Perceptual Learning in Second Language Learners

A Dissertation Presented

by

Katharina Sophie Schuhmann

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

December 2014

Stony Brook University
The Graduate School

Katharina Sophie Schuhmann

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Marie K. Huffman, Ph.D. – Dissertation Advisor
Associate Professor, Department of Linguistics

Ellen I. Broselow, Ph.D. - Chairperson of Defense
Professor, Department of Linguistics

Arthur G. Samuel, Ph.D.
Professor, Department of Psychology

Rachel Hayes-Harb, Ph.D.
Associate Professor, Department of Linguistics, University of Utah

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Perceptual Learning in Second Language Learners

by

Katharina Sophie Schuhmann

Doctor of Philosophy

in

Linguistics

Stony Brook University

2014

This dissertation studied the flexibility of linguistic representations in monolingual and bilingual speakers of English. We conducted four perceptual learning studies to determine how monolingual English and English-German bilingual listeners mentally represent fricative phonemes. Listeners first completed an auditory lexical decision task in English in which critical stimuli contained either an /f/ or /s/ that had been replaced with a mixture in between [f] and [s]. Subsequently, listeners completed forced-choice phoneme categorization tasks to test for perceptual learning on the trained English /f-s/ contrast and possible generalization to other within-language contrasts, and possible cross-language generalization. We hypothesized (a) that perceptual learning in monolinguals would generalize across phonological features if the relevant phoneme contrast is signaled by similar acoustic-phonetic cues, and (b) that perceptual learning would generalize from English to German because the phonetic properties important to fricative contrasts in the two languages are similar.

We found evidence of perceptual learning, and some generalization across phonemes and/or languages, in a complex pattern that suggests an important influence of type of bilingual experience. The monolingual English listeners showed perceptual learning on English /f-s/ and generalized the effect to the /v-z/ contrast, as predicted. Novice L1 English – L2 German speakers in the US (study 2 & 3) also showed perceptual learning on the trained English /f-s/ contrast. In addition, listeners in study 2 showed no perceptual learning in German, while participants in study 3, who were in a somewhat more bilingual language mode (Grosjean 1997, 2001), did show perceptual learning effects on German /f-s/ and German /v-z/. In study 4, intermediate-to-advanced L1 German – L2 English speakers in Germany who were in a bilingual language mode, showed perceptual learning on English /f-s/ and German /f-s/, but not on the voiced /v-z/ fricative contrast in either language.

These bilingual results are explained with a model in which phonemes common to two languages have separate but dynamically associated representations. A bilingual mode strengthens the interconnections between phonemes, thus facilitating cross-linguistic effects. Effects are strongest in L2 sound systems when perceptual learning generalizes from the dominant L1 to the non-dominant, novice L2. Finally, non-native listeners adjust the representation of phoneme boundaries in their L2 at the level of individual phoneme contrasts, and do not generalize these adaptation effects to phoneme contrasts that share relevant phonological features.

To my father.

Who has inspired me with his love for sports and for languages.

Who taught me my first English words:

“Come on!”

so I could play with the neighbors’ boy.

Who was the first to answer endless questions

about how we acquire foreign languages.

And who has taught me

how to persevere.

To my mother.

Who has inspired me with her dedication

to her family, to her career, to politics, and to community service.

Who has taught me the significance of

natural prowess,

contentedness, gracefulness, and dignity.

And who has taught me

how to assert myself,

and help those in need.

To Irina Izverna-Tărăbac (1970-2007).

Whose memory continues to inspire me,

because of her love for scientific endeavor

and her caring friendship.

Irina has shown me

how beautifully

science and benevolence can complement one another.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am tremendously grateful to everyone who has supported me, has been a friend or mentor to me, or has cheered me on while I have been working on my dissertation. My deepest gratitude is to my wonderful advisor, Dr. Marie Huffman, for her support and guidance throughout the entire dissertation process. Marie believed in me and in my ideas from day one, before I even thought about a dissertation topic, and continued to support me while I was far away, either collecting data in Germany or elsewhere in the US, while I lived in Oklahoma, and while I have been living and teaching in Pennsylvania. I will cherish our many hour-long conversations and discussions – whether in person, on the telephone, or via Skype – about perceptual learning, second language acquisition, career plans, and life in general. Marie managed to push me further than I ever thought possible, and she gave me the necessary criticism that made my NSF grant proposal, dissertation defense presentation, and written dissertation *so* much stronger! This dissertation would not have been possible without her mentorship, advice, and wonderful sense of humor. Shortly after Marie became my advisor, she "warned" me that a Ph.D. advisor is for life! Lucky me. What's more, I am fortunate enough to already be working with Marie (and some of my students) on our next project together! How much better can it get?

I am also extremely thankful to the chair of my dissertation defense and my dissertation committee member, Dr. Ellen Broselow. Ellen encouraged me to take more time to read broadly in the field of second language acquisition and phonology, which helped my understanding of and furthered my interest in the field of second language phonology tremendously. Ellen has vast knowledge of the relevant literature, and has continuously made important suggestions about which articles might be significant for my own research program. I aspire to become a linguist with as wide a grasp of the literature, from the classics, to what is hot off the press, or even still *in press*.

My dissertation would not have been possible in its current form without the seminal work of Dr. Arty Samuel (and his colleague Tanya Kraljic). Moreover, the design of the current study and my previous work in the area of perceptual learning and bilingualism benefitted greatly from many discussions with Arty – concerning everything from methodological considerations of the study design, relevant pieces of literature about perceptual learning and the psycholinguistics of speech perception, critical questions (not) to ask, and data interpretation. I am so grateful for Arty's inspirational work, his advice about how to conduct psycholinguistic studies, and his insightful and abundant comments about my research design, data analysis, and the implications of my research findings within the broader picture of psycholinguistic literature on speech perception and second language acquisition.

Finally, I am extremely lucky for Ellen's suggestion to ask Dr. Rachel Hayes-Harb to be my outside committee member. Rachel has provided a fresh look at, and an unbiased opinion of, my work, and has challenged me to be very clear and explicit about the notion of features whenever I refer to them in my work. Most importantly, I am looking forward to visiting Rachel's SLA lab in Utah to present some of my work, to engage with her very active team-members, and to collect additional data from L1 English – L2 German speakers. Words are hardly enough to describe my gratefulness to Rachel for extending this wonderful invitation to me. I cannot wait to see what my visits to Utah will bring.

I am also indebted to Dr. Susan E. Brennan, with whom I've had the pleasure of working on another NSF-funded project on speech adaptation in natural dialogues in various contexts. We are still actively in the process of analyzing some of the data, writing up additional analyses for publication, and are in the process of releasing the collected data as a corpus with the LDC. I have also had the pleasure of serving as a Teaching Assistant for her, and of taking several psycholinguistics courses with her, which ignited the initial spark in me to work on perceptual learning in a bilingual population. I look forward to many more hours of collaboration, many conversations about academia and life, and hopefully more opportunities to go kayaking together.

My transition to an American institution of Higher Education would not have been the same without Dr. Barbara Brownworth. She has played the most important role for me in my education about Teaching English as a Foreign Language and about the history and diversity of teaching philosophies and methodologies. Most importantly, she almost single-handedly trained me as a language assessment specialist for non-native English speakers and helped me to grow as a foreign language and ESL teacher. Ever since I have moved away from New York, I have missed our many heartfelt conversations about TESOL, teaching ESL, linguistics, family, and life. I hope our friendship and professional bond will last over the coming years and decades.

Among the many other professors who have supported me, inspired me, and cheered me on are: Dr. John F. Bailyn, Dr. Christina Bethin, Dr. Richard K. Larson, Dr. Lori Repetti, and Dr. Joy Janzen. I am so grateful to each and every one of them for *all* of the knowledge I gleaned from their wonderful courses, and for all the challenging assignments, exams, and papers. I know I have grown immensely in my intellectual capacities, including analytical skills and writing skills, and in my ability to handle criticism. Thank you for being such stimulating teachers – while being supportive and full of praise (when and where appropriate).

I am also thankful to the colleagues and friends at Stony Brook and around the globe who have supported me in various ways throughout my PhD program and my dissertation. I would like to thank, in particular, Vahideh Rasekhi, who's not only been my dearest friend in the program, but has also agreed, time and time again, to handle *a ton* of paperwork and administrative necessities for me, after I moved away from Stony Brook, not to mention the numerous times when she allowed me to crash on her couch when I came back for short visits. I am extremely happy for the many friendships and colleagueships that were maintained or developed throughout my time as a graduate student at Stony Brook. In particular, I'd like to mention Izverna-Tărăbac and her family in Romania, Dr. Meghan Sumner, Dr. Elizabeth John, Jay Ramachandran, Dr. Maruti Sharma, Dr. Sachin Jambawalikar, Reza Basseda, Dr. Allison Luengen, Dan Yang, Dr. Anna Kuhlen, Dr. Jadzia Jagiellowicz, Dr. Chih-hsiang Shu, Dr. Shih-chi (Stella) Yeh, Dr. Miran Kim, Dr. Jiwon Hwang, Dr. Karla Batres, Franco Amati, Dr. Julie Kolmogorova Weisenberg, Dr. Young-Ran An, Dr. Yu-An Lu, Dr. Andrei Antonenko, Dr. Roksolana Mykhaylyk, Dr. Xiaomeng (Mona) Xu, Ling Wang, Poppy Slocum, Dr. Georgia Weissman Zellou, Dr. Hyun-ju Kim, Achinya Jackson, Dr. Hijo Kang, Susana Huidobro, Ellen Guigelaar, Dr. Yukiko Asano, Dr. Tomoko Kawamura, Svitlana Antonyuk-Yudina, Dr. Yunju Suh, Dr. Tanya Scott, Hisako Takahashi, Zahra Basseda, Heather Ball, Mark Lindsay, Dr. Lori Moon, Dr. Anne-Laure Hentz, Dr. Betty Beran, Julian Ang, Roland Sämann, Gülay Alpaslan, Lude Distel, and Patricia Hellmuth.

I am also thankful for all the administrative support from the staff in the Linguistics Department at Stony Brook University, Michelle Carbone and Sandra Brennan, and Pam Glass in the Department of Languages, Cultures and Linguistics at Bucknell University. I am also thankful for the helpfulness and assistance I've received from everyone at International Services at Stony Brook University, in particular Christine Greenberg and Dr. Elizabeth Barnum (now at

UNC at Chapel Hill). Dr. Barnum championed for me ever since I first arrived at Stony Brook. I will forever cherish all her support as well as the many conversations we have had at various events for international students and scholars at Stony Brook University, at private gatherings, or while I was working as a Graduate Assistant in the Graduate School and in International Services.

I have also enjoyed collaborating with the Faculty Center at Stony Brook University, in particular with Nancy Wozniak. Finally, I am grateful to my colleagues with whom I have worked and shared many great moments at Nuance Communications in New York City in the summer of 2013, most importantly Dr. Jon Bloom, Dr. Alex Christodoulou, Dr. Tanya Kraljic, Gorm Amand, and Tom Hebner.

My dissertation process has been made a lot easier through the generous support of various research grants. With the support of Dr. Huffman, I was awarded an NSF Dissertation Improvement Grant (Award ID #: 1127718), a DAAD (*Deutscher Akademischer Austausch Dienst*/German Academic Exchange Service) two-month research grant to Germany, and a TOEFL Small Grant for Doctoral Research (ETS: Educational Testing Service, Princeton, NJ). Additionally, I cannot express enough of my gratitude to everyone who has allowed me to come visit and collect data in their labs, department, or universities. I will be forever indebted for Prof. Dr. Barbara Höhle's and Dr. Natalie Boll-Avetisyan's help at Potsdam University (Potsdam, Germany), Prof. Dr. Pompino-Marschall's and Peter Skupinski's help at Humboldt University (Berlin, Germany), Prof. Dr. Irwin Collier's and Sabine Riefler's help at the JFK Institute at the Freie Universität (Berlin, Germany), and Prof. Dr. Niklaus Largier's help at the University of California at Berkeley. Thank you so much!

For the last year and a half, my fiancé Kent and I have lived in Pennsylvania, where I have been teaching Linguistics at Bucknell University (Fall 2013 & Fall 2014). I am so thankful to the department for granting me an office space during the spring and summer of 2014, while I was finishing my first complete draft of this dissertation. The office, library privileges and, most importantly, colleagueship made me feel very much at home! I am glad that Or (Dr. Or Rogovin) was around most of the summer of 2014 to work on his scholarship. This helped me stay focused on my dissertation, but also allowed me to take breaks and have fun conversations from time to time. I have also enjoyed getting to know the other psycholinguist at Bucknell, Dr. Heidi Lorimor, and look forward to hopefully bringing our idea for a research project to fruition very

soon. I am grateful that my colleague in Linguistics, Dr. Jim Lavine, has become both a mentor and a friend to me throughout my time at Bucknell. I have had the pleasure of talking with Jim for countless hours, and value his time, advice, and humor tremendously. We will be sad to leave Bucknell, as we have made many new friends, including Dr. Jim Lavine & Dr. Ludmila Shleyfer Lavine, Dr. Nikki Young & Dr. Jodi Black, Dr. Katharina & Dr. Ben Vollmayr-Lee, Dr. Helen & Dr. Peter Morris-Keitel, Dr. Ann Tlusty & Dr. Helmut Graser, Dr. Logan Connors, Dr. Bastian Heinsohn, Laura Lanwermeyer & Dr. Matthew Slater, Dr. Bernhard Kuhn & Dr. Emek Ucarer, Dr. Cara Takakjian & Nate Smiley, Dr. Vincent Stephens, Dr. Ashli Baker, Dr. Rivka Ulmer, and all our other *Stammtisch*-friends.

Last but not least, I am truly indebted and forever thankful to each and every one of my family members for their relentless love and support, including all our furry family members who have reminded us to pause, play, and be compassionate! I am grateful to Dr. Edward & Cynthia Gustavson for allowing us to live in their charming A-frame on Lake Fort Gibson for a year and a half. We had a wonderful time in Oklahoma, which allowed us to spend more time with Kent's family, including Britta, Kevin (Dr. Kevin Gustavson) and sweet Svea, go canoeing on the lake nearly every day, explore Cherokee culture, and get used to running and racing in the hottest climate a German could possibly endure. During my time in Berlin (2011/2012), I was also able to spend a lot of fun time with Uler Knöß and Christoph Bast. Our many fun adventures in and around Berlin will remain with me forever, but most importantly, you helped me put things in perspective! I am lucky to have had four loving grandparents in my life, though I am saddened that none of them are with us this fall to share in this joyous moment. I am grateful to my brother Chris, for opening my eyes to the science of art and design, for showing me around Weimar during one of my visits to Berlin, and for many open and honest conversations over recent years.

I owe my biggest gratitude to my wonderful parents Hildegard Schuhmann-Knöß and Bernd Schuhmann, for always believing in me, for their relentless love, *Fürsorge*, regular post-cards, letters, and phone calls from *good ol' Germany*, and for supporting me in my dreams to travel around the world, study abroad – and then stay abroad for many more years than anyone expected when I packed my two suitcases for my Fulbright year! Thank you for always loving me, no matter what, and for being my biggest cheerleaders! And of course, I will be forever thankful to Kent[*], my soul mate and the love of my life, for helping me get to the finish line(s)! Whenever I did not feel like I wanted to push through, Kent found a way to build me up and push me on (yes, *sometimes* a bag full of chocolate and gummy bears will do the trick), he has found the most creative ways to distract me from my work when I needed to take a break, and he always has a broad shoulder to lean on, and a big hug after a long day! Thank you for loving me so much, despite knowing me more than anyone else does. I love you, *Schatz*.

The list goes on, and it is hard to stop myself here. For now, here is a heartfelt *Thank You,* a German-accented *Grazie,* and a Franconian *Gott vergelt's und a' herzliches Dankeschön!*

[*] Technically, Dr. Kent Samuel Gustavson

1. Introduction

One of the central goals of this dissertation is to establish how bilingual listeners mentally represent and adjust phoneme contrasts which are common to both their native and their non-native language and which do not differ considerably in their phonetic-acoustic realizations. The classic perceptual learning paradigm (reviewed below) provides a tool to probe the long-standing question about whether the languages of second-language learners are separate, merged, or interconnected. In short, perceptual learning studies expose listeners to non-canonical realizations of a particular speech sound and measure whether listeners retune their perception of the relevant – and possibly other related – phonemes. The results of the present work suggest that L1-L2 listeners have phonological representations that are separate for each language (reminiscent of Escudero 2007b) yet are part of a common phonological space (as in Flege's (1995b) Speech Learning Model, SLM, and Best & Tyler's (2007) Perceptual Assimilation Model, PAM-L2). Additionally, the results of the present work suggest that these equivalent phonemes are dynamically interrelated (cf. de Bot, Lowie, & Verspoor's (2005) application of Dynamic Systems Theory to SLA) such that the listeners' L1 can influence their L2 *and* that their L2 can influence their L1, similar to the bidirectional influence described in Flege's (1995b) SLM.

Speech is characterized by the infamous "lack of invariance" (e.g., Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967; Liberman, Harris, Hoffman & Griffith 1957; cf. Kraljic & Samuel 2005:167) yet listeners report largely stable speech percepts. This lack of invariance refers to the notion that phonemes are acoustically non-identical both in within-speaker and in between-speaker utterances. This is due to a multitude of factors, including co-articulation effects from the surrounding context, allophonic alternations, differences in speakers' gender and anatomical structure, and general effects of the rate of speech and the emotional state of the speaker. Listeners are typically neither aware of nor troubled by this lack of invariance, and are quite skilled at perceiving variable speech input.

It is well established that listeners are adept at listening to and interpreting non-canonical speech. Typically, linguistic context guides the perception of non-canonical speech. A well-known case is the Phoneme Restoration Effect, which demonstrates that listeners believe that they heard a missing phoneme (Warren 1970; Samuel 1981a,b). In these studies, listeners are presented with words in which one sound is spliced out of the recording and replaced with white noise. Listeners consistently report that they think they heard the missing phoneme as well as the white noise when listening to these kinds of modified recordings. In short, listeners are adamant that they heard a speech sound in a situation when the entire speech sound was replaced with white noise. These well-known studies illustrate the remarkable ability in human speech perception to deal with non-canonical speech, to perceive speech in adverse listening conditions, and to even perceptually "restore" sounds that have been excised from a recording.

The influence of lexical content on the ability to perceive and interpret non-canonical speech is particularly revealing when listeners judge ambiguous stimuli embedded within lexical items. For example, the same ambiguously voiced alveolar stop, [?td], i.e. an alveolar stop with a VOT value that is in-between a typical /t/ and a typical /d/ phoneme – is interpreted as /t/ or a /d/ by listeners, depending on whether it is occurs within a lexical item in which it takes the place of a /t/ sound or a lexical item in which it takes the place of a /d/ sound (Ganong 1980). When the ambiguous [?td] sound is followed by [i:k], listeners are more likely to report hearing /t/ – as in the word *teak (*deek)* – but when the ambiguous sound is followed by [i:p], listeners are more likely to report hearing /d/ - as in the word *deep (*teep)* (cf. Cutler 2012:392).

