70dB SPL.

Procedure

The experiment lasted between 35 and 45 minutes depending on the individual, with three experimental phases. The pre-test phase, training phase, and post-test phase were almost entirely equal in length and followed one another in rapid succession. All testing took place in a sound attenuated recording room at Chandler House, UCL. All test stimuli were presented via a PC over high quality headphones and responses logged via an experimental script running in Praat (Boersma & Weenink, 2004). Recordings were made using a high quality rode nt1-a microphone connected to a PC via a fireface UC sound card and were made with a sampling frequency of 44.5 kHz, 16-bit resolution. Participants were offered breaks at regular intervals, but only one took the opportunity.

The pre-test established a baseline VOT in English and French for each participant as a control. The first task required that they read out a list of fifteen words in English (see Table 1.4). The list had ten words with the target consonants /p/ and /b/ in word-initial position, as well as five distracting words without the targets. This established their baseline VOT for English /p/ and /b/ consonants, without exposure to any French stimuli.

Following the English recording, they recorded a list of fifteen French words (Table 1.5, Appendix 1) to establish a French baseline for VOT. As in the English VOT baseline recording, ten of the words had the target consonants and five were distractors. In both tests, words were recorded in a randomised order to eliminate any list effects. The final part of the pre-test phase was the minimal pair identification task. Participants listened to and identified French /p/-/b/ minimal pairs. In each trial, they heard a word and then identified whether the word began with a /p/ or /b/, by clicking on the associated letter on the PC. There were 262 trials, presented in a randomised order.

Participants then proceeded to the experimental phase. There were two training sessions. The first consisted of 88 trials. In each trial, participants heard and saw the word in isolation first and then in a short carrier sentence accompanied by an image that illustrated the meaning of the word (e.g. 'Baleine….Il a vu une baleine.'). The participants were recorded repeating the word in isolation and embedded in the sentences, and asked to mimic the accent of the speaker as closely as possible. The training was self-guided and participants advanced through it by clicking on the images at their own pace. The training took an average of six minutes for participants to complete.

After being offered a short break, the participants completed the second training session. For each trial, participants heard a sentence and then repeated the sentence as accurately as they could. The training consisted of 44 trials and took an average of four minutes to complete. The entire training phase of the experiment took between eight and eleven minutes for each participant. The order of presentation of stimuli in both parts of the training session was pseudo-randomised; stimuli were randomised and all subjects completed the experimental trials in this order.

The final phase of the experiment was the post-test. This phase was nearly identical to the pre-test, but with the addition of a short text to read aloud. This text was two paragraphs, drawn from Cendrillon (Cinderella), the Smashwords publishing company translation of the Brothers Grimm tales (KidLit-O, 2013; see Appendix 1 for full text). The intention of this recording was to give us a glimpse of connected speech with the target consonants in context, as well as to monitor how persistent any changes might be in the context of a more difficult reading task.

Following the short reading, participants completed the minimal pair identification task once more. Afterwards, they recorded two more lists of French and English words to monitor any potential changes to VOT production. Words were recorded in a randomised order to eliminate any list effects (see Table 1.6 and 1.7, Appendix 1). Finally, the participants completed the Tolerance of Ambiguity questionnaire to add an additional factor to the analysis of any changes. This questionnaire is standard for investigating the link between attitude and linguistic performance ((Chapelle and Roberts, 1986; Furnham and Ribchester, 1995; Lee, 1998).

This portion of the experiment took approximately ten minutes.

Analysis

Data from the minimal pairs identification task was extracted from Praat and saved. Recordings were converted to single-channel mono WAV files and then VOT measurements made using Praat’s annotation functions. This was accomplished by manual insertion of boundaries at the zero crossings nearest the release burst and the beginning of glottal pulses (onset of voicing). The measurements were consistent with the standard definitions of VOT (Lisker and Abramson, 1964; Abramson, 1975; Serniclaes et al., 2001). Duration was then measured using an automated script.

Each of the participants should have produced 172 recordings of words with /p/ or /b/ in word-initial, focused position. However, some of the participants appeared to struggle with the training tasks. They missed words or altered them so as to be unintelligible. One participant skipped eight of the sentences in the second training (18% of all trials in the training). These errors resulted in 1708 viable recordings in total, when the expected number should have been 1720.

Results

To investigate whether or not Tolerance of Ambiguity played a role in participant's ability to produce or perceive the French VOT contrast, all following analyses included the TOA score as a between-subjects factor. The TOA scores were split into two groups, High Tolerance (HT) and Low Tolerance (LT), based on the median score of all participants (3.73 out of 5, see Table 3.1, Appendix 2).

English VOT production

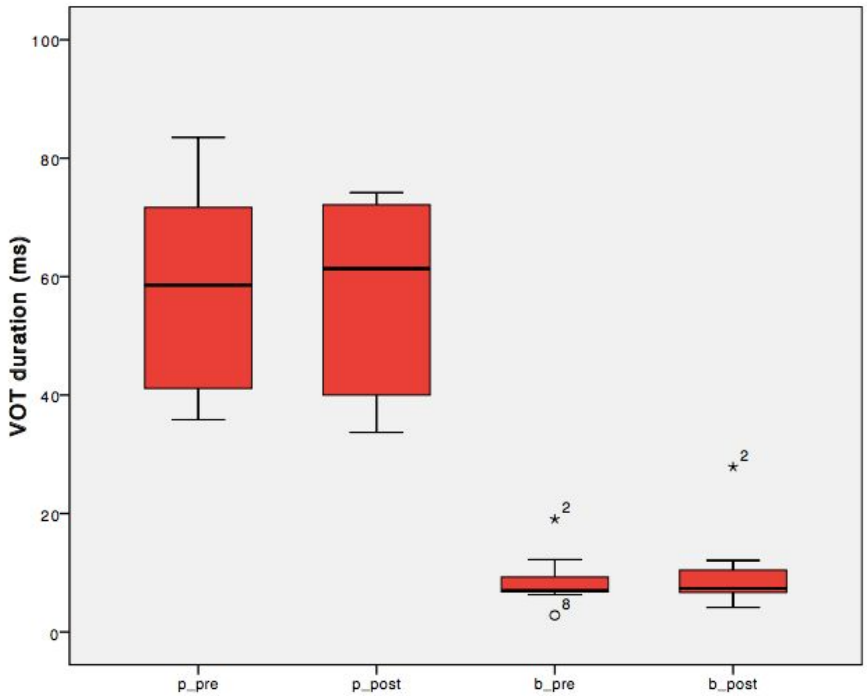

Figure 2.1- VOT measurements subjected to repeated measures ANOVA for English baseline and post-test recordings. Clear categorical difference between /p/ and /b/ productions for both.