The flexible nature of speech perception has been further studied in a growing body of literature on the topic of "perceptual learning" within the last decade. Overall, this research area has illustrated for various phoneme contrasts and various languages that adult listeners fine-tune their speech perception in their native language (L1) throughout their lifetime. Listeners systematically adapt phonological categories to accommodate accents, dialects, and idiosyncratic speech. Various studies have shown that this is a rapid and automatic adjustment process of which listeners do not seem to be aware.

The classic methodology for perceptual learning consists of two phases. In the first phase, the "training phase", listeners are exposed to words with unusual, non-canonical sounds. For example, in one condition in the studies of this dissertation, study participants hear words like "dinosaur" with an unusual pronunciation of /s/, [?sf], which is in-between [s] and [f]. In another condition, study participants hear a word like "daffodil" with an unusual pronunciation of /f/, [?fs], which is in-between [f] and [s]. The second phase, the "testing phase", examines whether the phoneme boundaries of the listeners are in fact enlarged such that they also include unusual pronunciations of the relevant phoneme boundaries that listeners heard in the first phase of the study. In order to establish whether the relevant perceptual phoneme boundaries have "stretched" in the relevant direction, listeners judge ambiguous sounds on a sound continuum. For example, if participants in a study heard either unusual pronunciations of /f/ or /s/, the participants in both conditions will hear and respond differently to ambiguous fricatives from the same sound continuum (/f/-/s/). Specifically, participants who heard unusual pronunciations of [?s] in words like "dinosaur" typically respond with more /s/ judgments compared to participants who heard unusual pronunciations of [?f] in words like "daffodil" (who, in turn, have more /f/ judgments).

This dissertation sets out to address several open questions in the existing research literature on perceptual learning. First of all, there is the question of the linguistic level at which perceptual learning takes place. The literature on perceptual learning has not arrived at a conclusion about this. While some have reasoned that perceptual learning targets individual segments (Eisner & McQueen 2005), other studies have concluded that perceptual learning takes place at the level of abstract, phonological features (Kraljic & Samuel 2006).

Additional research in this area has suggested that perceptual learning might be more sensitive to fine phonetic detail than these other studies imply, and in fact might not take place when the phonetic-acoustic cues of the relevant phonological features in the stimuli in phase one and those in phase two are not perceived to be sufficiently "similar". In the case of fricatives, perceptual learning has been argued to generalize across voices as long as the voices have acoustically similar fricatives, i.e., similar spectral means for the fricatives (Kraljic & Samuel 2005). Similarly, Reinisch & Holt (2014) demonstrate that perceptual learning on fricatives from the same speaker might or might not be successful, depending on whether the ambiguous fricatives in phase one and those in phase two are "sampled across a similar perceptual space" (Reinisch & Holt 2014:539). This dissertation is designed to test a number of ways in which perceptual learning might generalize, including generalization over voicing or manner contrasts at the same place of articulation. One general question we address is whether perceptual learning can generalize beyond individual segments on which listeners were trained, i.e. whether perceptual learning generalizes to all segments that share the same phonological feature as the ambiguous phoneme contrast on which listeners were trained? On the other hand, is the generalization limited to phoneme contrasts that share the same phonological contrast but also realize this phonological contrast with similar phonetic-acoustic cues? Findings in the current work suggest that listeners' adjustments to unusual pronunciations occur at the level of representation where abstract phonological features are connected to concrete acoustic-phonetic cues, the level seen by some as the interface between phonetics and phonology.

Secondly, this project extends research on perceptual learning in the area of second language learning. Most of the studies in the perceptual learning literature test monolingual, native speakers of the language under consideration. A major contribution of this dissertation to the literature is a group of studies with listeners who have knowledge in a second language and who are tested for perceptual learning effects in two languages after being exposed to non-canonical phones in just one of their languages. One set of studies tests whether native English speakers adjust their phoneme boundaries in English as well as in their L2 German after hearing non-canonical fricatives, [?fs], in English. The data is analyzed together with the listeners' language background, perception skills in their non-native language, and language mode at the time of the study. Another study tests whether native German speakers adjust their phoneme boundaries in their L2 English as well as their L1 German after hearing non-canonical fricatives, [?fs], in English.

Reinisch et al. (2013) showed that both native Dutch listeners as well as native German listeners adjust their f/s phoneme boundaries after listening to Dutch words with non-canonical voiceless fricatives, [?fs], in word-final position. While this study (Reinisch et al. 2013) provides initial evidence that perceptual learning in second language listeners (here L1 German – L2 Dutch listeners) is possible, there are important differences between the experiments reported in Reinisch et al. (2013) and those reported in this dissertation. Besides the difference in language combination and the position of the relevant phoneme-contrast within the word, the L2 Dutch listeners in Reinisch et al. (2013) lived in an L2 environment (in the Netherlands) and were, as a result, presumably advanced speakers of L2 Dutch. The L2 English listeners who participated in the study reported in this dissertation were intermediate-to-advanced L1 German – L2 English speakers who lived in Germany at the time of the study, i.e. in a predominantly German-speaking environment. As a result, the participants in the L1 German – L2 English study in this dissertation use their L2 English less often, are likely less proficient in their L2, and have a lower baseline activation of their L2 (English) compared to the L2 Dutch speakers in the Reinisch et al. (2013) study.

Moreover, the studies involving L2 speakers reported in this dissertation did not only test whether perceptual learning is possible in the listeners' non-native language, but also tested whether perceptual learning generalizes to similar phoneme contrasts within the language of training as well as to similar phoneme contrasts in the untrained language. Although Reinisch et al. (2013) conducted a study in which they tested for perceptual learning effects across languages, they had L1 Dutch speakers listen to non-canonical fricatives, [?fs], in English words which were uttered by an L1 Dutch-L2 English speaker. After exposure to these unusual fricatives in Dutch-accented English, the listeners showed perceptual learning both in the (Dutch-accented) English as well as in their L1 Dutch. The cross-linguistic effect – whereby training in L2 English affected L1 Dutch in these listeners – might have been a result of the fact that a native Dutch speaker produced the English critical stimuli. The listeners might therefore have assumed that any non-canonical speech sounds in the L2 English forms might be due to or linked to sounds in the speaker's L1 Dutch. In all the studies reported here, listeners are exposed to non-canonical fricatives, [?fs] in word-medial position in English words, which are produced by a native English speaker. The listeners are then tested for perceptual learning effects on this specific voiceless fricative contrast in English, as well as the voiced fricative contrast in English (except for the English-German study #1), and the voiceless and the voiced fricative contrasts in German.

Moreover, this dissertation was designed to concentrate on phoneme contrasts that exist both in the listeners' native and in their non-native language, namely the /f/-/s/ contrast and the /v/-/z/ contrast, which exist in both English and German and are realized very similarly in phonetic-acoustic terms[2]. The goal was to address to the question of whether bilinguals have two separate representations for (fricative) phonemes that are common to both of their languages, i.e. one for each language, or whether bilinguals have one representation, i.e. one shared representation for both languages.

In order to achieve these main research goals, we conducted studies with monolingual English speakers, L1 English – L2 German speakers in the U.S., and L1 German – L2 English speakers in Germany. In all studies, participants completed phase one, the training phase, in English, during which they heard English words with voiceless fricatives with an ambiguous place of articulation [?fs]. Study 1 involved monolingual English speakers who, in phase two, judged the ambiguous voiceless fricative contrast, [?fs], on which they were trained, as well as

[2] The voiced fricatives /v/ and /z/ can differ in their realization compared to their English voiced counterparts, but the differences are minor compared to other cross-linguistic comparisons (e.g. the realization of "voice" in English vs. Spanish or French).

two untrained phoneme contrasts: voiced fricatives with ambiguous place of articulation, [?vz], as well as voiceless stops with ambiguous place of articulation, [?pt]. This study tested whether perceptual learning applies to phonological features, such as LABIAL vs. CORONAL place of articulation, or whether the difference in how these place contrasts are realized acoustic-phonetically in different manner classes, fricatives versus stops, affects generalization patterns in perceptual learning. Specifically, does perceptual learning generalize from voiceless fricatives with ambiguous place cues to voiced *fricatives* with ambiguous place cues, and does it also generalize to voiceless *stops* with ambiguous place cues? All three phoneme contrasts share the same abstract, phonological place difference (LABIAL vs. CORONAL). Yet, due to the articulatory nature of fricatives compared to stops, the acoustic cues for place in fricatives resides primarily in the fricative spectrum and secondarily in the vowel transitions (Hughes & Halle 1956; Stevens 1960; Heinz & Stevens 1961; Harris 1958), while the acoustic cues for place in stops resides primarily in the vowel transitions and secondarily in the spectral characteristics of the burst noise – and possibly the aspiration portion – of stops (Stevens & Blumstein 1978; Halle, Hughes & Radley 1957).

Studies 2 and 3 involved beginner-level L1 English – L2 German speakers in the U.S. Study 4 involved intermediate-to-advanced L1 German – L2 English speakers in Germany. In phase two of each of these bilingual studies, participants also judged the voiceless fricative contrast in English with ambiguous place of articulation, [?fs], on which they were trained, as well as three untrained phoneme contrasts from English and German: English voiced fricatives with ambiguous place of articulation, [?vz] – except in study 2 (L1 English – L2 German study 1) – as well as German voiceless fricatives with ambiguous place of articulation, [?fs], and German voiced fricatives with ambiguous place of articulation, [?vz]. These studies tested whether perceptual learning applies to fricative contrasts with ambiguous place of articulation in both languages. Studies 2 and 3 tested this cross-linguistic perceptual learning effect in L1 English – L2 German listeners after training on the non-canonical fricative sounds in the listeners' native language (English). Study 4 tested for this cross-linguistic perceptual learning effect in L1 German – L2 English listeners after training on the non-canonical fricative sounds in the listeners' non-native language (English).

Chapter 2 sets out the perceptual learning study in monolingual English listeners in the U.S. and includes important details on the study design, the choice and characteristics of the stimuli for the studies, details about how the critical stimuli were mixed, how the mixed stimuli for the continuum were constructed and chosen, and details about the procedure of the study. Most of these aspects are identical between the studies reported in this dissertation, and are not repeated in each individual chapter. Instead, later chapters refer to the stimuli and procedure details in Chapter 2. Chapter 3 also tests native English speakers, but in this case L1 English speakers who have beginner-level knowledge of L2 German. This chapter discusses two studies that were conducted with L1 English – L2 German speakers, and the insights into cross-linguistic generalization of perceptual learning effects that can be gleaned from these studies. Chapter 4 reports a study that was conducted with L1 German – L2 English speakers in Germany and the insights that can be gained from this research for our understanding of how native and non-native speakers differ in terms of adjusting the perceptual boundaries between phonemes in their first and second language. Finally, Chapter 5 concludes with a summary of the four perceptual learning studies; it describes how this research fits into and expands the literature on perceptual learning in monolingual and bilingual listeners, discusses implications for models of bilingual grammar, and provides directions for possible future research.

All four studies are based on training with a non-canonical voiceless fricative contrast (/f-s/), which exists in both English and German and is acoustically-phonetically similar between the two languages. The study design intentionally employed a shared phonological contrast (with similar acoustic-phonetic cues) to increase the likelihood of potential cross-linguistic effects. The studies with bilinguals set out to determine whether training with a non-canonical speech sound in one language, either the listeners' native or non-native language, affects both languages with which the listener is familiar. The predictions for the various outcomes of our research questions depend on whether the grammars of the two languages are assumed to be shared, separate, or separate and interconnected, a classic and long-standing question in the field of bilingualism and second-language research (as will be discussed in sections 3.6 and 5.1). In summary, this dissertation uses the lens of perceptual learning to understand a core perceptual process, adaptive phoneme categorization, and through the workings of that process, to further our understanding of how multiple languages are represented in the minds of listeners.

2. Perceptual Learning in English

This chapter reports on a perceptual learning study that was conducted entirely in English with monolingual English speakers. All participants were trained on unusual pronunciations of the English f/s contrast, and were tested for category retuning on an English /f-s/ continuum, an English /v-z/ continuum, and an English /p-t/ continuum. The objective of this study was to determine whether phonetic adjustments of a fricative contrast with an ambiguous place of articulation ([?fs]) can generalize across the phonological feature *voice* to the voiced fricative contrast with the same place features ([?vz]), and across the feature *manner* to the voiceless stop contrast with the same place features ([?pt]). The results show that perceptual learning carried across the feature *voice* to voiced fricatives, but did not carry across the feature *manner* to stops. We argue that perceptual learning effects generalize across phonological features only if the relevant contrast – here, place of articulation – is signaled by similar acoustic-phonetic cues in both types of consonant contrasts (here, fricatives and stops).

This study tests whether perceptual learning manipulates an abstract phonological feature and can generalize, within the same language, to different sound contrasts that are distinguished by the same phonological feature. This study is designed to test whether there are limitations on the ability of perceptual learning effects to generalize, and whether there are limitations on the kind of phonological features, across which perceptual learning can generalize. It is hypothesized that perceptual learning effects on a phoneme contrast on place of articulation can generalize to other phoneme contrasts differing in the same place features; furthermore, it is hypothesized that perceptual learning effects can only generalize to those phoneme contrasts in which the shared phonological feature is realized with similar acoustic-phonetic properties.

Perceptual learning effects have been shown to generalize from trained to untrained phonemes that belong to the same manner class (Kraljic & Samuel 2005, 2006). In Kraljic & Samuel (2005, 2006) native listeners of American English were exposed to atypical realizations of voiced or voiceless alveolar stops, /t/ or /d/. Not only did listeners show a perceptual learning effect for the alveolar stop contrast on which they were trained, but also for a stop contrast at an untrained place of articulation in English, namely /p/ and /b/.

Kraljic & Samuel (2005, 2006) argue that training on ambiguous alveolar stops leads to perceptual learning effects that target the abstract phonological feature voice, which is shared across the tested stop contrasts. These findings raise the following questions: Can perceptual learning effects generalize across other phonological features besides the feature voice in stop contrasts in English? If perceptual learning affects a phonological feature in the context of a particular class of sounds because it is about the mapping of phonetic data onto categories, then just as it held for phonological feature voice (on stops) across different places, it should hold for place features (on fricatives) across different voicing values. If this is not just something special about stop voicing, this would predict that perceptual learning effects generated on the voiceless fricatives /f/ or /s/ would generalize to other fricatives with the same place features.

Under the hypothesis that perceptual learning affects phonological features (Kraljic & Samuel 2006), the details of the acoustic properties of the /f/-/s/[3] contrast and the /v/-/z/ contrast are not predicted to influence whether perceptual learning effects generalize. Other research, however, has argued for the role of similar acoustics in generalization of perceptual learning effects across voices (Kraljic & Samuel 2005, 2007). Voiced fricatives differ from voiceless fricatives on a number of phonetic properties: The fricative noise is weaker, the amplitude of thee mean and normalized fricative noise is higher, and the mean and normalized frication duration is smaller in voiced fricatives compared to their voiceless fricative counterparts (e.g. Stevens 1971; Delattre, Liberman & Cooper 1964; Fant & Gunnar 1960; Shadle 2012; Jongman et al. 2000). While voiced fricatives have some unique acoustic characteristics compared to voiceless fricatives, the /v/-/z/[4] contrast, like the /f/-/s/ contrast, can be established based on the mean and normalized noise amplitude (Jongman et al. 2000). Moreover, voiced and voiceless fricatives share frequency characteristics reflecting place of articulation. For instance, voiced and voiceless fricatives have same or similar mean spectral peak locations (Jongman et al. 2000). Overall, it is predicted that perceptual learning effects will generalize between voiced and voiceless fricatives, as the place contrast is based on relatively similar phonetic properties.

This study tests whether perceptual learning targets phonemic representations of sounds and sound contrasts, such as abstract contrastive features irrespective of their phonetic realizations. It is hypothesized that the details of the phonetic-acoustic information that cue

[3] The contrast between the phonetic categories for /f/ and /s/ can be based on several phonetic cues, such as the spectrum of fricative noise, including spectral peak location and spectral moments (e.g. Tomiak 1990:168; Jongman, Wayland & Wong 2000), noise duration (e.g. Behrens & Blumstein 1988; Jongman et al. 2000) transition information, including locus information and F2 onset values (e.g. Sussman, McCaffrey & Matthews 1991; Jongman et al. 2000), noise amplitude and dynamic amplitude (e.g. Jesus & Shadle 2002; Jongman et al. 2000). Most importantly, the /f/-/s/ contrast can be uniquely established based on normalized frication duration and normalized noise amplitude (Jongman et al. 2000): The normalized frication duration (ratio of fricative duration over word duration) for English /f/ vs. /s/ is 0.420 and 0.438, respectively; the other anterior fricative in the English inventory, /θ/, has a lower normalized fricative duration than both /f/ and /s/, namely 0.415 ms. Similarly, the normalized noise amplitude (noise amplitude minus vowel amplitude) for English /f/ and /s/ is -20.8 dB and -11.0 dB, respectively; /θ/, however, has a lower normalized noise amplitude than both /f/ and /s/, namely -21.9 dB.

[4] The /v/-/z/ contrast can be uniquely established based on normalized noise amplitude (Jongman et al. 2000). The normalized noise amplitude for /v/ vs. /z/ is -13.1 dB and -9.0 dB, respectively; the other voiced anterior fricative in the English inventory, /ð/, has a lower normalized noise amplitude than both /v/ and /z/, namely -14.0 dB.

specific phonemes and phoneme contrasts play a role in perceptual learning effects. While /f/ and /p/ share the phonological place feature [LABIAL], /f/ is labiodental and /p/ is bilabial. Furthermore, while /s/ and /t/ share the phonological place feature [CORONAL] and are both [+anterior], /s/ is [+strident] and /t/ is [-strident]. While place features in both fricatives and stops are cued by formant transitions, place features in fricatives are additionally cued by spectral cues during the consonant closure (e.g. Hughes & Halle 1956; Stevens 1960; Heinz & Stevens 1961; Harris 1958), whereas place features in stops are additionally cued by spectral characteristics of the burst (e.g. Stevens & Blumstein 1978; Halle, Hughes & Radley 1957).

The hypothesis that perceptual learning effects on the voiceless fricatives (/f/ and /s/) will generalize to fricatives with the same difference in place of articulation but opposite voicing features (/v/ and /z/) predicts that after training on /f/ or /s/ in English, English speakers will show shifts in the category boundary for /v/ or /z/ in English as well.