As displayed in Fig. 2.1, there was no significant change between the pre- and post-test in terms of VOT for the English words, indicating that training with the French VOT contrast did not affect production of their native category. As expected, there is a clear categorical difference between the /p/ and /b/ consonants is striking, with distinct, non-overlapping distributions for /p/ and /b/, in both the pre- and post-test.

French VOT production

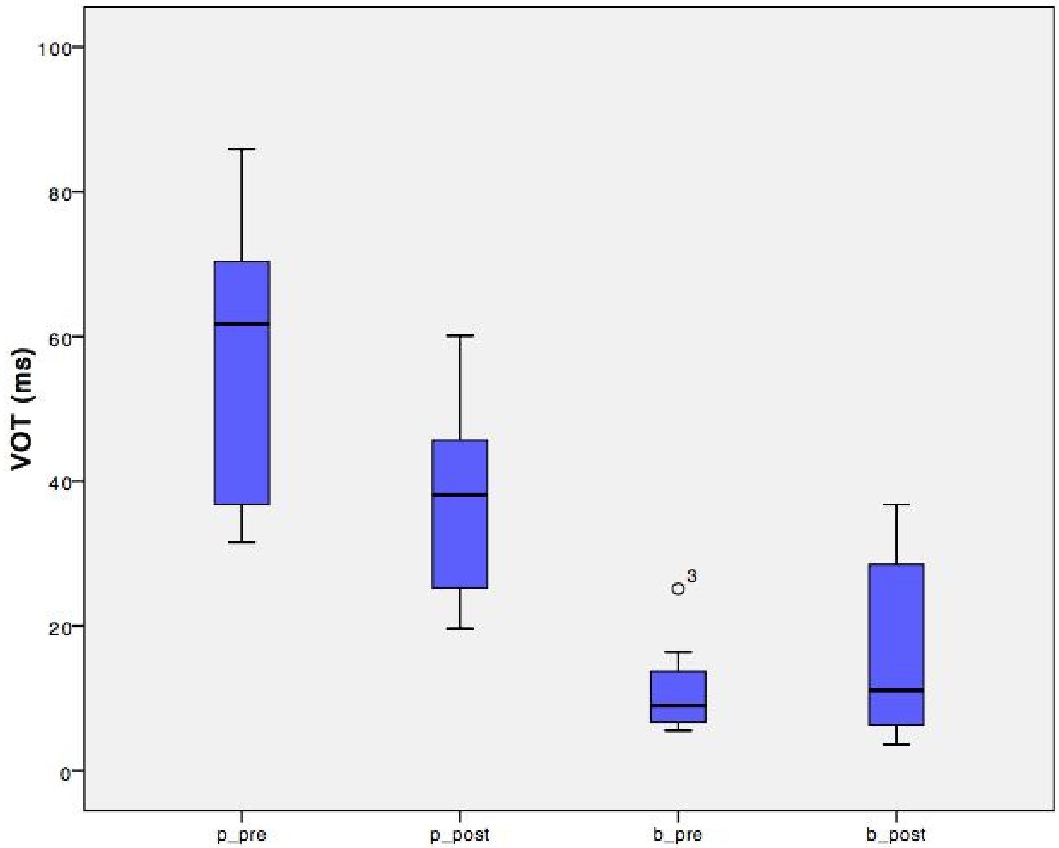

Figure 2.2- VOT measurements with repeated-measures ANOVA in SPSS. This box plot shows changes in VOT from the pre test to the post test. Note the significant change from the pre-test to the post-test for /p/ and the slight overlap between VOT duration for /p/ and /b/ in the post-test.

As displayed in Fig. 2.2, there appear to be some changes in the production of the French stops as a result of training. It appears that the VOT duration for the /p/ phoneme is reduced significantly after the training phase. In addition, the categorical difference between the consonants is less clear. The post-test /b/ values appear to overlap somewhat with the /p/ values in terms of VOT, which is interesting given that English speakers often confuse /p/ for /b/ in French.

These potential effects of training were tested in a repeated measures ANOVA with phoneme (/p/,/b/) and time (pre- post-test) coded as a within-subject variables, and ToA score as a between-subjects variable. There was a significant effect of phoneme

$F(1,8)=30.65$, $p<0.001$, confirming that participants produced /p/ and /b/ with different VOT values. There was also a significant effect of time, $F(1,8)=23.03$, $p<0.001$ and an interaction of time and phoneme, $F(1,8) = 23.03$, $p = 0.001$, confirming that participants changed in their production of /p/; after training they produced /p/ with shorter VOT values. There were no other significant effects or interactions, $p > 0.05$.

Any potential change in production of /b/ was also investigated by looking at the number of pre-voiced tokens in the pre- vs. the post-test. There was a slight increase in the number of pre-voiced tokens from the pre-test to the post-test, but a paired-samples t-test showed that this was not significant, $p > 0.05$ (Table 2.3).