To summarize, it is predicted here that perceptual learning based on exposure to atypical pronunciations of /f/ or /s/ can tune the way the perceptual system interprets cues to place features in fricatives, leading to possible extensions of category shifts to the voiced counterparts of the labial and coronal fricatives, namely /v/ and /z/. The hypothesis that perceptual learning targets purely abstract phonological features, such as place features (Kraljic & Samuel 2005, 2006), further predicts generalization of perceptual learning effects on place features across manner classes. However, the difference in how place contrasts are cued in fricatives and stops leads to the hypothesis that perceptual learning on place features in fricatives will not generalize to place features in stops. This study uses phonological place features to test whether perceptual learning effects based on training with voiceless fricatives differing in place will generalize to perceptual learning effects on voiceless stops differing in the same place features, /p/-/t/.

2.1 Methodology

2.1.1 Participants

Twenty-two English speakers participated in this study.[5] The participants were recruited from Stony Brook University, were at least 18 years old, provided written consent before participating, and were either paid for their participation or received course credit. No participant reported any current hearing impairments, language or learning disabilities. All participants classified English as their native language.

The majority classified themselves as monolingual native English speakers (whose parents were also native speakers of English). Eight of the 22 English speakers grew up with exposure to other languages besides English[6] ("bilinguals"), due to their parents' native languages(s): two participants had only minimal exposure to the other language when growing up and are best classified as heritage speakers of Korean and Spanish, respectively (the Korean heritage speaker reports using Korean more now than as a child; the Spanish heritage speaker reports not using Spanish at all now.) Six of the 22 participants classified themselves as English-dominant bilinguals or balanced bilinguals who grew up with at least one language besides English. The bilinguals reported having grown up with English and either Cantonese (three participants), or Spanish, Arabic, Malayalam, and Marathi (combined with perceptive skills in Hindi) (one participant each). All reported having been exposed to English from birth, except two (one of the Cantonese speakers, and the Arabic speaker), who started English around 3-4 years old in school but consider themselves as English-dominant bilinguals. All bilinguals reported using both of their languages on a regular basis (one Cantonese speaker reported having mostly receptive skills in Cantonese at the time of the study).

---

[5] The results of one participant are not included in the presented analysis, because this participant accepted less than half of the manipulated critical stimuli; see section 2.2.1 for details.

[6] The participant whose results were excluded from analysis reported being a balanced Cantonese-English bilingual, who started English in Kindergarten and still uses both languages on a regular basis.

2.1.2 Design and materials

The study was designed following Norris, McQueen & Cutler (2003) with an initial lexical decision task followed by a phoneme categorization task. During the auditory lexical decision task, participants heard 200 tokens, including 20 English words with one ambiguous fricative each, an unusual pronunciation of /f/ or an unusual pronunciation of /s/. The study had two conditions and participants participated in only one of two conditions. The only difference between the two conditions was whether the participants heard ambiguous fricatives within words with an underlying /s/ phoneme, such as *legacy*, or within words with an underlying /f/ phoneme, such as *microphone*. Specifically, in condition A ([?s]), participants heard 20 words – such as *legacy* or *rehearsal* – with ambiguous pronunciations of the phoneme /s/, and 20 words – such as *microphone* or *qualify* – with natural, unmodified pronunciations of the phoneme /f/. In condition B ([?f]), participants heard 20 words with ambiguous pronunciations of the phoneme /f/, and 20 word with natural, unmodified pronunciations of the phoneme /s/. The remainder of the stimuli, 60 filler words and 100 filler non-words, were identical for both conditions.

Following the lexical decision task, participants completed a phoneme categorization task. The participants in both conditions categorized the same ambiguous consonants on a continuum. Each participant performed category identification for the same seven-step continua, a /f/-/s/ fricative continuum, a /v/-/z/ fricative continuum, and a /p/-/t/ stop continuum. (Details about the materials and the procedure can be found in the following sections.)

2.1.3 Phase 1: Exposure to ambiguous stimuli (Lexical decision task)

The initial auditory lexical decision task exposed participants to 200 tokens, each one time. (See Appendix 1 and Appendix 2 for a list of all filler words and filler non-words.) One hundred of these tokens were made-up English sounding non-words, and 100 were (existing) English words. The 100 English words also included the 40 critical stimuli, half of which had exactly one /f/ phoneme (the /f/-stimuli), and half of which had exactly one /s/ phoneme (/s/-stimuli). Participants heard the same 200 non-words and lexical items, except that either all 20 critical stimuli containing /f/ or all 20 critical stimuli containing /s/ were manipulated.

In one condition, the [?s] training condition, the participants heard the 20 /s/-stimuli with manipulated fricative sounds: listeners heard 20 words like *legacy* with an underlying medial alveolar fricative /s/ ([lέgəs̠i]), but the fricative was manipulated to contain properties of a labial fricative [f]. Participants in this condition were also exposed to the 20 unmodified /f/-stimuli, such as *microphone*.

In the other condition, the [?f] training condition, participants heard the 20 /f/-stimuli with manipulated fricative sounds: listeners heard 20 words like *microphone* with an underlying medial labial fricative /f/ ([máɪkɹəf̠oʊn]), but the fricative was manipulated to contain properties of an alveolar fricative [s]. Participants in this condition were also exposed to the 20 unmodified /s/-stimuli, such as *legacy*.

The 20 critical /f/- stimuli and the 20 critical /s/-stimuli were selected so that each /f/-stimulus contained exactly one /f/ phoneme, and each /s/ stimulus contained exactly one /s/ phoneme (cf. Table 2.1). The critical stimuli were chosen so that the crucial fricative phoneme occurred word-medially, in a syllable-initial position followed by a vowel, and preceded by a vowel or a sonorant. Moreover, the critical stimuli were chosen so that they did not contain any other instances of /f/ or /s/ phonemes, nor any instances of /v/, /z/, /p/ or /t/. The critical /f/- and /s/- stimuli were further selected to show relatively matching mean syllable length (3.65 for /f/-stimuli, 3.25 for /s/-stimuli) and frequency (40.07 for /f/-stimuli, 41.74 for /s/-stimuli; frequency ratings based on SFI, the standard frequency index reported in Zeno, Ivens, Millard & Duvvury (1995).

Table 2.1: Critical /s/-stimuli and /f/-stimuli

	/s/-stimuli	syllables	frequency	/f/-stimuli	syllables	frequency
1	accuracy	4	50.9	amphibian	4	38.8
2	aerosol	3	13.2	beneficial	4	47.3
3	Arkansas	3	47.1	cacophony	4	25.1
4	chromosome	3	47.2	calligraphy	4	29
5	coliseum	4	32.4	chlorophyll	3	45.6
6	condensation	4	44.6	clarification	5	41.7
7	condescend	3	29.3	daffodil	3	36.3
8	connoisseur	3	32.4	endorphin	3	31.3
9	cul-de-sac	3	32.3	gleeful	2	36.4
10	delicacy	4	42.6	glorification	5	27.4
11	democracy	4	53.1	manufacture	4	51.7
12	dinosaur	3	49.9	meaningful	3	49.9
13	embassy	3	41.9	microphone	3	46.8
14	eraser	3	42.9	modification	5	45.2
15	indecision	4	39	Newfoundland	3	46.6
16	Johnson	2	54	orthography	4	22.1
17	legacy	3	40.7	perform	2	56.2
18	medicine	3	57.1	qualification	5	39.8
19	reconcile	3	40.9	qualify	3	46.3
20	rehearsal	3	43.2	unofficial	4	37.9
	Average:	***3.25***	**41.74**	***Average:***	***3.65***	**40.07**

About half of the stimuli in each group also had the critical fricative – /f/ and /s/, respectively – in the onset of a syllable with primary or secondary stress (8/20 /f/-stimuli, 11/20 /s/-stimuli). The context for the crucial fricatives was chosen to facilitate relative ease and clarity in the production and perception of these sounds. The crucial fricatives never occurred in word-initial position, and usually appeared in the last or next-to last syllable of the stimuli (20/20 /s/-stimuli; 16/20 /f/-stimuli – /f/ appears in the ante-penultimate syllable in four /f/-stimuli, three of which are five syllables long), so that the word context preceding the critical fricative would help narrow down possible lexical representations/competitors in word recognition (Kraljic & Samuel 2006).

The filler words were selected to show relatively matching mean syllable length (3.60) and token frequency compared with the critical stimuli (41.65[7]). The filler non-words were composed of English phonemes, adhered to English phonotactics, and showed relatively matching mean syllable length (3.23) compared with the critical stimuli and other filler words. None of the filler words or filler non-words contained the phonemes /f/, /s/, /v/, /z/, /p/, or /t/. The filler words and filler non-words are listed in Appendix 1 and Appendix 2, respectively.

2.1.4 Phase 2: Phoneme categorization task and stimulus construction

In the second phase of the study, each participant completed a phoneme categorization task for three English phoneme contrast continua, English /f/-/s/, /v/-/z/, and /p/-/t/. Each of the three continua consisted of seven individual disyllables of the form CVCV (/ɹiCi/) with a medial consonant that was ambiguous between the two respective endpoints, for example /ɹifi/ to /ɹisi/. All seven-step continua ranged from a relatively [LABIAL] consonant to a relatively [CORONAL] consonant; the endpoints were also modified fricatives (rather than *un*modified instances of [f] or [s]). Each of the seven steps of each continuum was presented ten times in semi-random order to each participant (see 2.1.5 below for details). In effect, each participant was exposed to 70 stimuli per continuum.

A male, 31-year old native speaker of American English was recorded for all of the stimuli used in Phase 1 and Phase 2 of this study (and the other stimuli used in the second phases used in similar studies described later in this dissertation). The speaker was also fluent in (Standard High) German and used German on a regular basis. The speaker grew up partially in Louisiana, partially in Minnesota, and had lived in New York for about eight years at the time of the recording.

This speaker said all the critical stimuli, filler words and filler non-words for the lexical decision task, as well as the disyllabic non-words for the phoneme categorization task English-only study (as well as those for the English-German studies described in subsequent chapters). For the critical stimuli, the speaker produced each of the 40 critical words in a natural way with the correct /f/ or /s/ phoneme, respectively. In addition, the speaker also produced a second

[7] See Appendix 1 and Appendix 2 for a list of all filler words and filler non-words, respectively, as well as each of their syllable lengths and – in the case of the filler *words* – each of their frequency levels.

version for each critical stimulus in which the crucial /f/ fricative had been replaced with an /s/ phoneme, and vice versa. In effect, for critical /s/-stimuli like *legacy*, the speaker produced both one version containing [s], [lɛgəs̠i], and one version containing [f], [lɛgəf̠i]. For critical /f/-stimuli like *microphone*, the speaker produced both one version containing [f], [maɪkɹəf̠oʊn], and one version containing [s], [maɪkɹəs̠oʊn].

The stimuli were recorded (on a Marantz Portable Solid State Recorder and an X microphone with a bit depth of 16 bit and a sampling rate of 44.1 kHz) in a sound-treated room. The critical stimuli with the ambiguous fricatives were created by segmenting and mixing the [f] and [s] fricatives obtained from the two types of recording for each stimulus word. For example, for the /f/-stimulus *microphone*, the [f] segment from [maɪkɹəf̠oʊn], and the [s] segment from [maɪkɹəs̠oʊn] were segmented using Praat (Boersma & Weenink 2011). For most stimuli (13/20 /f/-stimuli, 20/20 /s/-stimuli), the [f] and [s] segments did not include vowel transition information; in the other cases, the segments included one to two cycles of vowel transition information. To make the /f/-stimuli more natural sounding, the intensity of the segmented [f] sounds for nine of the stimuli was increased (multiplied by the multiplication factor 1.5) before the [f] and [s] segment were mixed for each stimulus. For one /f/-stimuli and one /s/-stimuli, the intensity of the [s] sound needed to be increased (multiplied by 1.5), instead to achieve desired ambiguous but natural sounding stimuli.

For each stimulus word, these two [f] and [s] segments were then mixed in Praat by means of interpolation, using a (modified) script by Mitterer (Mitterer 2011). "The script uses PSOLA to equate duration and pitch contour, and then interpolates between the manipulated sounds" (*Mitterer, in the introduction to the script used for the construction of the stimuli*). The script was used to create 21 [f/s] fricative mixtures for each of the 40 stimulus words. The mixtures were then spliced into the original recordings with the labial fricative pronunciation ("[f]-frame") for both the /f/-stimuli and the /s/-stimuli, e.g. [maɪkɹəf̠oʊn] for /f/-stimuli like *microphone*, and [lɛgəf̠i] for /s/-stimuli like *legacy*. The mixture was inserted in place of the original labial fricative and short (one to two cycle) vowel transitions.[8]

The author and one to two native English speakers judged, for each stimulus item, which of the 21 mixed fricatives spliced into the f-frame sounded most ambiguous while also still

[8] Four of the modified /f/-stimuli without transition information in the mixed fricative portion or the /f/-frame were additionally modified by lengthening the surrounding vowel (or sonorant) by replicating one waveform cycle, because the vowel (or sonorant) sounded short. This step was not necessary for any of the /s/-stimuli.

sounding natural rather than machine-generated. Mixtures were numbered as mixtures #0-20, going from relatively [f]-like to relatively [s]-like. For the /f/-stimuli, mixes in the range of #14-20 were chosen. For the /s/-stimuli, mixes in the range from #6-20 were chosen. For each stimulus items, the mixture that met the criteria of sounding most ambiguous yet natural was chosen. The individual mixes are listed in Table 2.2 below.

Table 2.2: Mixtures chosen for critical stimuli

/s/-stimuli	**mix#**	**/f/-stimuli**	**mix#**
accuracy	16	amphibian	18
aerosol	13	beneficial	18
Arkansas	17	cacophony	18
chromosome	6	calligraphy	18
coliseum	7	chlorophyll	15
condensation	17	clarification	18
condescend	15	daffodil	18
connoisseur	10	endorphin	19
cul-de-sac	10	gleeful	18
delicacy	12	glorification	18
democracy	14	manufacture	19
dinosaur	13	meaningful	18
embassy	13	microphone	14
eraser	18	modification	18
indecision	19	Newfoundland	20
Johnson	12	orthography	18
legacy	9	perform	16
medicine	19	qualification	17
reconcile	20	qualify	20
rehearsal	8	unofficial	18
Average:	*13.4*	*Average:*	*17.8*

In a preliminary study containing 17 of the above /f/-stimuli, participants often did not accept the modified /f/-stimuli as words. The mean acceptance rate was below 50% for the 11 participants in the [?f] training condition: 9.73/20. This result demonstrated that the critical /f/-stimuli did not sound [f]-like enough to listeners. In order to make the /f/-stimuli sound more [f]-like, the intensity of the mixed fricatives was lowered for 17 of the /f/-stimuli; the other three /f/-stimuli were replaced with new words. In the same preliminary study, participants in the [?s] training condition accepted the modified /s/-stimuli to a very high degree (19.20/20). This result demonstrated that the /s/-stimuli appeared to be natural sounding to the listeners and were therefore not adjusted.

The fricative continua for the phoneme categorization task (Phase 2 of the study) were constructed in a similar manner. The phoneme categorization task made use of two disyllabic non-words of the type CVCV, specifically /ɹifi/ and /ɹisi/. The [f] and [s] fricatives were segmented from the disyllabic non-words and mixed with the Mitterer script (Mitterer 2011), yielding [ɹi?i] non-words with ambiguous fricatives. For each continuum, the script created 21 fricative mixtures, which were spliced one at a time into the [f]-frame of the disyllabic non-word, replacing the [f] segment. The same basic procedure was applied to the disyllabic recordings for the /v/-/z/ continuum: [v] and [z] were segmented from the non-words (including brief transition periods of approximately 0.03s from the previous and following vowel), these segments were then mixed with a script, and the resulting mixtures were spliced into the [v]-frame of the disyllabic non-word, in place of the [v] segment and its vowel transitions. For the /p/-/t/ continuum, [p] and [t] were segmented from the disyllabic non-word endpoints: this included the closure duration, the aspiration, and a brief vowel transition period of the preceding vowel (approx. 0.02s). The [t] segment (which included only the amount of aspiration duration so that [t] was approximately the same length as [p]) was scaled down in intensity (by a factor of 0.7). The intensity-enhanced [p]-aspiration and the [t]-aspiration were then also mixed with the same script, and the resulting mixtures were inserted into the [p]-frame of the disyllabic non-word, in the place of the [p] segment and its vowel transition.

The same seven mixture steps between the two endpoints served as the seven steps of the continuum for all three continua used during the phoneme categorization task, i.e. for the English /f/-/s/, /v/-/z/, and /p/-/t/ continua. This means that the same number of mixture steps (three steps) was skipped between the chosen steps in all continua. In other words, the chosen mixtures were always three steps apart. The diagram in (1) below illustrates which of the 21 mixtures were chosen to form an evenly spaced seven-step continuum. (Note that the numerous mixtures generated by the script are referred to as “mix#”, whereas the seven steps chosen for the continua are referred to as “step#” henceforth.)

(1) The seven mixtures chosen as steps (inserted into disyllables) for the continua

mix#:	mix1 …	mix4 …	mix7 …	mix10 …	mix13 …	mix16 …	mix19
step#:	step1	step2	step3	step4	step5	step6	step7

more [f]-like ← ---------------------------------- ------------------------------ → *more [s]-like*

Besides these seven steps, none of the other mixtures were presented to the listeners during the phoneme categorization task. These mixtures did of course not stand on their own, but occurred with the /f/-frame, /v/-frame and /p/-frame of the disyllabic non-words (/ɹi?i/), respectively. Details about the presentation and randomization of the stimuli are provided in 2.1.5 below.

2.1.5 Procedure

Participants completed the study in a sound-treated room and were randomly assigned to one of the two training conditions. They were not told that the study contained ambiguous or modified sounds. The experiment was programmed and displayed through SuperLab software (Cedrus, Phoenix, Arizona) on a MacBook Pro laptop. The participants listened to the stimuli at a comfortable (low) volume through Sennheiser HD 555 headphones. The experimenter left the room before the actual study began, so that the participants would not attempt to speak with the experimenter during the study. This measure was taken because evidence suggests that perceptual learning effects might disappear once listeners hear unmodified speech input, although such resetting does not seem to occur very often, and only if the unmodified tokens of the relevant fricatives are produced by the same voice (Kraljic & Samuel 2005:168). The participants were encouraged to ask any clarification questions during and after the practice part and were told that the experimenter could not answer any questions once the study had begun. Participants were told that they could take short breaks throughout the experiment, whenever it was indicated on the computer screen. During the Lexical Decision Task, participants were given three options to take a break (after 50, 100, and 150 stimuli, respectively). During the Phoneme Categorization Task, participants were given the option to take a few moments before moving on to the next section (i.e. the next phoneme contrast continuum), but could not take breaks while judging the stimuli of each continuum.