Changes to VOT During Training

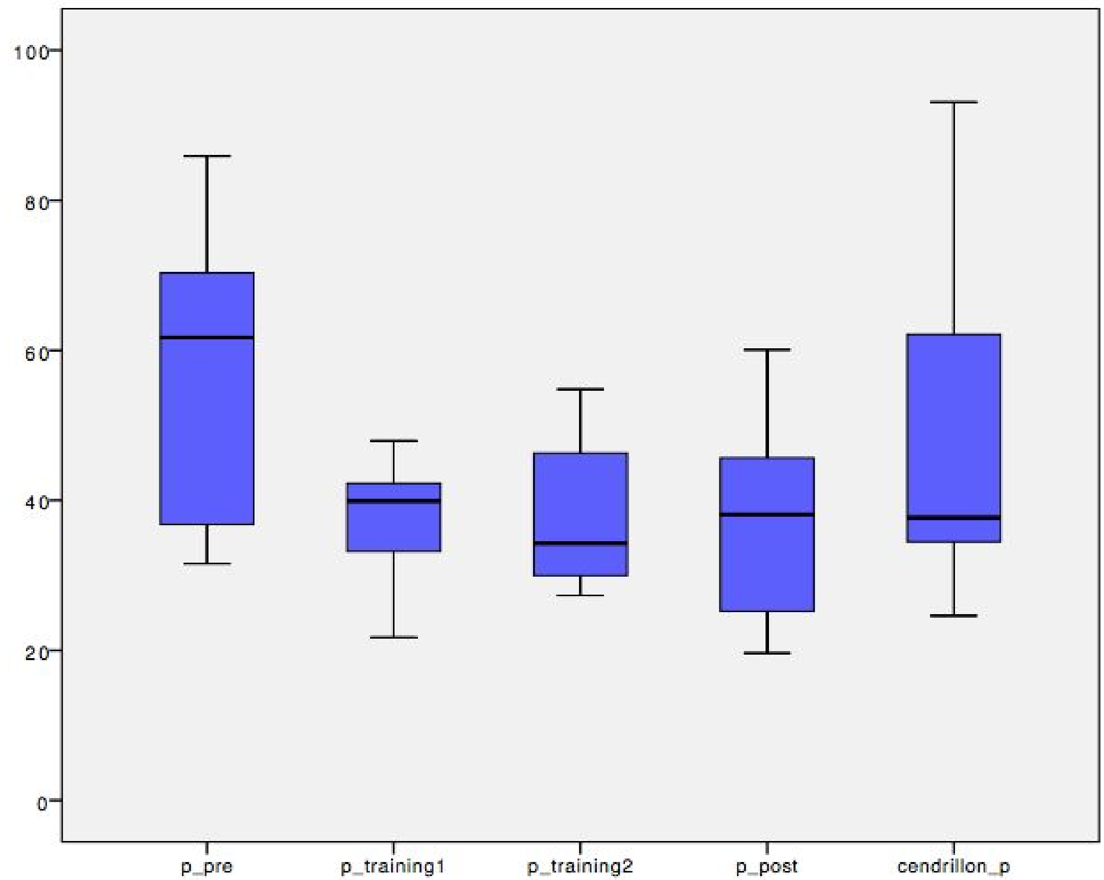

Figure 2.4- VOT for /p/ phoneme through five portions of the experiment. Average values for VOT changed rapidly and remained lower than the baseline (p_pre) for the duration of the two trainings, the post-test, and the connected speech sample (cendrillon_p).

The results presented in Fig. 2.4 demonstrate that the change in the production of /p/ was almost immediate in the first training session. This change persisted throughout the rest of the experiment, with all of the following /p/ productions falling into statistically different categories from the pre-test. There did not appear to be much change from the first training onward.

This was tested in a repeated measures ANOVA with time coded as within-subject factors and ToA as a between-subjects factor. The analysis revealed that for phoneme the values were $F(1,8)=97.94$, and $p<0.001$, confirming that there was no significant change between the two trainings. There were no other significant interactions or effects, $p > 0.05$.

This confirms that the change in production of /p/ was very rapid and although the VOT varied slightly for the different parts of the experiment, the change persisted even into the connected speech sample in the Cendrillon text. The range is larger in this case, and more similar to the baseline English VOT productions, but the average remains near those of the other tasks.

Speed of /p/ Production Change

A further investigation of the production data for /p/ aimed to determine how quickly the VOT adjusted to the French pattern. Previous studies (German, Carlson and Pierrehumbert, 2013) indicate that this change should have occurred very rapidly, within the first few trials of Training 1. As such, the VOT from the French pre-test and the first five /p/ trials of the training were compared. On average, the participants had changed the duration of their VOT for /p/ by more than 16 milliseconds from the baseline by the time 40 seconds of the training had passed (See Table 2.5).

	Averages Pre	Averages Training (First five trials)	Difference
s01	68.3966	40.147	28.2496
s02	60.6825	39.4194	21.2631
s03	70.3254	37.8188	32.5066
s04	72.4498	50.3248	22.125
s05	47.2138	42.582	4.6318
s06	32.2162	33.4374	+1.2212
s07	36.7738	17.8572	18.9166
s08	31.556	64.0494	+32.4934
s09	85.913	31.626	54.287
s10	62.7626	47.6942	15.0684
Average	56.82897	40.49562	
Difference	-16.33335		

Table 2.5- Differences in the VOT for /p/ productions, comparing the five trials in the pre-test to the first five trials of training 1. Average VOT for /p/ tokens was reduced by 16.3 milliseconds on average within the first five trials.

The largest change from pre-test to the beginning of the trainings was an average decrease of 54 miliseconds for s09. Strangely, some of the participants appeared to increase the duration of their VOT from the pre-test to the first part of the trainings. Despite this, the previous data demonstrates that the average /p/ VOT was significantly lower in the two trainings, the connected speech sample, and the post-test overall.

Perception: Minimal Pairs Identification Task

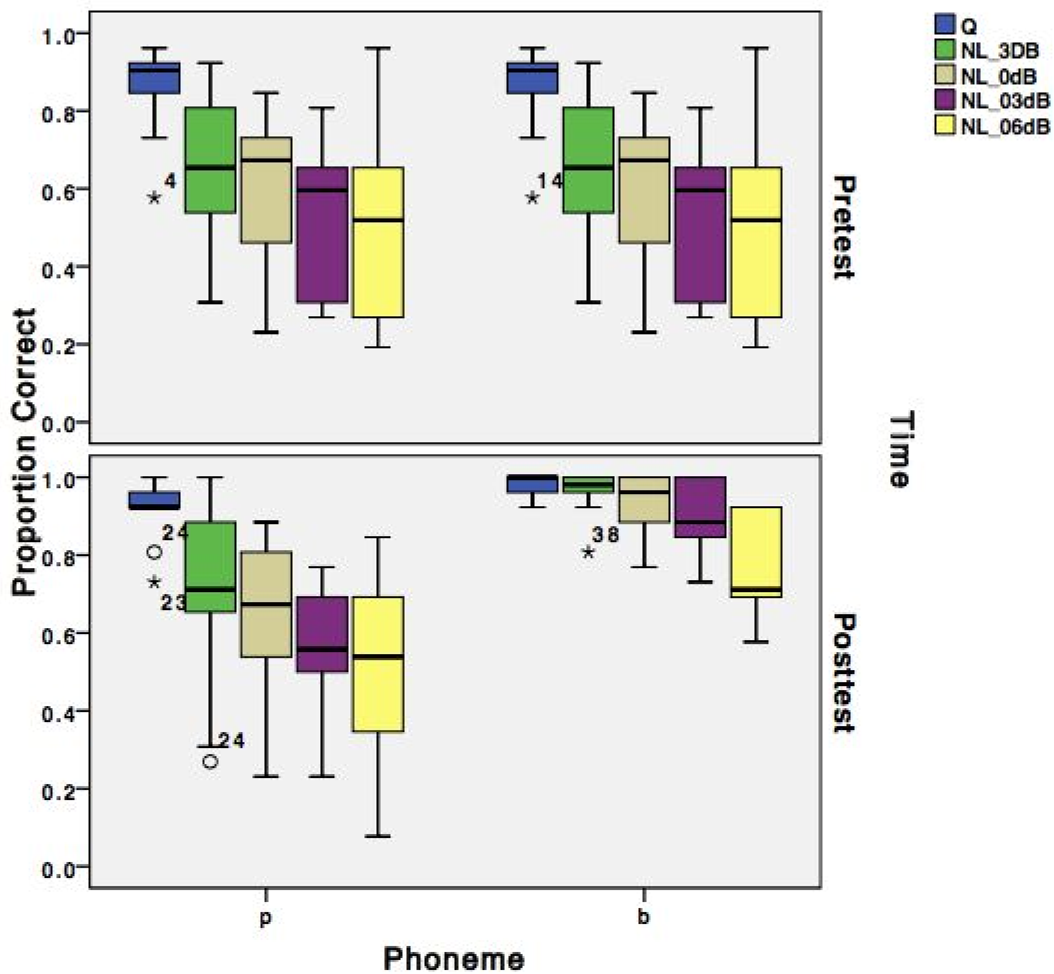

Figure 2.6- Statistical analysis for minimal pairs identification tasks in the pre and post tests. Overall marginal improvement driven by the improvement in the /b/ identification, with no marked improvement for both after the training.

The results for the perception portion of the experiment are similarly clear-cut (please see Table 2.6). The data from the minimal pairs tasks was subjected to an ANOVA, with noise-level coded as a within-subjects variable and TOA score as a between-subjects variable.

There was a significant effect of phoneme ($F(1,8)=16.49$, $p<0.01$), a marginally significant effect of time ($F(1,8)=4.96$, $p=0.057$). Furthermore, there was a significant effect of noise level ($F(4,32)=77.29$, $p<0,001$), and a significant interaction between noise and phoneme ($F(4,32)=5.28$, $p<0.01$).

The participants did worse with /p/ perception than with that of /b/. The participants became marginally better at knowing when /b/ appeared in the minimal pairs, which appears to have driven an overall improvement (albeit a marginal one). When it came to the noisiest of the samples, the participants clearly guessed (50% correct, 50% wrong). There was no effect of TOA on these results, $p > 0.05$.

Discussion

These results demonstrate that participants were able to change their production of the French /p/-/b/ contrast. Specifically, participants changed their production of French /p/ such that it was produced with a short VOT. There was no evidence for any change in the production of /b/. The change in /p/ was rapid and persistent, and participants were able to generalise this learning to new stimuli.

However, despite what some models of speech perception might predict, the results do not indicate that the participants' perception was significantly altered by the same training that generated the production change. Does this change in /p/ production in French result in a change in English VOT for /p/? Confidently, it does not appear to have an effect. Tolerance of ambiguity does not appear to be as closely correlated with L2 success in this study as has been asserted in the past.

Clearly, the French consonant training had effects on the production of the participants. These effects were almost immediate. The difference on average between the pre-test VOT duration for /p/ and that of the first training (Figure 2.4, Table 2.5) demonstrates that mere seconds of instruction persuaded participants to change their production. All of the participants changed their production significantly by the end of the first training, and this change was consistent across the three tasks that followed (see Table 2.6). This effect was not seen in the /b/ VOT values, which is unusual.

Given the somewhat unconscious nature of the difference between French and English /p/ (compared with the perhaps more obvious prevoicing for French /b/) the fact that the /p/ VOT changed significantly for all participants is remarkable. It remains to be seen why the participants changed only their /p/ values. Speculation could include the fact that French /b/ values, when not prevoiced, are too similar to English ones for participants to notice a difference, or that they merged the new category with a neighbouring similar one (Flege, 2005). It has been shown that prevoicing is used in English (e.g., Docherty, 1994) but that this is used differently by different speakers. There is some evidence that this is the case here; the participant with the biggest difference in pre-voicing in the post test (s02) was also the one with the most /b/ pre-voicing in the English baseline recording.

The /p/ VOT duration in the final parts of the experiment was perhaps most interesting of all. The participants remarked anecdotally that the second training session and the connected speech sample (Cendrillon) were particularly difficult. Indeed, one of the participants apparently found the second training so difficult that she skipped nearly 20% of the sentences (s05). Given their remarks and the difficulty of the tasks, it would have been reasonable to assume that the VOT duration would creep back up toward the baseline as the cognitive load involved increased.