Participants first completed a practice portion with feedback for both parts of the two-phase experimental study. In the practice for phase 1 (Lexical decision task), the participants were exposed to ten existing English words and ten English non-words in the accented English of a Hindi-English bilingual whose /k/ sounds had been modified so that the phonemes sounded ambiguous between [k] and [g] in five of the ten English words (*volcano, implication, publication, percussion, recover*). The practice part was devised to include accented English to encourage participants to adopt a "tolerant" and accommodating mode when listening to unusual or different speech pronunciations. This procedure was one strategy adapted to overcome the low acceptance rate of ambiguous (/f/-) stimuli in the preliminary study referred to above. The participants were told that they would hear a different voice during the actual study.

The practice part for phase 2 (category identification) consisted of two disyllabic non-words (/ɹimi/, /ɹini/) repeated three times each, or until participants selected the right answer; participants had to push one of two relevant keys to indicate whether the middle consonant they heard was /m/ or /n/. Note that participants were not exposed to a continuum and were not exposed to ambiguous sounds in the non-word syllables. This practice part was included to familiarize participants with the process of monitoring CVCV non-words for specific sounds. This step was necessary, as numerous participants in the preliminary study did not respond to several of the initial continuum steps in the phoneme categorization task. The Table 2.3 below presents an overall summary of the experimental design for the English-only study following the practice parts, including stimuli details about both phases of the study.

Table 2.3: Summary of experimental design for English-only study

<table>
<tr><th>Phase</th><th>Language</th><th>Stimuli</th><th>Cond. A ([?s])</th><th>Cond. B ([?f])</th><th>Order</th></tr>
<tr><td rowspan="4">Phase 1: Lexical Decision Task</td><td rowspan="4">English</td><td rowspan="2">critical stimuli</td><td>20 [?s] words</td><td>20 [?f] words</td><td rowspan="4">randomized (except first three filler words)</td></tr>
<tr><td>20 [f] words</td><td>20 [s] words</td></tr>
<tr><td rowspan="2">fillers</td><td>60 filler words</td><td>60 filler words</td></tr>
<tr><td>100 non-words</td><td>100 non-words</td></tr>
<tr><td rowspan="3">Phase 2: Phoneme Categorization</td><td rowspan="3">English</td><td rowspan="3">7-step continua</td><td>/v-z/ continuum</td><td>/v-z/ continuum</td><td rowspan="2">counter-balanced</td></tr>
<tr><td>/p-t/ continuum</td><td>/p-t/ continuum</td></tr>
<tr><td>/f-s/ continuum</td><td>/f-s/ continuum</td><td>always last</td></tr>
</table>

In phase 1 of the study, participants completed an auditory lexical decision task with the 40 critical stimuli (20 ambiguous, 20 unmodified stimuli), 60 filler words, and 100 filler non-words (cf. Table 2.3 above). The stimuli were randomly selected by the stimuli presentation software (SuperLab); the first three items, however, were always three English words (*barbecue, legendary, reliable*) to ensure that all participants would initially hear unmodified speech of the speaker, and to avoid losing a response to one of the critical stimuli if participants are hesitant or unsure about responding at the start of the study. The listeners heard one English word or English-like non-word at a time and had to decide within 4 seconds whether they heard an existing English word or a made-up English word. The English-like non-words were defined as made-up words consisting of sounds (phonemes) and sound combinations that occur in English; they were told that these words sound English-like but were not words in English. Participants were told that the non-words were not simply words taken from other languages, and that loanwords or English words with roots in other languages would be considered "English words" for the purpose of this study. Participants were instructed to press the key marked "Y" (for *Yes, this was an English word*) on the keyboard when they heard an existing English word, and "N" (for *No, this was not an English word*) when they heard an English-like non-word. The participants did not receive feedback about their answer choice, but received a message if they did not respond in time, or if they pressed a button other than the two answer options.

In phase 2 of the study, participants completed an auditory phoneme categorization task in three parts. In each part, participants listened to disyllabic non-words containing an ambiguous consonant word-medially (/ɹi?i/) and were instructed to indicate which sound the non-word contained in medial position by pressing one of the two answer options. Participants had 2.5 seconds to select an answer choice for each non-word they heard. The answer options were labeled "1" and "2" on the keyboard. The self-paced instructions on the computer for each of the three phoneme categorization parts explained for which sound the numbers "1" and "2" stood in each case, by making reference to the first sound of an English word. For example, for the /f/-/s/ continuum, participants read in the directions that the key with a green label and the number "1" represents the sound [f]:

(2)
The GREEN "1" stands for the first sound in the ENGLISH word "FRANK".
The GREEN "2" stands for the first sound in the ENGLISH word "SAND".

For the English /v/-/z/ continuum and the /p/-/t/ continuum, the reference words in the instructions were *veal* and *zeal,* and *Perry* and *Terry*, respectively. The three continua were presented in separate blocks, with separate introductions (including the reference words). While participants heard the stimuli items, the computer screen presented a static message as in (2) above, reminding the listeners about the sounds for which the numbered answer keys stood in each case.

For each continuum, the seven steps were presented to the participants in random order; once all seven steps had been presented once, the program presented the seven steps of the continuum again in randomized order. There were a total of ten repetitions for each 7-step continuum, for a total of seventy tokens heard per continuum. The /f/-/s/ continuum was always presented last to participants in order to avoid any (design-induced) carry-over effects from the trained /f/-/s/ contrast to the untrained contrasts (Kraljic & Samuel 2006:265). The order of the other two continua – the English /v/-/z/ continuum and the English /p/-/t/ continuum – was counterbalanced.

2.2 Results

2.2.1 Lexical Decision Task

The participants in the study performed well overall in the lexical decision task with a mean overall accuracy rate of 93.95%, ranging from 89% to 96.5%. Table 2.4 below shows that the mean accuracy for all words (critical stimuli and filler words) was 95% (SD 3%); individual scores ranged from 89% to 100%. Table 2.4 below further shows that the mean accuracy for all non-words (filler non-words) was 92.9% (SD 4.1%); individual scores ranged from 81% to 98%. Participants were also faster to correctly accept words (286.3 ms, SD 79.9) than to correctly reject non-words (393.5 ms, SD 90.9).

Note that one participant only accepted seven of the twenty critical stimuli and was replaced. This participant's data is excluded from all results reported and analyzed here, because this participant did not accept at least 50% of the critical stimuli (although the participant had a passable score of 88.0% correct overall on the lexical decision task).

Table 2.4: English-only study: Results for Lexical Decision Task (*all* stimuli)

	words	**non-words**
% correct* (SD)	95% (3.0%)	92.9% (4.1%)
RT (in ms)** (SD)	286.3 (79.9)	393.5 (90.9)

** Mean accuracy (i.e. "yes" responses to critical items)*

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Table 2.5 on page 24 illustrates the mean overall accuracy for the unmodified and modified critical stimuli. Overall, the response time (RT) and acceptance rate data suggest that the critical stimuli sounded acceptable to the listeners, although participants responded about 117 ms slower to ambiguous [?f] items, and accepted slightly fewer of the ambiguous [?f] compared to natural [f] items. Participants' responses to ambiguous [?s] items did not slow down, and participants accepted slightly more of the ambiguous [?s] items compared to natural [s] items.

(Note, though, that each participant heard either only ambiguous [?f] and natural [s] words OR ambiguous [?s] words and natural [f] words.)

Table 2.5: English-only study: Results for Lexical Decision Task (*critical* stimuli)

	natural stimuli		**ambiguous stimuli**	
	[f]	**[s]**	**[?f]**	**[?s]**
% correct* (SD)	93.6% (4.5%)	95.0% (4.7%)	86.5% (8.8%)	98.2% (2.5%)
RT (in ms)** (SD)	242.9 (93.6)	252.1 (102.4)	359.4 (61.2)	251.6 (89.9)

* *Mean accuracy (i.e. "yes" responses to critical items)*

** *Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Separate RM ANOVAs were carried out on the lexical decision results, one on the acceptance rates of the critical stimuli, and one on the RT data, of the critical stimuli. The analyses were performed separately using participants (F1) or the critical stimuli (F2) items as repeated measures. Additional factors in the ANOVAs were 1) training condition, i.e. whether participants heard ambiguous [?f] and unambiguous [s] or ambiguous [?s] and unambiguous [f] pronunciations, a between-participants but within-items factor, and 2) stimulus type, i.e. whether the critical stimuli contain an /f/ or an /s/ phoneme, a between-item but within-participant factor.

Participants accepted slightly fewer of the ambiguous [?f] items as words (86.5%) than of the natural [f] items (93.6%); Participants accepted slightly more ambiguous [?s] items as words (98.2%) than natural [s] items (95.0%). The acceptance rates of the critical stimuli reveal no significant interaction between condition and stimulus type in the by-participant analysis ($F1(1,19)=1.682$, $p=.210$), although there is a significant interaction in the by-item analysis ($F2(1,38)=5.633$, $p=.023$). The acceptance rates show a main effect of stimulus type and a main effect of condition in the by-participant analysis ($F1(1,19)=18.306, p<.001$; $F1(1,19)=7.482, p=.013$, respectively) but not in the by-item analysis ($F2(1,38)=3.641, p=.064$; $F2(1,38)=.827$, $p=.369$).

The RT data show an interaction between condition and stimulus type (F1(1, 19)=9.33, p=.007; F2(1,38)=15.958, p<.001), indicating that participants responded slower to ambiguous [?f] words than to natural [f] words (359.4 ms vs. 242.9 ms, respectively) while they responded equally fast to ambiguous [?s] words and natural [s] words (251.6 ms and 252.1 ms, respectively). The RT data also show main effect for stimulus type in the by-participant analysis (F1(1,19)= .6747, p=.018), but not in the by-item analysis (F2(1,38)=2.393; p=.130). The RT data show a main effect of condition in the by-item analysis (F2(1,38)=16.309; p<.001, but no effect in the by-participant analysis (F(1,19)=3.030, p=.098).

2.2.2 Categorical identification after exposure phase

The analysis is based on the average percent labial (%f, %v, %p) responses calculated for each continuum for each participant. A RM ANOVA was performed on percent labial (%f, %v, %p) responses to the continua. The analysis included one within-participant repeated measures factor – "phoneme contrast" – with three levels: f/s contrast, v/z contrast, p/t contrast (all in English). The analysis also included the between-participant factor of condition (training condition A with ambiguous [?f] words, or training condition B with ambiguous [?s] words).

There was a significant interaction between condition and phoneme contrast (F(2, 38)=4.348, p=.020*), a main effect of condition (F(1,19)=9.771, p=.006*) and a main effect of contrast (F(2, 38)=19.657, p<.001*). The interaction indicates that the effect of condition – the differences in percent labial responses between the two conditions – was not the same size for all three levels of the factor contrast. This interaction is shown in Figure 2.1: The figure indicates that the training effect of condition in the two fricative contrasts (E_vz and E_fs, on the left and in the middle of the figure) is larger than any training effect of condition in the stop contrast (E_pt, on the right hand-side of the figure).

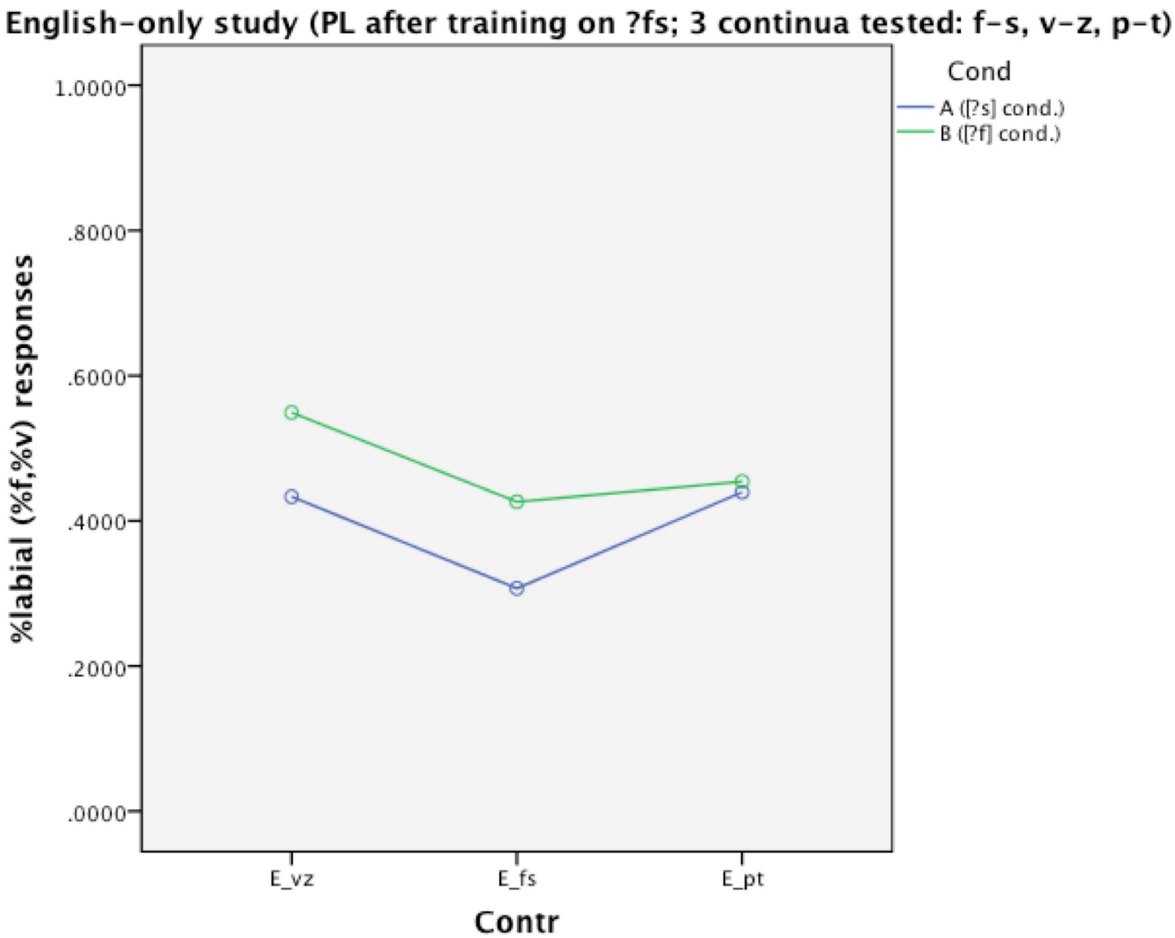

Figure 2.1: English-only study: Condition*Phoneme Contrast interaction

Further analyses address the specific hypothesis and predictions of this study, namely whether perceptual learning effects took place on the trained English f/s contrast, whether it generalized to the English v/z contrast as predicted, and whether it would show evidence for generalization of perceptual learning to English stop contrasts.

As predicted, participants showed a perceptual learning effect on the English f/s contrast: Participants in the ambiguous [?f] condition categorized more items of the f-s continuum as the labial consonant "f" (42.6%) than participants in the ambiguous [?s] condition (30.7%), a significant simple effect ($F(1,19)=16.066$, $p=.001$*) obtained through a planned pairwise comparison. The difference in percent "f" responses between the two conditions amounts to 11.9%. Figure 2.2 illustrates perceptual learning on the English [f]-[s] continuum.

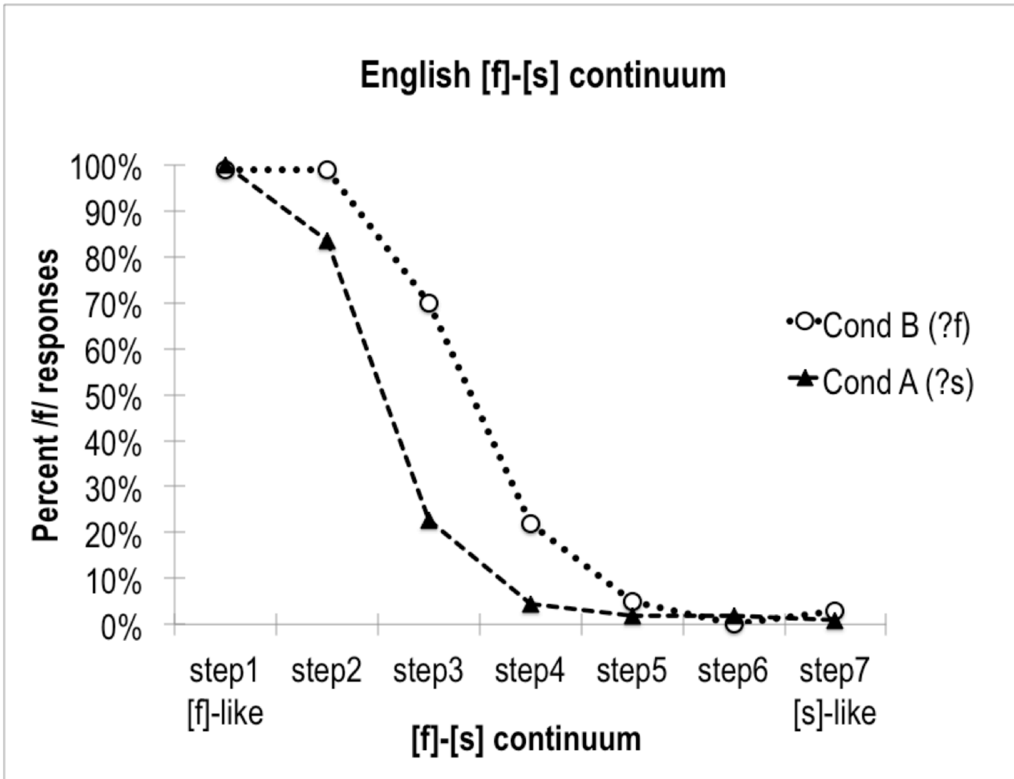

Figure 2.2: English-only study: Perceptual Learning on the [f]-[s] continuum

As predicted, participants also showed a perceptual learning effect on the English /v/-/z/ contrast: Participants in the ambiguous [?f] condition categorized more items of the [v]-[z] continuum as the voiced labial consonant "v" (54.9%) than participants in the ambiguous [?s] condition (43.3%), a significant simple effect ($F(1,19)=7.260$, $p=.014^*$) obtained through a planned pairwise comparison. The difference in percent "v" responses between the two conditions is 11.6%. Figure 2.3 below illustrates perceptual learning on the English [v]-[z] continuum.