However, this was not the case. The average duration of /p/ VOT for the anecdotally 'more difficult' tasks is still within the range of the first training, and significantly different from that of the pre-test. This indicates that the changes established in seconds in the first training continued into the final tasks. Likewise, the post-test average was not significantly different from those in the previous three tasks (see Figure 2.4). This is strong evidence that the participants had learned a new way of producing /p/ that, despite no additional French input remained usable during the connected speech sample.

Even more importantly, the connected speech sample included a number of novel words that participants had not encountered in the training (‘parasait,’ ‘pauvre,’ etc.). The fact that they were able to produce them with similar accuracy to those in the training suggests they were able to generalise their newly-acquired knowledge to novel words. This may be evidence for changes in the underlying representations necessary for speaking French, a finding that replicates results from the German, Carlson and Pierrehumbert study (2013).

Furthermore, the VOT duration changes observed for the /p/ productions in French had no significant effect whatsoever on the English VOT in the post-test. The post-test English VOT results indicate that the participants did not change their English category for /p/. Beyond that, the French production is intermediate in duration between a native French VOT and that of an English /p/. This indicates that L2 learners may create intermediate categories that resemble neither language, perhaps in an attempt to differentiate between their L1 and L2 and preserve phonetic contrasts. Corpus studies with bilingual adults demonstrate this; it appears that part of L2 acquisition may be the creation of a unitary system which is distinct from both L1 and L2 (Pavlenko, 2004). This is a specific prediction of the Speech Learning Model (SLM) (Flege, 2005).

If confirmed, these results would be consistent with the SLM. The same processes through which the learners acquired their L1 would be active and available even as adults, and consonant production would not be a static and unchangeable neural commitment (Flege, Schirru, and MacKay, 2003; Kuhl et al., 2008). Aspects of the NLM-e are questioned by these results, although the intermediate nature of the participants’ /p/ VOT leaves open the possibility of L1 commitments preventing a full shift to native French /p/ VOT. The rigidity of MT fares worse.

Still more damaging to the intimate relationship between perception and production proposed by Motor Theory are on the perception results. There is little indication that the production training had an effect on the participants’ ability to differentiate between

French /p/ and /b/, especially in noise (see Table 2.6). The participants did improve slightly in the post-test, but the marginally significant effect of time on the results could simply be due to habituation. The fact that the words in the minimal pairs task were not in the trainings or baseline lists could also have had an effect, indicating that the generalisation to novel words seen in the production of /p/ in 'Cendrillon' was not carried over into perception.

According to MT, the production training should have instructed the participants' perceptual module with the same gestures that encode French /p/ for both. It should not have been necessary to train the participants directly for perception. According to these initial results, this is not the case. The production training was not sufficient to change the perception of the participants in a significant way. In fact, it appears that on average the participants got worse at identifying /p/ and a bit better at identifying /b/ (driving the overall marginal improvement). The participants also did worse on average with the /p/ identification than with /b/, despite the fact that their production VOT did not change significantly for the latter.

This is perhaps not surprising. A wealth of L2-learning literature supports the idea that training production is effective only for changing production (Hattori and Iverson, 2010; Alshangiti and Evans, 2014). The results for the perceptual minimal pairs tasks in this experiment support the idea that the production and perception processes function somewhat independently when training for improvements in an L2.

These results may provide answers for some of the basic questions about speech perception in L2 learning, but further questions arise. What exactly did the participants in this experiment learn, and how does this relate to L2 learning? Theories about learning as described above could help to shed light on the results of this experiment. It is important to tease out whether the observed differences are categorical in nature (indicating a change on the representational level) or if there might be a simpler explanation. It is possible that the participants simply learned to manipulate their English /p/ to sound more 'French-like.' It is well-documented that conversational convergence effects happen in human social interactions, especially in the case of phonetic convergence (Pardo, 2006). The evidence suggests that people often manipulate the acoustic properties of the speech signal to converge with other speakers (Pardo et al., 2012).

The imitation required in this experiment may also be of interest from the standpoint of language learning. As discussed previously, there is some evidence to suggest that imitation of non-native speech (D2 or L2) can influence the attitude of speakers toward the target D2/L2 (Adank et al., 2013). It has also been claimed that imitation of an unfamiliar accent can aid in later comprehension, and furthermore that the areas of the brain that imitation activates (the Inferior Frontal Gyrus, Supplementary Motor Area, and left Superior Temporal Sulcus) are all related to sensorimotor integration (Adank, Rueschemeyer and Bekkering, 2013). Motor activity of this type during perception would appear to support MT and its claims.

According to that imitation research, the results of this experiment should have indicated that imitation of the French /p/ and /b/ would directly aid the participants to comprehend them better. The results presented here seem to contradict those assertions. No access to brain imaging is necessary in this case to investigate what parts of the brain may have been activated by the tasks in French; the lack of any significant change in perception that correlated with the observed change in production is sufficient to challenge the claims. Imitation did not improve comprehension in this case, indicating that the ties between production and perception are not as close as some may claim.

What could be happening in this experiment is that the change observed is the result of the participants learning a new motor pattern for /p/ production. This motor pattern could be associated through the trainings with both the orthographic and auditory French /p/, as indicated by the rapid and persistent shift in VOT duration. There would be no change to the underlying representation of the phoneme in the participants, save for some marking for ‘French’ that corresponded to the new motor pattern. Despite being partially based on motor learning, the new /p/ production pattern would not instruct the perceptual systems of the participants (as MT would propose).

This learned motor pattern could be activated when the participants consciously went into ‘French mode,’ a conscious decision that they were attempting to change their production. In a sense, the additional new /p/ would require ‘translation’ into the surface French state, with steps between the underlying representation (closer to English) to the production of a French consonant. This may be reflected in the intermediate character of the /p/ VOT observed in the participants, distinctly different from the baseline English values and yet not an entirely perfect match for a native French VOT value.

A final consideration is the apparent lack of correlation between tolerance of ambiguity and changes in the participants’ learning of this French consonant contrast. There was no statistical interaction between the TOA scores and the improvements in production or any perception. Viewed in light of the NLM-e’s emphasis on social interaction as a crucial part of the language learning process, this is a curious result.

Despite the long establishment of TOA as a personality indicator and its previously-demonstrated correlation to L2 success (e.g. Lee, 1998), it appears that in this study there was no effect of this particular social characteristic ($p > 0.05$ for production, perception, no significant interactions). This is possibly due to the small number of participants in this initial study, or may be influenced by the close grouping of participant scores around the median (see Table 3.6, Appendix 2). There may not have been sufficient differentiation in this sample of the population to demonstrate an effect.

One specific score from this group is of interest; the participant with the lowest overall TOA score (s05--2.82) was also the one who essentially refused to complete the second training. This participant gave no indication at the time of the experiment that she had particular difficulty with the task. Given that the TOA score is meant to reveal attitudes about new experiences and novel social settings, the low score may indicate that this participant was particularly likely to dislike the experiment and refuse to complete the task (Bochner, 1965).

There are limitations to these results, however. It remains to be seen whether the significant production effects demonstrated in the training and post-test tasks would be retained. There is some evidence to indicate that any improvement might disappear after time away from the training, as appears to occur in short-term L1 accent accommodation (e.g., Alshangiti and Evans, 2011).

Nonetheless, in the accent training study from which this experiment drew some inspiration (German, Carlson and Pierrehumbert, 2013), it appeared that the rapid acquisition of a new phoneme could persist into multiple training sessions over a period of weeks. Further work with more participants, using the same experimental procedure with the addition of at least one more training session with baseline recordings would address this. In addition, further generalisation tasks could be added to investigate the extent to which the /p/ production changes were applicable to novel words.

Another limitation to these rather robust results is the minimal pairs task itself. Due to initial pilots of this task resulting in near 100% accuracy with monolingual English speakers, it was necessary to increase the difficulty by adding speech-shaped noise. Despite the fact that this is standard practice for dialect and language comprehension studies (e.g. Adank et al., 2009), it is possible in retrospect that in this particular case the noise may have been slightly problematic. It is possible that the type of noise used in this experiment could have masked relevant information about the /p/ phoneme, and that this is why, despite the improvement in /p/ production, participants did better at the pre and post-test with /b/.The placement of the burst is highly significant to /p/ comprehension, and this could have been lost in the noise, especially at higher intensities, where it is clear that participants were simply guessing. This could be investigated further by measuring the participants' reaction times for the identification of /p/ and /b/ or using a categorical boundary task to investigate whether or not participants changed their VOT boundary as a result of training.

Conclusion

Further investigation of these results is needed. While the exact mechanism for the learning observed is not clear, it should be obvious that the discrepancies between perception and production could have significant implications for major theories. Motor Theory in particular would be threatened if these results can be replicated and expanded to a large sample size. If it could be shown that imitation training changes only production and not perception, then the Speech Learning Model would be closer to vindication. Given that this experiment demonstrated a clear ability on the part of all the participants to adjust their production to an L2 despite their ages, it is already better supported than MT and the NLM-e.

Beyond the scope of laboratory research, this experiment and others like it give insight into L2 learning in general. If replicated and expanded it might be possible to demonstrate the best ways to improve perception and production in a foreign language, especially for adult learners. This would apply directly to the immediate and pressing needs of the estimated billions attempting to learn additional languages as adults in the 21st century. Greater collaboration between disciplines (neurolinguistics, sociophonetics, speech perception, and L2-learning disciplines such as EFL) will lead to the richest picture of this complex and important process.

Word Count: 9828 (excludes tables, captions, bibliography)

Appendix 1: Stimuli

Table 1.1-Words for Training 1, with translations to English

Word	Translation	Word	Translation
baleine	whale	pain	bread
ballet	ballet	pantalon	trousers
barre	bar/railing	poisson	fish
bébé	baby	porte	door
bâton	stick	paille	straw
balle	bullet	pied	foot
barbe	beard	poupée	doll
boite	box	père	father
banc	bench	parc	park
bandeau	flag	pomme	apple
but	goal	poule	hen
beurre	butter	policier	police officer
piscine	pool		

Table 1.2- Translated BKB sentences for Training 2, with translation to English.

Standardised Sentence	English Sentence
La MAISON avait NEUF PIECES	The HOUSE had NINE ROOMS

Les TOMATES VERTES sont PETITES	The GREEN TOMATOES are SMALL
Le GARCON FAIT le POIRIER	The BOY DID a HANDSTAND
L'ECHELLE est PRES de la PORTE	The LADDER'S NEAR the DOOR
Ils REGARDENT le TRAIN PASSER	THEY'RE WATCHING the TRAIN
Le CHIEN JOUAIT avec un BATON	The DOG PLAYED with a STICK
Ils DISENT des CHOSES BETES	THEY SAY some SILLY THINGS
Elle AVAIT son ARGENT de POCHE	SHE had her POCKET MONEY
Les TROIS filles ECOUTENT BIEN	The THREE GIRLS are LISTENING
La TABLE a TROIS PIEDS	The TABLE has THREE LEGS
Il ECOUTE SON PERE	HE LISTENS TO his FATHER
Les DEUX FERMIERS se PARLENT	The TWO FARMERS are TALKING
Elle ATTEND SON BUS	SHE'S WAITING for her BUS
Un HOMME a PREVENU la POLICE	A MAN TOLD the POLICE
Il SUCE SON POUCE	HE'S SUCKING his THUMB
Le CHIOT JOUE avec une BALLE	The PUPPY PLAYS with a BALL
L'enfant VEUT son BIBERON	The BABY WANTS his BOTTLE
La TARTE au FROMAGE était BONNE	The CHEESE PIE was GOOD
Ils se SONT ASSIS sur un BANC	THEY SAT on a BENCH
Le CHAT mange dans un BOL	The CAT EATS from a BOWL.