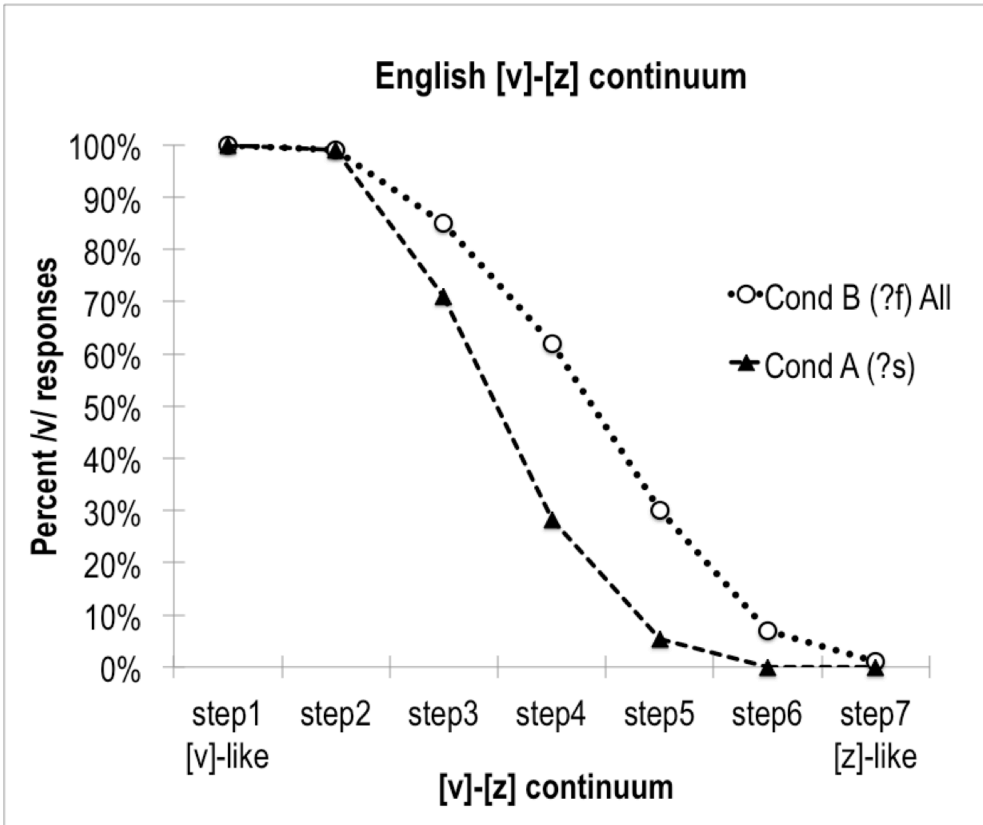

Figure 2.3: English-only study: Perceptual Learning on the English [v]-[z] continuum

As further expected, participants did not show a perceptual learning effect on the English /p/-/t/ contrast: Participants in the ambiguous [?f] condition categorized only slightly more items of the /p/-/t/ continuum as the labial stop "p" (45.4%) than participants in the ambiguous [?s] condition (44.0%), not a significant simple effect ($F(1,19)=.214$, $p=.649$), tested for through a planned pairwise comparison. The difference in percent 'p' responses between the two conditions amounts to 1.5%. The non-significant perceptual learning effect for the English [p]-[t] is illustrated in Figure 2.4 below.

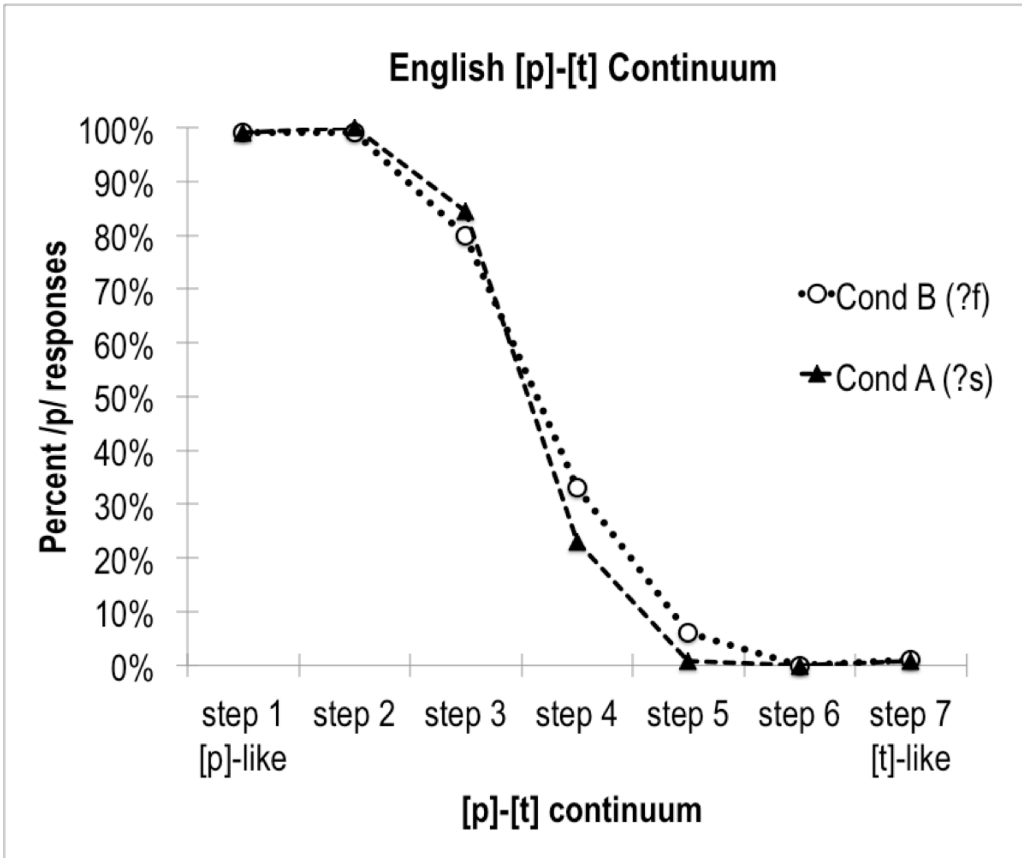

Figure 2.4: English-only study: No Perceptual Learning on the English [p]-[t] continuum

Planned comparisons/contrasts further showed that the perceptual learning effect on the English /f/-/s/ contrast (11.9%) is not significantly larger than the perceptual learning effect on the English /v/-/z/ contrast (11.6%) ($F(1,19)=.005$, $p=.944$). (This suggests that perceptual learning generalized from the voiceless fricative contrast /f/-/s/ to the voiced fricative contrast /v/-/z/). However, the perceptual learning effect on the English /f/-/s/ contrast (11.9%) is significantly different from the results on the English /p/-/t/ contrast (1.5%) ($F(1,19)=7.465$, $p=.013$*). [This suggests that the perceptual learning effect did not generalize from a labial-coronal contrast in fricatives (/f/-/s/) to a labial-stop contrast in stops (/p/-/t/).] The results of these planned comparisons show that the statistically significant interaction between condition and phoneme contrast is due to a smaller effect size on the [p]-[t] continuum.

2.3 Discussion

This study set out to test the prediction that perceptual learning based on training on the /f/-/s/ fricative place contrast tunes the way the perceptual system interprets cues to place features in fricatives in English, including the /v/-/z/ contrast. Further generalization of perceptual learning effects on place features across manner classes (such as from fricatives to stops) is predicted by the hypothesis that perceptual learning targets purely abstract phonological features, such as *place* or *voice* features (Kraljic & Samuel 2005, 2006). We hypothesized, however, that perceptual learning effects do not generalize to another contrast that differs in the same phonological feature, if this feature has different phonetic realizations and different perceptual cues, such as the /p/-/t/ contrast (e.g. fricatives: Hughes & Halle 1956; Stevens 1960; Heinz & Stevens 1961; Harris 1958; stops: Stevens & Blumstein 1978; Halle et al. 1957).

As expected, listeners who were exposed to unusual pronunciations of /f/ or /s/ in English words showed a statistically significant perceptual learning effect for the English /f/-/s/ phoneme contrast. Exposure to unusual pronunciations of /f/ or /s/ did not only lead to a perceptual learning effect on the /f/-/s/ continuum, but also to a statistically significant perceptual learning effect on the /v/-/z/ continuum. As expected, perceptual learning effects from fricatives did not generalize to stops. This limited generalization of perceptual learning is argued to be due to the different phonetic cues with which place contrasts are realized in stops and fricatives. While formant transition cues are relevant for the place contrast in both fricatives and stops, place features in fricatives are also cued by spectral cues during the consonant closure for both voiced and voiceless fricative contrasts (e.g. Hughes & Halle 1956; Stevens 1960; Heinz & Stevens 1961; Harris 1958), while place features in stops are cued by the spectral characteristics of the bursts (e.g. Stevens & Blumstein 1978; Halle et al. 1957).

The findings in this study also provide evidence that perceptual learning can generalize across various phonological features. Perceptual learning can generalize not only to different places in the English voiced-voiceless stop contrast (Kraljic & Samuel 2006), but it can also generalize across the phonological feature *voice* in the English alveolar – labial-dental fricative contrast. In other words, perceptual learning holds on a contrast of place features (labial vs. coronal) on fricatives across different phonological voicing values in English.

These findings add to the discussion about whether listeners' adjustments to unusual pronunciations occur at a generalized, abstract, phonological level, or at a detailed, acoustic-phonetic level of representation, or at the interface between phonetics and phonology. This study provides additional evidence that adaptations to atypical speech affect the perception of sound contrasts that share a number of characteristics with the sound contrasts that were produced in an atypical way. However, adaptations to atypical speech are not completely general or abstract: adaptations do not affect speech sound contrasts that, while also sharing a number of characteristics with the atypically pronounced sounds, differ in terms of the acoustic-phonetic realization of the relevant abstract phonological features. The results from this study suggest that perceptual learning operates at the phonetics-phonology interface, at a level where phonetic data are mapped onto categories. In other words, listening to atypical pronunciations leads to adjusted representations of phonological features and their acoustic-phonetic realizations, as relevant for the class of sounds sharing these phonological features and their acoustic-phonetic realizations: listeners adjust how place features (labial vs. coronal) in fricatives are to be interpreted, while listeners do not adjust how place features (labial vs. coronal) in stops are to be interpreted. This asymmetrical generalization pattern was hypothesized because the acoustic-phonetic realization of place features (labial vs. coronal) differs greatly between fricatives and stops.

The successful generalization of perceptual learning effects within fricative contrasts in English brings up the question of whether perceptual learning effects might generalize across languages. For example, would perceptual learning generalize from the English /f/-/s/ contrast to the German /f/-/s/ contrast, or from the English /f/-/s/ contrast to the German /v/-/z/ contrast? Are non-native listeners able to adjust as flexibly to unusual speech sounds as native speakers? Chapters 3 and 4 report on cross-language perceptual learning studies which address these and related questions.

To summarize, this study was designed to test the hypothesis that perceptual learning on place features in fricatives generalizes to other fricatives, but does not generalize from fricatives to stop contrasts because of the differences in phonetic realizations and perceptual cues for place features in fricatives and stops. This hypothesis was tested on native English speakers who showed, as predicted, shifts in the category boundary for /f/ or /s/ and /v/ or /z/, but no shifts in the category boundary for /p/ or /t/.

The findings show that there is generalization at an abstract, phonological level, as perceptual learning generalizes from the voiceless fricatives with an ambiguous place of articulation ([?fs]) to voiced fricatives with the same ambiguous place of articulation ([?vz]), but at the same time there are limits on the generalization of perceptual learning effects across phoneme contrasts. The results indicate that perceptual learning takes place at the phonetics-phonology interface, where phonetic-acoustic cues are interpreted in terms of phonological features and phonemic distinctions.

3. Perceptual Learning Studies in L1 English – L2 German speakers

This chapter reports on two perceptual learning studies that were conducted with novice L2 German speakers (L1 English). All participants were trained on unusual pronunciations of the English f/s contrast, and were tested for perceptual learning effects on an English /f-s/ continuum (Study #2 also included the English /v-z/ continuum), as well as German /f-s/ and German /v-z/ continua, to test for generalization of perceptual learning effects across languages. The first L1 English – L2 German study did not provide unambiguous evidence for cross-linguistic generalization effects from English to German, whereas the second study provided evidence for large perceptual learning effects on the two German continua. It is argued that this apparent difference in results might stem from a more bilingual language mode and stronger activation of the German phonological subsystems in the participants in the second study, because of their more regular use of L2 German and the situational context and bilingual instructions in the second study. These findings shed light on the complex nature of phonological and phonetic representation of phonemes common to two languages in second language learners; it is suggested that the representations of phonemes are not statically either separate or merged, but profoundly interrelated and *dynamic* in nature (cf. De Bot, Lowie & Verspoor 2005; Cook 1995; Kroll, Dussias, Bogulski & Valdes Kroff 2012; cf. also Li & Farkaš 2002), with stronger inter-relatedness between the representations when listeners use the non-dominant L2 more frequently (especially recently) and when they are engaged in a more bilingual language mode during the experiment task, and thus listening with a more activated L2 system. Finally, the findings suggest that perceptual learning effects lead to stronger perceptual learning effects in a novice, and more malleable L2 sound system, especially when the effects are generalized from the dominant L1 to the non-dominant and novice L2 (cf. De Bot 2004:26f, quoted in De Bot et al. 2005:48f).

3.1 Introduction

The previous study (reported in Chapter 2), along with previous research on speech perception (e.g. Eisner & McQueen 2005; Maye, Aslin & Tanenhaus 2008; Kraljic & Samuel 2005, 2006, 2007; Kraljic, Samuel & Brennan 2008) found that listeners do not ignore variation in their speech input, but systematically adapt to a speaker's idiosyncratic accent in their perception. The study in Chapter 2 shows that listeners can extrapolate perceptual learning effects from atypically sounding *voiceless* coronal versus labial fricative contrasts [?fs] to *voiced* contrasts [?vz]. In other words, perceptual learning effects can generalize to the place contrast in voiced fricatives even after training on voiceless fricatives only. To the best of our knowledge, generalization of perceptual learning effects to untrained phoneme contrasts has previously only been documented for stops (Kraljic & Samuel 2006). The study in Chapter 2 was conducted entirely in English to test for generalization of perceptual learning effects across the phonological feature voice to other fricative contrasts within the language of training (here, English). This chapter reports on two studies that were conducted to test whether perceptual learning (PL) can generalize across languages and influence category boundaries in second language learners' untrained novel second language. Both the German and the English phoneme inventory contain the voiceless and voiced labial fricatives /f/ and /v/, respectively, as well as the voiceless and voiced coronal fricatives /s/ and /z/, respectively. Two studies tested whether exposing English-German bilinguals to intermediate versions of a labial or coronal voiceless fricative in English will lead not only to a shift of this phoneme boundary in English, but also to a shift in the phoneme boundary of the labial and coronal voiceless as well as voiced fricatives in German. Thus, these studies are designed to test whether perceptual learning can not only generalize across the phonological feature voice within the language of training (as was done in the study in Chapter 2), but also whether perceptual learning can generalize to equivalent phoneme contrasts in an untrained language – here the /f/-/s/ contrast in the listeners' L2 German – and across languages *as well as* across the phonological feature voice to voiced fricatives with an ambiguous place of articulation [?vz] in the listeners' untrained language, L2 German.

Thus, the main question in this chapter is whether listeners can retune phoneme contrasts within *and* across languages, when they are trained in their native language. On the one hand, it is conceivable that perceptual learning might be a language-specific phenomenon and that effects would not materialize outside the language of training. This would be expected under the hypothesis that the phonological systems of a listener's native language and second language are completely separate. In fact, Escudero (2005, 2007, 2009) and Escudero & Boersma (2004) argue for two separate and autonomous sound representations and perception systems for L1 and L2. The basic claim in this approach is that second language acquisition begins with a copy of L1 grammar and (lexical) representations, which will then be adjusted based on input in L2 and the guiding principles of Universal Grammar.[9] This reflects a common notion that a bilingual's two language system make up two separate, "discrete coexisting language systems […] without links between them" (Cook 1992:566), which Grosjean (1989) labeled the "separatist position" (Cook 1992:566).

On the other hand, one might expect that two languages share a representation for speech sound categories that occur in both language inventories, based on theoretical grounds, such as Occam's razor (or *lex parsimoniae*). Indeed, early research on the bilingual double phonemic

[9] In response to studies that find that the results of bilingual listeners are intermediate between perception results of either monolingual group, Escudero (2009) proposes that this is not the result of shared representations, but of parallel activation of two separate and autonomous perception grammars.

boundary reported that English-French (Caramazza, Yeni-Komshian, Zurif & Carbone 1973) and Spanish-English (Williams 1977, 1979) bilingual listeners did not show different responses to a VOT continua in their two language contexts, and at the same time differed from each group of monolinguals in their categorization of the stop VOT values (i.e , in their perceptual cross-over point which signifies the phoneme boundary). While these early data suggested that bilingual listeners have one merged category – e.g., for VOT on stops in both of their languages – later research demonstrated that bilinguals do perceive VOT continua different in different language contexts such as French or Spanish vs. English, provided that the language contexts are well established. These data provided clear evidence that at least some bilinguals can achieve a double phonemic boundary in at least certain specific circumstances. Further research suggested that separate representations and perceptual results do not necessarily imply that equivalent phones in a bilinguals' two languages are independent of each other. Flege & Eefting (1987a, 1987b) found that highly proficient L2 English speakers – their native language being Spanish and Dutch – produced more 'extreme' VOT values in their native language, namely extremely short-lag VOT values that contrasted with their L2's long-lag VOT values and were shorter than the VOT values of monolingual Spanish/Dutch speakers. This indicates that representations of similar or equivalent phonemes in the two languages of a bilingual might be separate yet still influence each other at least in a dissimilatory manner (cf. Flege 1995b, 1995a:102) .

More recent models in SLA, Flege's L2 Speech Learning Model (SLM) (Flege 1992, 1995b, 2002, 2003, 2007) and Best & Tyler's (2007) Perceptual Assimilation Model for L2 speech perception (PAM-L2) similarly argue that second language learners equate phonetically similar phonemes that are common to L1 and L2. Flege claims that "bilinguals cannot fully separate their L1 and L2 phonetic subsystems" (Flege 2003:326), which could lead to assimilatory or dissimilatory effects (cf. Flege 1995b, 1995a:102). Research in second language acquisition and psycholinguistic approaches to multilingualism have also provided a large amount of evidence for shared perceptual categories or the mutual influence of languages in multilingual language listeners. Watson (1995), for example, found that French-English bilingual children differ from English monolingual children in their VOT category boundary perception. Sundara & Polka (2008) list numerous publications that found that the perceived category boundaries in the native language of second language listeners (who started learning their L2 at an early age, between 3 and 6 years of age) differed from monolingual L1 listeners (and monolingual listeners in their L2).