Table 1.3- Minimal pairs, for the minimal pair identification task. Translated to English.

Minimal Pair	Translation to English
bois-pois	wood-peas
bas-pas	low-step
bar-par	bar-through
baisse-paisse	drop-graze
boire-poire	to drink-pear
bond-pond	jump-lay eggs (3rd person singular)
boudoir-pouvoir	boudoir- to be able
boue-pou	mud-louse
banne-panne	awning-failure
bot-pot	club foot-jar
ban-pan	banns(marriage)-section
belle-pelle	pretty-shovel

battre-pâtre	to battle-shepard
bâti-pâtis	assembly-grazing

Table 1.4 - Words for English baseline recordings.

Word	Distractor or Target
peak	Target
beak	Target
pack	Target
back	Target
park	Target
peck	Target
log	Distractor
buck	Target
beck	Target
seem	Distractor
thumb	Distractor
luck	Distractor
puck	Target
go	Distractor
bark	Target

Table 1.5- Words for the French baseline recordings.

Word	Distractor or Target
pique	Target
père	Target
bec	Target
mieux	Distractor
bulle	Target
passe	Target
parc	Target
livre	Distractor
basse	Target
jeudi	Distractor

puce	Target
bique	Target
gorge	Distractor
barque	Target
temps	Distractor

Table 1.6 - Words for English post-test recordings.

Word	**Distractor or Target**
beak	Target
bark	Target
park	Target
back	Target
peck	Target
seem	Distractor
buck	Target
pack	Target
beck	Target
log	Distractor
thumb	Distractor
luck	Distractor
puck	Target
go	Distractor
peak	Target

Table 1.7 - Words for French post-test recordings.

Word	**Distractor or Target**
barque	Target
père	Target
mieux	Distractor

bec	Target
bulle	Target
passe	Target
livre	Distractor
basse	Target
jeudi	Distractor
parc	Target
puce	Target
bique	Target
gorge	Distractor
temps	Distractor
pique	Target

Text Sample: Cendrillon
Cendrillon

La femme d'un homme riche tomba malade. Elle bien savait qu'elle allait mourir, et puis elle demanda a sa fille unique de venir a son chevet. Elle lui dit: "Si tu restes toujours bonne, je veillerai sur toi depuis le ciel."

Peu apres, elle ferma les yeux et mourut. On l'enterra dans le jardin. La petite fille se rendait sur la tombe chaque jour pour pleurer, et était toujours bonne et gentille avec toutes les personnes de son entourage. Lorsque la neige tombait, la tombe paraissait belle, mais quand le primtemps arriva son père s'etait remarrié. Sa nouvelle femme avait deja deux filles qui vinrent vivre avec elle. Elles étaient jolies à l'éxterieur, mais horrible à l'interieur, et pour la pauvre petite fille ce fut le commencement de mauvais traitement.

KidLit-O, (2013)

Appendix 2- Results

Table 3.1- Tolerance of Ambiguity scores for participants. 'High tolerance' in bold.

Subject	ToA score
S01	3.45
S02	4.00
S03	3.18
S04	4.09
S05	2.82
S06	4.18
S07	4.00
S08	3.45
S09	3.45
S10	4.18
MEDIAN	3.73

Table 3.2 - Prevoicing tokens in French for the pre and post tests. T-test indicates that the results are not significant.

Pre-voicing

nPrevoiced	pre-test	post-test
S01	0	0
S02	0	4
S03	0	0
S04	0	0
S05	1	2
S06	1	1
S07	0	0
S08	2	0
S09	0	0
S10	3	4
	7	11
	T-test	0.42257163

Bibliography

Abramson, A. (1975). Laryngeal timing in consonant distinctions. pp.17--23.

Adank, P., Evans, B., Stuart-Smith, J. and Scott, S. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), p.520.

Adank, P., Hagoort, P. and Bekkering, H. (2010). Imitation improves language comprehension. *Psychological science*, 21(12), pp.1903--1909.

Adank, P., Rueschemeyer, S. and Bekkering, H. (2013). The role of accent imitation in sensorimotor integration during processing of intelligible speech. *Frontiers in human neuroscience*, 7.

Adank, P., Stewart, A., Connell, L. and Wood, J. (2013). Accent imitation positively affects language attitudes. *Frontiers in psychology*, 4.

Alshangiti, W. and Evans, B. (2011). Regional accent accommodation in spontaneous speech: Evidence for long-term accent change?. In: Proceedings of the 17th International Congress of Phonetic Sciences.

Andersen, S. and Schwartz, A. (1992). Intolerance of ambiguity and depression: A cognitive vulnerability factor linked to hopelessness. *Social Cognition*, 10(3), pp.271--298.

Aoyama, K., Flege, J., Guion, S., Akahane-Yamada, R. and Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r. *Journal of Phonetics*, 32(2), pp.233--250.

Aslin, R., Pisoni, D., Hennessy, B. and Perey, A. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child development*, 52(4), p.1135.

Bench, J., Kowal, \. and Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British journal of audiology*, 13(3), pp.108--112.

Best, C. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, 167, p.224.

Bleakley, H. and Chin, A. (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. *American economic journal. Applied economics*, 2(1), p.165.

Bochner, S. (1965). Defining intolerance of ambiguity. *Psychological Record*, 15.3, pp.393-400.

Brown, A. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, 14(1), pp.107--133.

Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of personality*.

Chapelle, C. and Roberts, C. (1986). Ambiguity tolerance and field independence as predictors of proficiency in English as a second language. *Language learning*, 36(1), pp.27--45.

Clopper, C. and Bradlow, A. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and speech*, 51(3), pp.175--198.

Delvaux, V. and Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, 64(2-3), pp.145--173.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language learning*, 27(2), pp.315--330.

Eimas, P., Siqueland, E., Jusczyk, P., Vigorito, J. and others, (1971). Speech perception in infants. *Science*, 171(3968), pp.303--306.

Ellis, R. (1994). *The study of second language acquisition*. 1st ed. Oxford: Oxford University Press.

Evans, B. and Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*, 121(6), pp.3814--3826.

Flege, J. and Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6(3), pp.185--202.

Flege, J. and Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24(2), pp.125--146.

Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sung, H. and Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, 34(2), pp.153--175.

Flege, J., MacKay, I. and Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 106(5), pp.2973--2987.

Flege, J., Schirru, C. and MacKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, 40(4), pp.467--491.

Flege, J. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of phonetics*, 15(1), pp.47--65.

Flege, J. (1999). Age of learning and second language speech. *Second language acquisition and the critical period hypothesis*, pp.101--131.

Flege, J. (2005). Origins and Development of the Speech Learning Model. In: *1st ASA Workshop on L2 Speech Learning*. Vancouver, BC.

Fowler, C., Sramko, V., Ostry, D., Rowland, S. and Hall\'e, P. (2008). Cross language phonetic influences on the speech of French--English bilinguals. *Journal of Phonetics*, 36(4), pp.649--663.

Furnham, A. and Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current Psychology*, 14(3), pp.179--199.

Galantucci, B., Fowler, C. and Turvey, M. (2006). The motor theory of speech perception reviewed. *Psychonomic bulletin \& review*, 13(3), pp.361--377.

German, J., Carlson, K. and Pierrehumbert, J. (2013). Reassignment of consonant allophones in rapid dialect acquisition. *Journal of Phonetics*, 41(3), pp.228--248.

Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105(2), p.251.

Gottfried, T. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*.

Hattori, K. and Iverson, P. (2010). Examination of the relationship between L2 perception and production: An investigation of English/r/-/l/perception and production by adult Japanese speakers.

Hattori, K. (2010). Perception and production of English/r/-/l/by adult Japanese speakers. *Doctoral thesis.*, UCL (University College London).

Herschensohn, J. (2000). *The second time around*. 1st ed. Amsterdam: J. Benjamins Pub.

Hirschfeld, L. and Gelman, S. (1994). *Mapping the mind*. 1st ed. Cambridge: Cambridge University Press.

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. and Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87(1), pp.47--57.

KidLit-O, (2013). *Le meilleur des frères Grimm En français d'aujourd'hui (French Edition)*. 1st ed. CreateSpace Independent Publishing Platform.

Komarova, N. and Nowak, M. (2001). Natural selection of the critical period for language acquisition. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1472), pp.1189--1196.

Kuhl, P., Conboy, B., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. and 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 B: Biological Sciences*, 363(1493), pp.979--1000.

Kuhl, P. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception \& psychophysics*, 50(2), pp.93--107.

Kuhl, P. (1992). Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes. *American Psychological Association*.

Kuhl, P. (2000). Language, mind, and brain: Experience alters perception. *The new cognitive neurosciences*, 2, pp.99--115.

Kuhl, P. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience*, 5(11), pp.831--843.

Lee, J. (1998). Language learning strategies and tolerance of ambiguity of Korean midshipmen learning English as a foreign language.

Leyro, T., Zvolensky, M. and Bernstein, A. (2010). Distress tolerance and psychopathological symptoms and disorders: a review of the empirical literature among adults. *Psychological bulletin*, 136(4), p.576.

Liberman, A. and Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), pp.1--36.

Liberman, A. and Whalen, D. (2000). On the relation of speech to language. *Trends in cognitive sciences*, 4(5), pp.187--196.

Liberman, A., Cooper, F., Shankweiler, D. and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, 74(6), p.431.

Liberman, A. (1996). *Speech*. 1st ed. Cambridge, Mass.: MIT Press.

Lisker, L. and Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), pp.384--422.

Mac Donald Jr, A. (1970). Revised scale for ambiguity tolerance: Reliability and validity. *Psychological reports*, 26(3), pp.791--798.

McCandliss, B., Fiez, J., Protopapas, A., Conway, M. and McClelland, J. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, \& Behavioral Neuroscience*, 2(2), pp.89--108.

McLaughlin, B. (2013). *Second Language Acquisition in Childhood*. 1st ed. Hoboken: Taylor and Francis.

Morgan-Short, K., Finger, I., Grey, S. and Ullman, M. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PloS one*, 7(3), p.32974.

Pallier, C., Dehaene, S., Poline, J., LeBihan, D., Argenti, A., Dupoux, E. and Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first?. *Cerebral cortex*, 13(2), pp.155--161.

Pardo, J., Gibbons, R., Suppes, A. and Krauss, R. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40(1), pp.190--197.

Pardo, J. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), pp.2382--2393.

Pardo, J. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), pp.2382--2393.

Pavlenko, A. (2004). L2 influence and L1 attrition in adult bilingualism. *First language attrition: Interdisciplinary perspectives on methodological issues*, pp.47--59.

Piske, T., MacKay, I. and Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of phonetics*, 29(2), pp.191--215.

Rice, D. and Barone Jr, S. (2000). Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environmental health perspectives*, 108(Suppl 3), p.511.

Schmid, M. and Keijzer, M. (2009). First language attrition and reversion among older migrants. *International Journal of the Sociology of Language*, 2009(200), pp.83--101.

Schmid, M. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, 13(01), pp.1--7.

Scobbie, J. (2006). Flexibility in the face of incompatible English VOT systems. *Mouton de Gruyter*.

Serniclaes, W., Sprenger-Charolles, L., Carr\'e, R. and Demonet, J. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, 44(2), pp.384--399.

Tikofsky, R. (1968). Eric H. Lenneberg, Biological foundations of language. New York: John Wiley and Sons, 1967. *Wiley Online Library*.

Upshur, J. (1962). Language proficiency testing and the contrastive analysis dilemma. *Language Learning*, 12(2), pp.123--127.

Yao, Y. (2009). Understanding VOT variation in spontaneous speech. *Proceedings of the 18th International Congress of Linguists (CIL XVIII)*. 2009.