More generally, there is growing evidence that the two languages in a bilingual influence each other and interact with each other (e.g. Sebastián-Gallés & Kroll 2003:286), and that when bilinguals are selecting and using one language at a particular point in time, the other language is not entirely inactive (Green 1986; Grosjean & Soares 1986; Grosjean 1989, 1997; Spivey & Marian 1999; Dijksta 2005; Costa 2005; cf. Schwartz & Kroll (2006) for a review on bilingual processing in general). Sridhar & Sridhar (1980:413) already pointed in this direction: "The right approach, therefore, seems to be to avoid both the strong linguistic independence model and the merged system model in favor of an *interactionist* model of overlapping systems" (quoted in Cook (1992:570), my emphasis). In particular, when bilinguals listen to lexical items, lexical and sublexical items from both languages are activated (Nas 1983, quoted in Sebastián-Gallés & Kroll (2003:295); Schwartz & Kroll 2006:975; Jared & Kroll 2001; Ju & Luce 2004; Spivey & Marian 1999; Marian, Spivey & Hirsch 2003). The relative amount of activation of a language is influenced by various interacting factors, such as language dominance, the language user's language mode (Grosjean 1997, 1998, 2001) and, under the inhibitory control model (Green 1998), the degree of suppression of the non-target language (Kroll & Dijkstra 2002:317). However, "[a] very counterintuitive aspect of this body of research is that the activity of the unintended language is not simply a matter of proficiency. Both languages appear to be active in even highly proficient bilinguals." (Schwartz & Kroll 2006:990). These findings are in line with the well known claim in bilingualism research that a "bilingual is not two monolinguals in one person" (Grosjean 1989): bilinguals processing one of their languages will almost certainly differ from monolinguals on psycholinguistic measures in either of their languages. In addition, more recent findings support the notion that the language systems in bilinguals and second-language learners are constantly interacting and are mutually influencing each other in a dynamic system (De Bot et al. 2005; Flege 1995b, 2007; Cook 1995; Kroll et al. 2012; Li & Farkaš 2002).

We hypothesized that second language listeners (at least for an L1 English – L2 German population), would show cross-linguistic perceptual learning effects due to shared or interacting phoneme representations (and/or interrelated phoneme processing mechanisms), especially in novice L2 users. We therefore predicted that perceptual learning effects on the English /f-s/ contrast would also impact similar German fricative contrasts, because listeners have been argued to equate similar phones in L1 English and L2 German – according to Flege's (1995b, 2002, 2003, 2007) SLM and Best & Tyler's (2007) PAM-L2 – and because the language (phonological) systems in bilinguals and second language learners have been shown to be concurrently activated (Green 1986; Grosjean & Soares 1986; Grosjean 1989,1997; Spivey & Marian 1999; Dijkstra 2005; Costa 2005; cf. Schwartz & Kroll (2006) for a review on bilingual processing in general), and dynamically interrelated and interacting (De Bot et al. 2005; Kroll et al. 2012; cf. also Li & Farkaš 2002's Self-Organizing Connectionist Model of Bilingual Processing).

The studies on second-language leaners/bilinguals in the current and the next chapter (Chapter 4: Perceptual Learning in L1 German – L2 English speakers) make use of two closely related languages (English and German), and two sets of phoneme pairs that are phonetically (relatively) comparable between the two languages: /f/, /s/ and /v/, /z/. The studies were specifically designed with phoneme pairs that are very similar to each other across the two languages, in order to increase the likelihood of finding evidence for interrelated phonological representations or processing mechanisms (cf. Flege's equivalence classification (Flege 1989, 2005)). A previous study with Hindi-English bilinguals did not find evidence for cross-linguistic perceptual learning effects from English to Hindi, which might have been due, at least in part, to the large phonetic-acoustic differences between the chosen stop phonemes in English and Hindi (cf. Schuhmann 2012, for discussion).

Similarly, the studies reported here made use of mostly novice L2 learners, thereby increasing the likelihood of finding cross-linguistic effects from the native to the non-native language: Flege's "interaction hypothesis" (Flege 1999,1992; Walley & Flege 1999), which is part of Flege's SLM, claims that bilinguals who learned their L2 at a younger age are less likely to show a cross-linguistic interaction between their L1 and L2 than bilinguals who learned their L2 at an older age (quoted in Baker & Trofimovich (2005:22)). A growing body of research has confirmed that the largest cross-linguistic interactions occur in late as opposed to early bilinguals (cf. Bosch & Sebastián-Gallés 2003; Gildersleeve-Neumann & Wright 2010; Sundara, Polka & Baum 2006; Watson 2007; listed in Barlow et al. (2013:69)).

The long-standing question in Bilingualism and SLA research about whether bilinguals and second-language users have one or two language systems has morphed into a more complex question. This transition reflects the field's increased understanding about the flexible, interactive, continually adaptive language subsystems in both monolinguals and bilinguals (including second language learners), and the influence of variable language use and language background among listener-speakers (cf. Hall et al. 2006; Cook 2002:13). There might be various different ways in which the speech perception system handles phonemes that occur in multiple languages. Certain bilinguals might have a shared phoneme system with shared phoneme representations; other bilinguals might have separate phoneme systems and separate phoneme representations with no clear evidence for interaction between the phoneme representations across languages. Whether the phoneme systems are separate, interrelated, or integrated might further change over time within an individual, depending on language use, code-switching practices, language dominance, and other situational or task contexts (cf. Cook 2002:1) – including language mode (Grosjean 1989,1997,1998,2001). These kind of findings would be in line with a growing body of evidence for a usage-based view of multilingual (as well as monolingual - cf. Hall et al. 2006) listeners, who are argued not to have a fixed, stable, discrete or even homogenous system of language knowledge, but rather one that is dynamic and variable (Hall et all 2006:225; Kroll et al. 2012), for example Dynamic Systems Theory for second language acquisition (De Grot et al. 2005) and the notion of multicompetence (Cook 1992,1995). In Kroll et al. (2002)'s words about bilingual processing: "the emerging findings suggest a language system that is far more dynamic than previously understood." (Kroll et al. 2012:231).

In this chapter, inferences about the status of phoneme representations and their interactions in L1 English – L2 German second language learners are made based on whether participants show perceptual learning effects both in the language of training and another language, or in just the language of training. The analyses take into account the listeners' activation level of each language (in particular their L2 German), as evidenced by self-reported language background and language usage details, as well as the manipulation of language mode in the two studies. Additionally, each participant completed an additional L2 German perception proficiency test, which was designed for this dissertation to assess each listener's ability to discriminate non-native phoneme contrasts in a same/different task. The relevant phonemes occurred in non-words, which were slightly masked by white noise.

3.2 Methodology

3.2.1 Participants

We conducted two studies with relatively novice L1 English – L2 German listeners at two different universities (1) Stony Brook University and (2) University of California at Berkeley. The results are analyzed separately – rather than as *one* combined study – due to important differences in the set-up and in the volunteers that participated in each study. Participants in the second study completed an additional phoneme categorization task (an English /v/-/z/ continuum), in addition to the three other continua that participants in both studies completed. As will be detailed below, the second L1 English – L2 German study was more likely to achieve a more bilingual language mode in the listeners compared to the first L1 English – L2 German study. Further, the language background of the participants differed such that participants in the second study used more German on a regular basis at the time of the study, compared to participants in the first L1 English – L2 German study.

3.2.1.1 Participants English-German Study 1

In the first English-German study, 20 native speakers of English with mostly beginner-level skills in German as a second language participated in the study. The participants were recruited from Stony Brook University, provided written consent before participating, were at least 18 years old, and were paid for their participation or received course credit. No participant reported any current hearing impairments, language, or learning disabilities. All participants classified English as their native language and had studied (or were at the time studying) German as a foreign language.

All participants in this study reported having grown up in the US or in Canada (two participants) and six of the participants reported having grown up with more then one language in their childhood. All of these six participants also report having heard and learned English from birth, indicating that they should be considered simultaneous bilinguals (although not each participant continued learning and/or speaking the non-English language). All of these six bilinguals consider English to be their dominant language or to be a relatively balanced bilingual with English and (one of their) other language(s) (one to two participants). The other language(s) of these simultaneous bilinguals are: French (2), Bengali & Hindi (1), Hindi (& Gujarati) (1), Armenian (& Turkish) (1), and German (1).

3.2.1.2 Participants English-German Study 2

Similarly, in the second English-German study, 21 native speakers of English with mostly beginner-level skills in German as a second language participated in the study. The participants were recruited from the University of California at Berkeley, provided written consent before participating, were at least 18 years old, and were paid for their participation. No participant reported any current hearing impairments, language, or learning disabilities. All participants classified English as their native language and had studied (or were at the time studying) German as a foreign language.

All but three of the participants in this study reported having grown up in the US. (One person grew up and lived in the US except between the ages of 2-8 year old, when the participant lived in Germany; one participant grew up in China before moving to the US at age 3, and one person grew up and lived in Sri Lanka until 1 year before the study.) Ten of the participants reported having grown up with more then one language in their childhood. Six of the bilinguals in this study were raised by parents or caretakers whose native language(s) was/were not English, and four of these ten bilingual did not learn English quite from birth, but around the ages of 3-4 years old. All of the ten bilinguals in this study consider English to be their dominant language today. The other language(s) of these bilinguals are: Mandarin (4), Korean (2), German (2), Sinhalese (& Sindhi) (1), and Filipino (1).

3.2.2 Design, materials, and procedure

The design of the English-German studies followed the design of the English-only study, described in Chapter 2, which was modeled after Norris et al. (2003): Participants first completed an auditory lexical decision task (100 words, 100 non-words), conducted entirely in English, followed by a phoneme categorization task in both English and German. The difference between the English-only study in Chapter 2 and the English-German studies in this chapter (and the German-English study in Chapter 4) resides in the phoneme categorization tasks: in the studies with both English and German, the phoneme categorization task included continua in both English and German. The lexical decision task and the stimuli used within the lexical decision task were the same in the English-only study, the English-German studies, and the German-English study: it included the filler words and 100 filler non-words listed in Appendix 1 and Appendix 2, including the same 20 English words with ambiguous fricatives (depending on the condition, either ambiguous /f/ phonemes within /f/-stimuli, or ambiguous /s/ phonemes within /s/-stimuli). The English-German (and the German-English) studies used the same recordings and modified sound files for the English stimuli as in the English-only study from Chapter 2.

The category identification task in the English-Germany and German-English studies includes the two English fricative continua (English_fs and English_vz) from the English-only study in Chapter 2, as well as two new German fricative continua, German_fs and German_vz. The same speaker produced all of the English and the German stimuli for these studies; the stimuli were produced, recorded, and manipulated by the same principles. The speaker was a native speaker of American English, started learning German as a second language at the age of 11/12, was fluent in (Standard High) German, taught German as a foreign language at the university level and used German on a regular basis at the time of the recording. The stimuli for the phoneme identification task were chosen so that the first consonant /r/ would identify the language of these non-words. This is important, because the four phonemes in each of the four non-words are part of the language inventory of both English and German. The speaker produced the German nonwords ([rifi, risi, rivi, rizi]) with a uniquely non-American /r/, namely an alveolar trill [r], one of the possible German /r/-realizations.

As a reminder, each of the fricative continua were constructed from the recordings of the two nonwords, for example [rifi] and [risi]: The [f] and [s] fricatives were segmented from the German disyllabic non-word endpoints and mixed with the same Mitterer script as above for each continuum. The script created 21 fricative mixtures, which were spliced one at a time into the [f]-frame of the disyllabic non-word, replacing the [f] segment. The same basic procedure was applied to the disyllabic recordings for the /v/-/z/ continuum: [v] and [z] were segmented from the non-words (including brief transition periods of approximately 0.03s from the previous and following vowel), these segments were then mixed with a script, and the resulting mixtures were spliced into the [v]-frame of the disyllabic non-word at a zero-crossing, in place of the [v] segment and its vowel transitions.[10]

The general procedure also followed the English-only study: participants were randomly assigned to one of the two training conditions and were not told that the study contained ambiguous sounds. The experiment was presented through SuperLab software on a MacBook Pro laptop, and participants listened to the stimuli at a comfortable (low) volume through Sennheiser HD 555 headphones. The experiment took place in a sound-treated or quiet room. As in the English-only study, participants conducted two practice parts, and the experimenter left the room before the actual experiment began. As indicated earlier, the two English-German studies differed in phase 2 of the study. The specific design for each study is described below in the next two sections.

3.2.2.1 Design and Procedure for English-German Study 1

The English-German study #1 at Stony Brook included a three-part auditory phoneme categorization task. In each part, participants listened to disyllabic non-words containing an ambiguous consonant word-medially and were instructed to indicate which sound the non-word contained in medial position by pressing one of the two answer options. In the English-German study 1, all the instructions were provided in English. One of the categorization tasks was the English /f/-/s/ continuum that was used in the English-only study in Chapter 2. The other two categorization tasks were German continua, namely a German /f/-/s/ continuum and a German /v/-/z/ continuum.

Participants read self-paced instructions on the computer screen for each of the three phoneme categorization parts. Most importantly, these instructions explained for which sounds the colored answer options "1" and "2" on the participants' keyboard stood. For the English f/s continuum, participants were instructed that "1" and "2" referred to the first sound of the English

[10] Any clicks or non-smooth transitions that resulted from splicing the fricative (with minor mixed vowel transition portions) into the /v/-frame with the surrounding vowels (and their vowel transitions portions) were cut.

words *Frank* and *sand*, respectively. Similarly, for the German f/s continuum, participants were instructed that “1” and “2” referred to the first sound in the German words *Frank* and *Sand*, respectively. Finally, for the German v/z continuum participants were instructed that “1” and “2” referred to the first sound of the German words *wie* and *sie*, respectively.

As in the English-only study, the English /f/-/s/ continuum was always presented last to participants in order to avoid any (design-induced) carry-over effects from the trained /f/-/s/ contrast to the untrained contrasts (Kraljic & Samuel 2006:265). The order of the other two continua – the German /f/-/s/ continuum and the German /v/-/z/ continuum – were counterbalanced. Table 3.1 below presents an overall summary of the experimental design for the English-German study #1 (at Stony Brook University), including stimuli details about both phases of the study.

Table 3.1: Summary of the experimental design for the English-German study 1

Phase	**Stimuli**	**Language**	**Cond. A ([?s])**	**Cond. B ([?f])**	**Order**	*****
Phase 1: Lexical Decision Task	**critical stimuli**	English	20 [?s] words	20 [?f] words	randomized *(except first three filler words)*	*same*
			20 [f] words	20 [s] words		
	fillers		60 filler words	60 filler words		
			100 non-words	100 non-words		
Phase 2: Phoneme Categorization	**7-step continua**	German	/f-s/ continuum	/f-s/ continuum	counter-balanced	*new*
		German	/v-z/ continuum	/v-z/ continuum		*new*
		English	/f-s/ continuum	/f-s/ continuum	always last	*same*

*** Same as in E-only study?**

3.2.2.2 Design and Procedure for English-German Study 2

The English-German study at the University of California at Berkeley was a slightly modified version of the first English-German study; it included the same Lexical Decision Task

(Phase 1) as study 1, and an expanded four-part auditory phoneme categorization task (Phase 2): Two continua of the categorization tasks were the two English fricative continua that were used in the English-only study, namely the English /f/-/s/ continuum and the English /v/-/z/ continuum. The other two categorization tasks were German continua, namely a German /f/-/s/ continuum and a German /v/-/z/ continuum. In each part, participants listened to disyllabic non-words containing an ambiguous consonant word-medially and had to categorize this word-medial consonant by pressing one of the two answer options.

A heightened bilingual language mode was facilitated by a bilingual practice part and bilingual instructions: The initial phoneme categorization practice part prior to the start of the study was done with English non-words (as in study 1) and *additionally* with German non-words. This second English-German study also included additional German instructions (preceded by the English instructions) for the German continua portions. Moreover, this particular study was conducted in a more "German environment", a quiet room within the German library in (on the hallway of) the German Department, where study participants were surrounded by books, media, and posters in German and relating to the German language and culture. (Study 1 at Stony Brook, in contrast, was conducted in a quiet room within the Linguistics lab in the Linguistics Department.) These modifications were meant to help listeners activate both English and German at the outset of the study, and to facilitate switching into their German language mode when listening to German stimuli in the phoneme categorization task after having heard only English for about 5-7 minutes in the Lexical Decision part (Phase 1 of the study).

Participants read self-paced instructions on the computer screen for each of the three phoneme categorization parts. Most importantly, these instructions explained for which sounds the colored answer options "1" and "2" on the participants' keyboard stood. For the English f/s continuum, participants were instructed that "1" and "2" referred to the first sound of the English words *fun* and *sun*, respectively. Similarly, for the German f/s continuum, participants were instructed that "1" and "2" referred to the last sound in the German words *auf* and *aus*, respectively. Next, for the English v/z continuum, participants were instructed that "1" and "2" referred to the first sound of the English words *veal* and *zeal*, respectively. Finally, for the German v/z continuum participants were instructed that "1" and "2" referred to the first sound of the German words *wie* and *sie*, respectively.

As in the English-only study, the English /f/-/s/ continuum was always presented last to participants in order to avoid any (design-induced) carry-over effects from the trained /f/-/s/ contrast to the untrained contrasts (Kraljic & Samuel 2006:265). To prevent any carry-over effects from the trained language (English) to the untrained language (German), the other English continuum, the English /v/-/z/ continuum, was always presented as the next-to-last continuum, directly before the English /f/-/s/ continuum. The order of the other two continua – the German /f/-/s/ continuum and the German /v/-/z/ continuum – were counterbalanced. Table 3.2 below presents an overall summary of the experimental design for the English-German study #2 (at the University of California at Berkeley), including stimuli details about both phases of the study.

Table 3.2: Summary of the experimental design for the English-German study #2.

Phase	**Stimuli**	**Language**	**Cond. A ([?s])**	**Cond. B ([?f])**	**Order**	*
Phase 1: Lexical Decision Task	**critical stimuli**	English	20 [?s] words	20 [?f] words	randomized *(except first three filler words)*	*same*
			20 [f] words	20 [s] words		
	fillers		60 filler words	60 filler words		
			100 non-words	100 non-words		
Phase 2: Phoneme Categorization	**7-step continua**	German	/f-s/ continuum	/f-s/ continuum	counter-balanced	*same*
		German	/v-z/ continuum	/v-z/ continuum		*same*
		English	/v-z/ continuum	/v-z/ continuum	always next-to-last	*new*
		English	/f-s/ continuum	/f-s/ continuum	always last	*same*

* Same as English-German study #1?

To summarize, the English-German study 2 (in Berkeley) was a modification of study 1 (in Stony Brook); most importantly, participants in study 2 completed an additional phoneme categorization task (an English /v/-/z/ continuum, as in the English-only study), and participants were more likely to be in a bilingual language mode in study 2 compared to study 1.

3.3 Results

3.3.1 Results English-German Study 1

3.3.1.1 Lexical Decision Results (English-German Study 1)

The participants in the study performed well overall in the lexical decision task with a mean overall accuracy rate of 94.10% (SD 3.65), ranging from 84.5% to 98.5%. Table 3.3 below shows that the man accuracy for all words (critical stimuli and filler words) was 95.2% (SD 3.3%); individual scores ranged from 88.0% to 99.0%. Table 3.3 further shows that the mean accuracy for all non-words (filler non-words) was 93.0% (SD 5.9%); individual scores ranged from 79.0% to 100%. Participants were also faster to correctly accept words (251.8 ms, SD 75.1) than to correctly reject non-words (390.2 ms, SD 97.4).

Table 3.3: English-German study 1: Lexical Decision (all stimuli)

	words	**non-words**
% correct* (SD)	95.2% (3.3%)	93.0% (5.9%)
RT (in ms)** (SD)	251.8 (75.1)	390.2 (97.4)

** Mean accuracy (i.e. "yes" responses to critical items)*

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Table 3.4 below illustrates the mean overall accuracy for the unmodified and modified critical stimuli. Overall, the response time (RT) and acceptance rate data suggest that the critical stimuli sounded acceptable to the listeners, although participants accepted slightly fewer of the ambiguous critical stimuli and responded slightly slower compared to the natural critical stimuli. Participants accepted slightly fewer (7%) of the ambiguous [?f] compared to natural [f] items (86.5% compared to 93.5%, respectively) and responded about 37.4 ms slower to ambiguous [?f] items. Participants also accepted slightly fewer (2.5%) of the ambiguous [?s] items compared to natural [s] items (96.5% vs. 99%, respectively) and responded slightly slower to ambiguous [?s] items (29.4 ms slower). (Note, though, that each participant heard either only ambiguous [?f] and natural [s] words OR ambiguous [?s] words and natural [f] words, just like in the English-only study.)

Table 3.4: English-German study 1: Lexical Decision (critical stimuli)

	Natural stimuli		**Ambiguous stimuli**	
	[f]	**[s]**	**[?f]**	**[?s]**
% correct* (SD)	93.5% (4.7%)	99.0% (2.1%)	86.5% (5.8%)	96.5% (4.1%)
RT (in ms)** (SD)	207.7 (129.8)	225.5 (66.6)	245.1 (67.2)	254.9 (80.3)

** Mean accuracy (i.e. "yes" responses to critical items)*

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Separate RM ANOVAs were carried out on the lexical decision results, one on acceptance rates of the critical stimuli and one on RT data. The analyses were performed separately using participants (F1) or the critical stimuli (F2) items as repeated measures. Additional factors in the ANOVAs were 1) training condition, i.e. whether participants heard ambiguous [?f] and unambiguous [s] or ambiguous [?s] and unambiguous [f] pronunciations, a between-participants but within-items factor, and 2) stimulus type, i.e. whether the critical stimuli contain an /f/ or an /s/ phoneme, a between-item but within-participants factor.

The acceptance rates of the critical stimuli show a significant interaction for condition and stimulus type in the by-person analysis (F1(1,18)=17.190, p=.001) but not in the by-item analysis (F2(1,38)=1.132, p=.294). Participants overall accept more /s/-stimuli as words (ambiguous [?s] items: 96.5%, natural [s] items: 99.0%) than /f/-stimuli (ambiguous [?f] items: 86.5%, natural [f] items: 93.5%): Both the by-subject and the by-item analysis show a significant main effect of stimulus type (F1(1,18)=45.762,p=.000; F2(1,38)=10.038, p=.003).

The by-item analysis also shows a significant main effect of condition ($F2(1,38)=5.047, p=.031$), but the by-subject analysis does not ($F1(1,18)=1.976, p=.177$).

The RT data of the by-subject analysis show an interaction between condition and stimulus type ($F1(1, 18)=6.289$, $p=.022$) while the by-item analysis does not show a significant interaction ($F2(1,38)=.236$, $p=.630$). The RT data also show a main effect for condition in the by-item analysis ($F2(1,38)= 5.610$, $p=.023$), but not in the by-subject analysis ($F1(1,18)=.011$; $p=.917$). The RT data do not show a main effect of stimulus type ($F1(1,18)=1.078$; $p=.313$; $F2(1,38)= .037$, $p=.849$). Participants' reaction times did not differ significantly between critical items with /f/ phonemes and critical items with /s/ phonemes; participants also did not respond significantly slower to ambiguous phonemes overall, or only in the case of /f/-stimuli (245.1 ms for [?f] stimuli vs. 207.7 ms for [f] stimuli; compared to 254.9 ms for [?s] stimuli vs. 225.5 ms for [s] stimuli).

3.3.1.2 Perceptual Learning Results (English-German Study 1)

The analysis is based on the average percent labial (%f, %v) responses calculated for each continuum in each language for each participant. A RM ANOVA was performed on percent labial (%f, %v) responses to the continua. The analysis included one within-participant repeated measures factor – “phoneme contrast” – with three levels: the English f/s contrast, the German f/s contrast, and the German v/z contrast. The analysis also included the between-participant factor of condition (training condition A with ambiguous [?f] words, or training condition B with ambiguous [?s] words).

There was no significant interaction between phoneme contrast and condition ($F(2,36)=1.418$, $p=.255$), no overall significant effect of condition ($F(1,18)= 3.053$, $p=.098$), but a significant main effect of phoneme contrast ($F(2,26)=21.098$, $p<.001$). The missing interaction indicates that the effect of condition – the differences in percent labial responses between the two conditions – was not significantly different between any of the three levels of the factor contrast. The non-significant effect of condition indicates the lack of an overall training effect (a training effect across all three phoneme contrasts tested). The effect of phoneme contrast is not relevant;

it indicates that the average percent labial responses across both conditions differ between the three phoneme contrasts.

Further analyses (planned pairwise comparisons) address the specific predictions of this study, namely whether perceptual learning effects took place on the trained English f/s contrast, and if so, whether it generalized to the German f/s and v/z contrast as predicted. As predicted, participants showed a perceptual learning effect on the English f/s contrast: Participants in the ambiguous [?f] condition categorized more items of the f-s continuum as the labial consonant "f" (44.6%) than participants in the ambiguous [?s] condition (38.0%), a significant simple effect ($F(1,18)=9.056$, $p=.008$*) obtained through a planned pairwise comparison. The difference in percent "f" responses between the two conditions is 6.6%. Figure 3.1 below illustrates successful perceptual learning on the English [f]-[s] continuum in the English-German study 1.

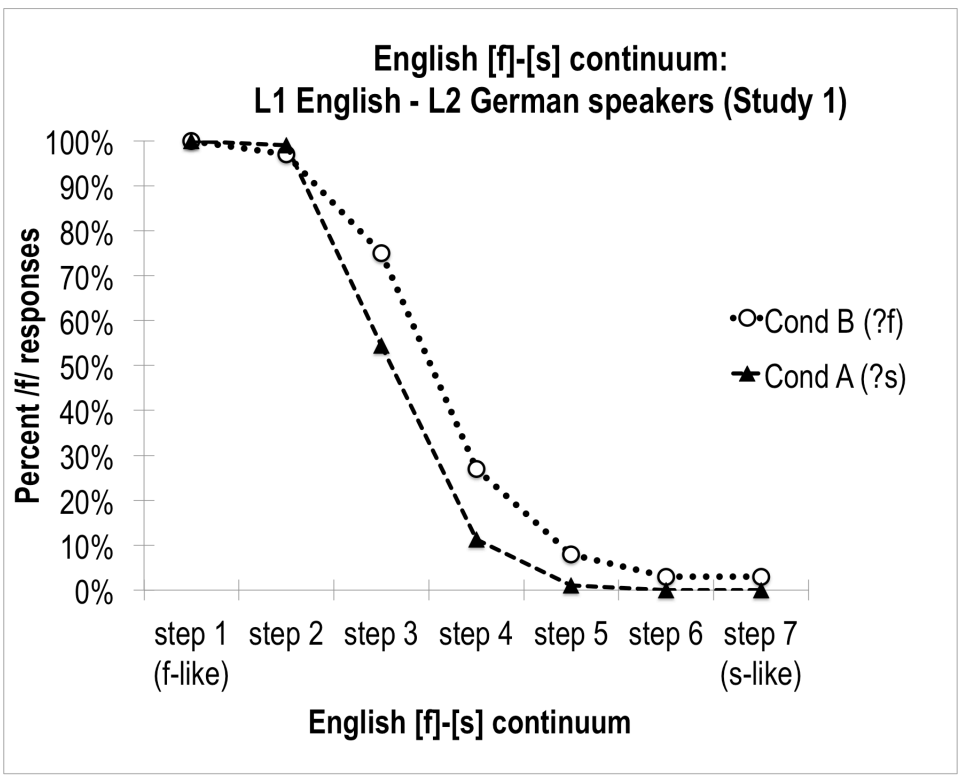

Figure 3.1: English-German study 1: Perceptual Learning on the English [f]-[s] continuum

Participants did not show a statistically significant perceptual learning effect on the German /f/-/s/ contrast: Although participants in the ambiguous [?f] condition categorized more items of the German [f]-[s] continuum as the voiceless labial consonant "f" (45.2%) than participants in the ambiguous [?s] condition (40.4%), the difference did not reach significance (F(1,18)=2.534, p=.129), a simple effect obtained through a planned pairwise comparison. The difference in percent "f" responses between the two conditions amounts to 4.8%. Figure 3.2 below illustrates the non-significant perceptual learning effect on the German [f]-[s] continuum in the English-German study 1.

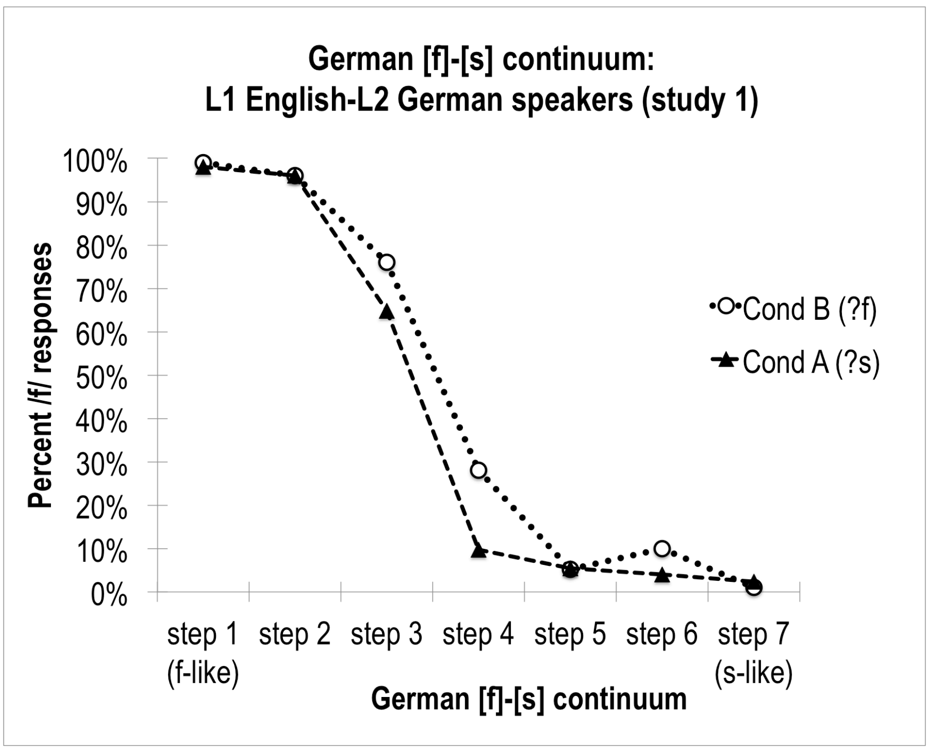

Figure 3.2: English-German study 1: Perceptual Learning on the German [f]-[s] continuum (not significant).

Participants also did not show a statistically significant perceptual learning effect on the German /v/-/z/ contrast: Participants in the ambiguous [?f] condition categorized only slightly more items of the German [v]-[z] continuum as the voiced labial consonant "v" (33.6%) than participants in the ambiguous [?s] condition (32.4%), not a significant simple effect

(F(1,18)=.103, p=.752) obtained through a planned pairwise comparison. The difference in percent "v" responses between the two conditions amounts to 1.2%. Figure 3.3 below illustrates the non-statistical perceptual learning effect on the German [v]-[z] continuum in the English-German study 1.

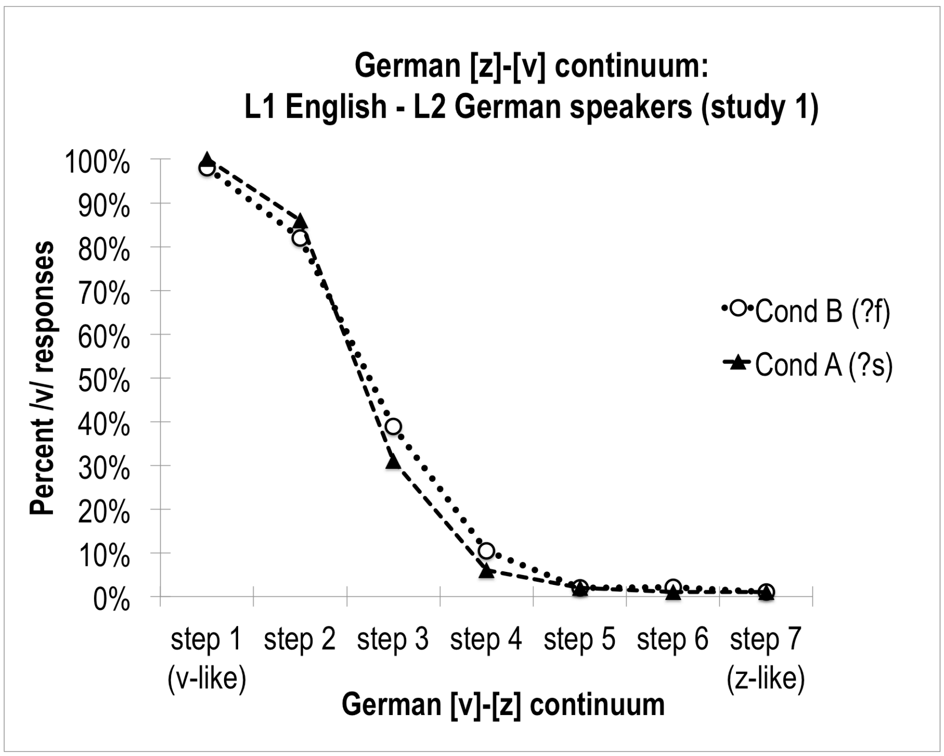

Figure 3.3: English-German study 1: Perceptual Learning on the German [v]-[z] continuum (not significant).

Planned (within-subjects) contrasts further show that the difference of percent labial responses between the two conditions on the English [f]-[s] continuum is not significantly larger than the difference between conditions on the German [f]-[s] continuum (F(1,18)=.359, p=.557) or the German [v]-[z] continuum (F(1,18)=2.671, p=.120). These non-significant planned contrasts and the non-significant interaction of phoneme contrast and condition reported above suggest that the perceptual learning effect on the English f/s contrast is *not* significantly larger than the perceptual learning effects on either the voiceless or voiced fricative contrasts in

German. These non-significant differences between the English f/s contrast and either of the two German fricative contrasts (German f/s and German v/z) hint that the perceptual learning effects on the German fricatives were smaller, albeit not absent.

To summarize, English native speakers with second language learning experience in German also show perceptual learning effects on the trained English /f/-/s/ phoneme contrast, similar to monolingual English listeners (cf. Chapter 2). However, the participants in the first L1 English –L2 German study did not generalize – at least not at a statistically significant effect – the perceptual learning effects from English /f/-/s/ phoneme contrasts to labial-alveolar fricative contrasts in German, neither to the German /f/-/s/ phoneme contrast, nor to the German /v/-/z/ phoneme contrast.

3.3.2 Results English-German Study 2

3.3.2.1 Lexical Decision Results (English-German Study 2)

The participants[11] in this study also performed well overall in the lexical decision task with a mean overall accuracy rate of 95.22% (SD 3.3%), ranging from 84.5% to 98.5%. Table 3.5 below shows that the mean accuracy for all words (critical stimuli and filler words) was 97.1% (SD 1.8%); individual scores ranged from 93.0% to 100%. Table 3.5 below further shows that the mean accuracy for all non-words (filler non-words) was 93.4% (SD 6.7%); individual scores ranged from 70.0% to 100%. Participants were also faster (on average, 175.6 ms.) to correctly accept words (300.0 ms, SD 155.8) than to correctly reject non-words (475.6 ms, SD 225.2).

[11] This analysis is based on the results of all participants except three “low acceptors”, who accepted fewer than 70% of the ambiguous critical stimuli – see below.

Table 3.5: English-German study 2 (w/o low acceptors): Lexical Decision (all stimuli)

	words	**non-words**
% correct* (SD)	97.1% (1.8%)	93.4% (6.7%)
RT (in ms)** (SD)	300.0 (155.8)	475.6 (225.2)

** Mean accuracy (i.e. "yes" responses to critical items)*

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Table 3.5 below illustrates that the mean overall accuracy for the unmodified and modified critical stimuli. Overall, the response time (RT) and acceptance rate data suggest that the critical stimuli sounded acceptable to the listeners: on average, participants responded merely 31.5 ms slower to ambiguous [?f] items compared to natural [f] items (330.6 ms. compared to 299.1 ms., respectively), and accepted only slightly fewer (5.4%) of the ambiguous [?f] compared to natural [f] items (91.4% compared to 96.8%, respectively). Participants also responded slower to ambiguous [?s] items (127.9 ms slower) but accepted slightly more (0.5%) of the ambiguous [?s] items compared to natural [s] items (99.1% vs. 98.6%, respectively). (Note, though, that each participant heard either only ambiguous [?f] and natural [s] words OR ambiguous [?s] words and natural [f] words, just like in the English-only study.)

Table 3.6: English-German study #2: Lexical Decision (critical stimuli)

	Natural stimuli		Ambiguous stimuli	
	[f]	**[s]**	**[?f]**	**[?s]**
% correct* (SD)	96.8% (3.4%)	98.6% (3.8%)	91.4% (4.8%)	99.1% (2.0%)
RT (in ms)** (SD)	299.1 (180.8)	201.7 (140.6)	330.6 (118.8)	329.6 (178.3)

** Mean accuracy (i.e. "yes" responses to critical items)*

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Separate RM ANOVAs were carried out on the lexical decision results, one on acceptance rates of the critical stimuli and one on RT data. The analyses were performed separately using participants (F1) or the critical stimuli (F2) items as repeated measures. Additional factors in the ANOVAs were 1) training condition, i.e. whether participants heard ambiguous [?f] and unambiguous [s] or ambiguous [?s] and unambiguous [f] pronunciations, a between-participants but within-items factor, and 2) stimulus type, i.e. whether the critical stimuli contain an /f/ or an /s/ phoneme, a between-item but within-participants factor.

The acceptance rates of the critical items do not show an interaction between condition and stimulus type ($F1(1,16)=3.521$, $p=.079$; $F2(1,38)=2.600$, $p=.115$). As in the first L1 English – L2 German study, participants accepted overall more /s/-words ([?s] items: 99.1%, [s] items: 98.6%) than /f/-items ([?f] items: 91.4%, [f] items: 96.8%), as indicated by a main effect of stimulus type ($F1(1,16)=13.161$, $p=.002$; $F2(1,38)= 7.700$, $p=.009$). The acceptance rate of critical stimuli also shows an effect of condition in the by-subject analysis ($F1(1,16)=8.215, p=.011$) but not in the by-item analysis ($F2(1,38)=1.766$, $p=.192$).

The RT data show a significant interaction in the by-subject analysis ($F1(1,16)=24.492$, $p<.001$) but not in the by-item analysis ($F2(1,38)=4.050, p=.051$). The RT data also show a significant effect of stimulus type ($F1(1,16)=9.317$, $p=.008$; $F2(1,38)=4.107$, $p=.050$), indicating that participants responded slower to /f/-stimuli ([f] stimuli: 299.1; [?f] stimuli: 330.6) than to /s/-stimuli overall ([?s] stimuli: 329.6, [s] stimuli: 201.7). The RT data also show an effect for condition in the by-item analysis (F2(1,38)=13.323,p=.001) but not in the by-subject analysis (F1(1,16)=.391,p=.540).

3.3.2.2 Perceptual Learning Results (English-German Study 2)

The data were analyzed after exclusion of three participants who accepted fewer than 70% of the ambiguous critical stimuli. The acceptance rate of the ambiguous critical stimuli of these three participants appears quite different from the acceptance rate of the ambiguous critical stimuli of the remaining participants: the low acceptors accepted between 50% and 65% of the ambiguous critical [?f] stimuli, while all the remaining participants accepted between 85% and 100% of the ambiguous critical [?f] or [?s] critical stimuli.[12]

A RM ANOVA on Percent Labial (%f, %v) responses to the continuum was carried out. The results are based on Percent Labial (%f, %v) responses calculated for each participant. The analysis included two within-participant repeated measures factors: language (English, German) and phoneme contrast (f/s contrast, v/z contrast). The analysis also included the between-participant factor of condition (training condition with ambiguous [?f] words or ambiguous [?s] words). (As in the literature: RT is not analyzed; F2 (by item) analysis is also not conducted on a single continuum with 7 steps/items.)

There was no significant three-way interaction between language, phoneme contrast, and condition (F(1,16)= 1.863, p=.191). There was, however, a statistically significant interaction between condition and language (F(1,16)=7.258, p=.016), indicating the difference in percent labial responses between conditions was significantly larger in German than English: the effect of condition was larger in German (49.3% vs. 31.6%, i.e. a difference of 17.7%) than in English (53.1% vs. 45.2%, i.e. a difference of 7.8%). There was also a significant main effect of condition (F(1,16)=19.449; p<.001). Keeping in mind that there was also a (dominating)

[12] Norris et al. (2003:214) and Sjerps & McQueen (2010) excluded participants with fewer then 50% acceptance of the ambiguous critical stimuli. Kraljic & Samuel (2005:150) excluded participants with fewer then 75% acceptance of *all* critical items (i.e. modified and unmodified critical items) or below 75% acceptance rate of all fillers items. Kraljic & Samuel (2006:265) exclude all participants with lower than overall 70% correct acceptance rate in the lexical decision task. Each participant in all the studies reported in this dissertation had an acceptance rate of over 70% correct in the overall Lexical Decision task (the lowest overall acceptance rate in the Lexical Decision task was achieved by a L1 German-L2 English participant in Berlin with a 70.5% acceptance rate; this study is discussed in Chapter 4).

interaction between condition and language, this main effect of condition indicates an overall effect of training condition: Participants who heard ambiguous [?f] words categorized overall more items on the English and German f-s and v-z continua as labial (%f and %v) (51.2%) than participants who heard ambiguous [?s] words (38.4%).

Finally, there was a significant main effect of language (F(1,16)=22.449, p<.001). This effect indicates that the mean percent labial responses for both contrasts and both conditions in English (49.2%) were significantly larger than in German (40.5%). (There was also a statistically significant interaction between language and phoneme contrast (F(1,16)=47.876, p<.001). This effect shows that the v/z phoneme contrast had a larger difference in percent labial responses between languages (56.7% in English vs. 33.6% in German) than the f/s phoneme contrast (41.6% in English vs. 47.3% in German). Moreover, the interaction between phoneme contrast and condition was not significant (F(1,16)=.005, *p* =.946), and neither was the main effect of phoneme contrast (F(1,16)=.141, *p* =.712).)

The most important results, namely perceptual learning effects on each of the four individual contrasts, were achieved through simple effects. As expected, participants showed a perceptual learning effect on the English f/s contrast, the language and phoneme contrast of training. Participants in the ambiguous [?f] condition categorized more items of the [f]-[s] continuum as the labial consonant “f” (47.01%) than participants in the ambiguous [?s] condition (36.16 %), a significant simple effect (F(1,16)=8.988, p=.009*) obtained through a planned pairwise comparison. The difference in percent “f” responses between the two conditions amounts to 10.85%. Figure 3.4 below illustrates successful perceptual learning on the English [f]-[s] continuum in the English-German study #2.

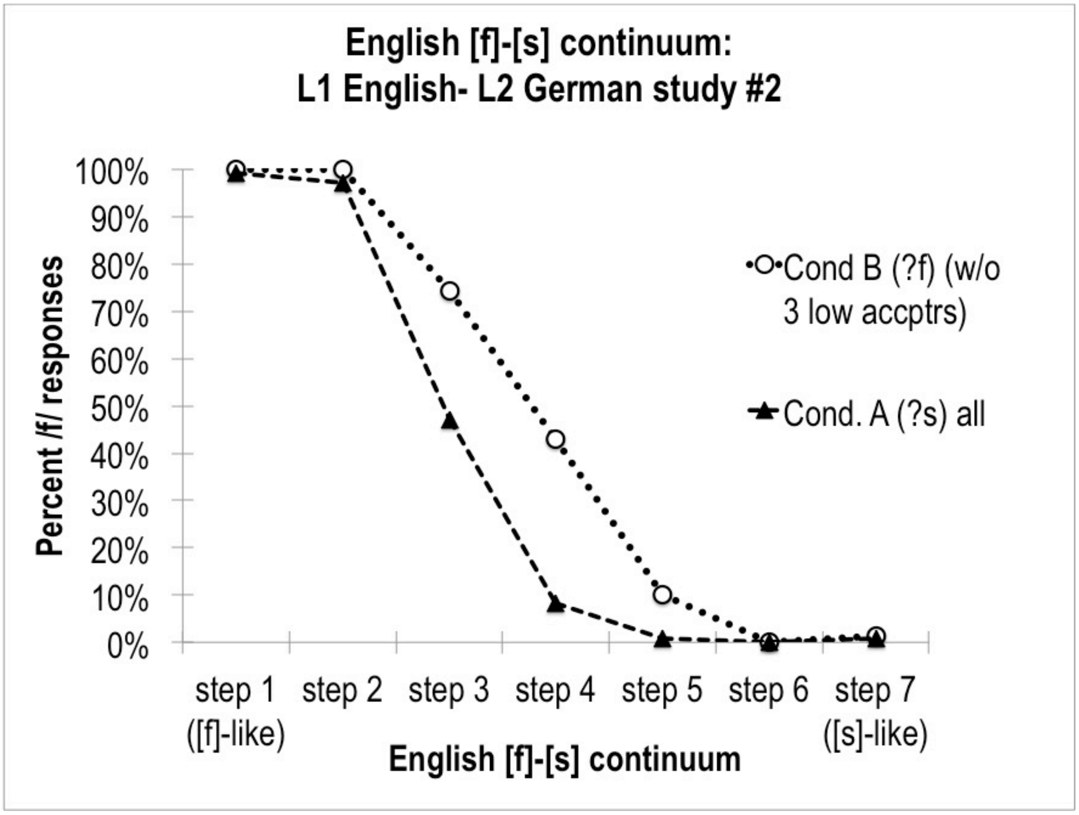

Figure 3.4: English-German study 2 (without low acceptors): Perceptual Learning on the English [f]-[s] continuum.

Participants also showed a perceptual learning effect on the German f/s contrast, the phoneme contrast of training in an untrained language: Participants in the ambiguous [?f] condition categorized more items of the f-s continuum as the labial consonant "f" (54.81%) than participants in the ambiguous [?s] condition (39.74%) a significant simple effect ($F(1,16)=8.497$, $p=.010$*) obtained through a planned pairwise comparison. The difference in percent "f" responses between the two conditions amounts to 15.07%. Figure 3.5 below illustrates successful perceptual learning on the German [f]-[s] continuum in the English-German study #2.

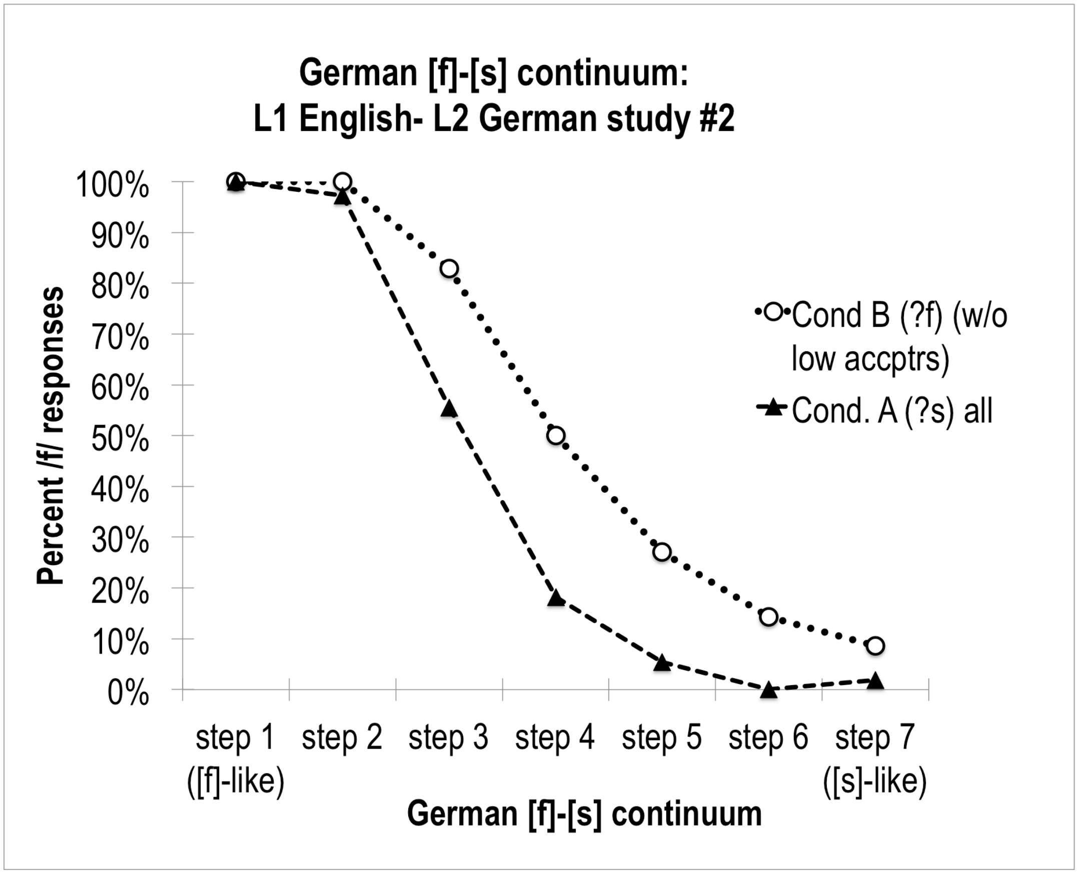

Figure 3.5: English-German study 2 (without low acceptors): Perceptual Learning on the German [f]-[s] continuum.

Moreover, participants showed a perceptual learning effect on the German v/z contrast, generalizing to an untrained phoneme contrast in an untrained language: Participants in the ambiguous [?f] condition categorized more items of the German [v]-[z] continuum as the labial consonant "v" (43.88%) than participants in the ambiguous [?s] condition (23.42%) a significant simple effect ($F(1,16)=21.369$, $p<.001$*) obtained through a planned pairwise comparison. The difference in percent "v" responses between the two conditions amounts to 20.46%. Figure 3.6 below illustrates successful perceptual learning on the German [v]-[z] continuum in the English-German study #2.

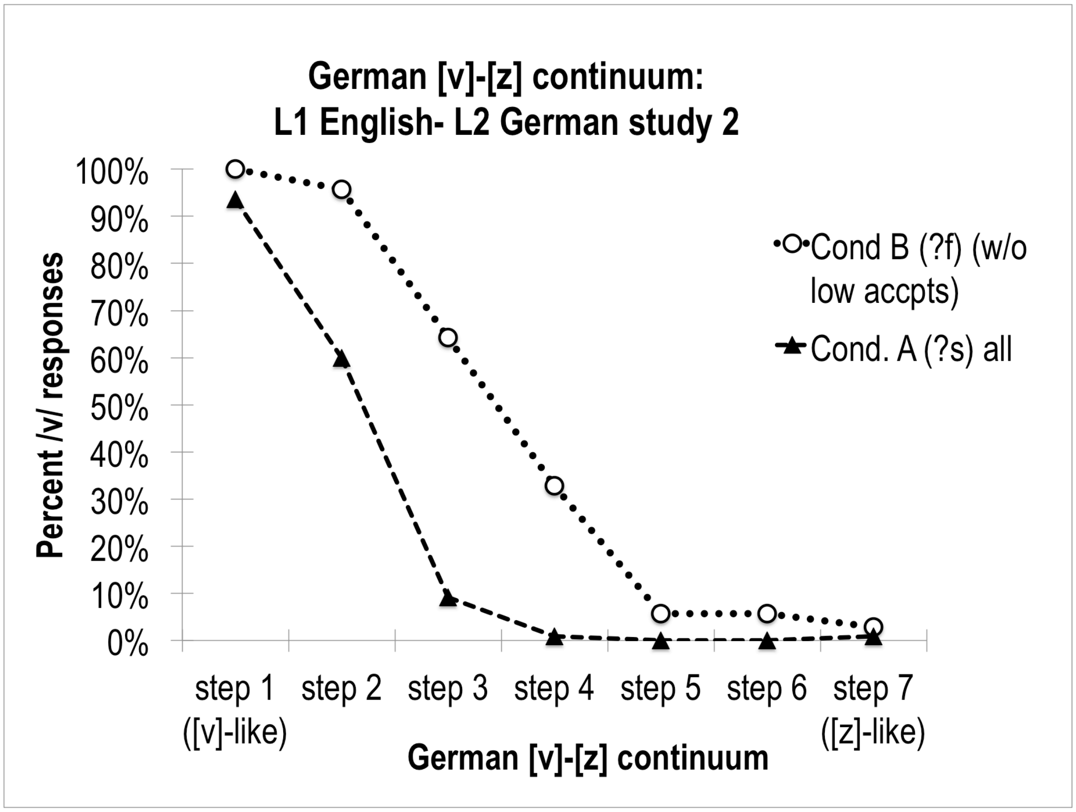

Figure 3.6: English-German study 2 (without low acceptors): Perceptual Learning on the German [v]-[z] continuum.

Rather unexpectedly in this context, participants did not show a statistically significant perceptual learning effect on the English v/z contrast, an untrained phoneme contrast in the language of training[13]: Participants in the ambiguous [?f] condition categorized only slightly more items of the English [v]-[z] continuum as the labial consonant "v" (59.18%) than participants in the ambiguous [?s] condition (54.29%), a non-significant simple effect ($F(1,16)=1.102$, $p=.309$) obtained through a planned pairwise comparison. The difference in percent "v" responses between the two conditions amounts to 4.89%. Figure 3.7 below illustrates the non-significant perceptual learning on the English [v]-[z] continuum in the English-German study #2.

[13] One possible reason for this unexpected missing effect on the English /v-z/ continuum might be that the voiced fricative /v/ has been found to be especially difficult to distinguish reliably (in particular from the voiced fricative /ð/) (Miller & Nicely 1955, quoted in Wright 2004:37f)

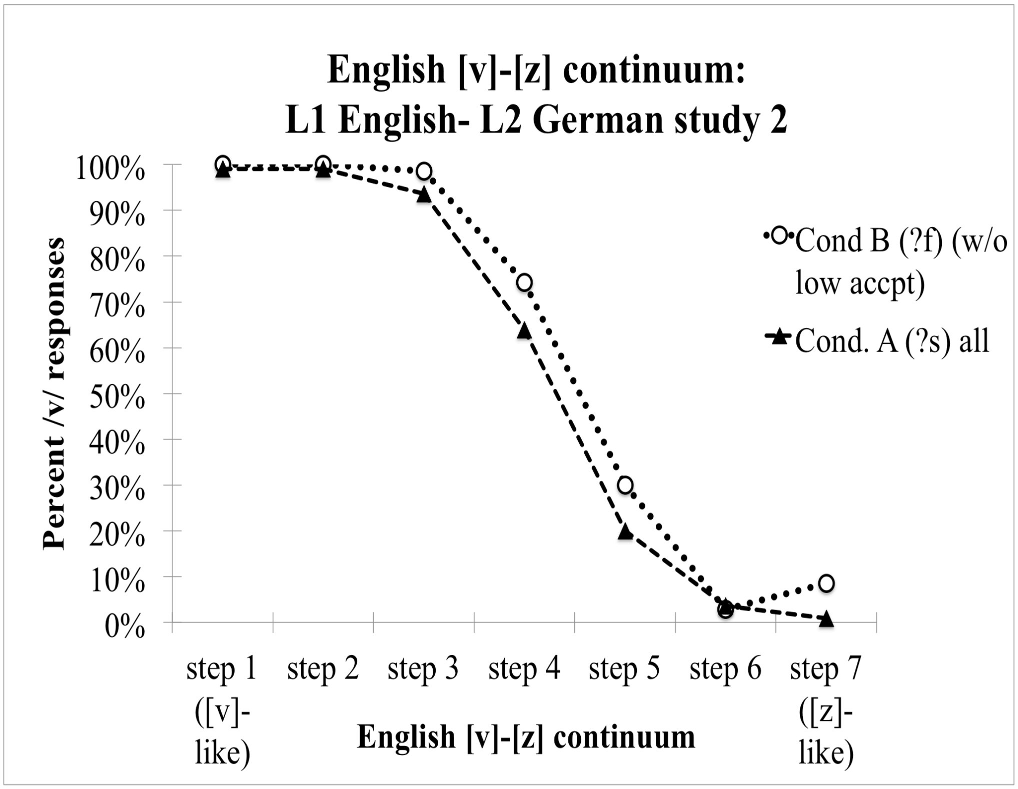

Figure 3.7: English-German study 2 (without low acceptors): Perceptual Learning on the English [v]-[z] continuum (not significant).

If all participants of the English-German study – including the three low acceptors – are included in the analysis, the results change slightly, because the German f/s contrast is only marginally significant (cf. Appendix 3.1). As indicated above, all of the participants accepted at least 50% of the ambiguous critical stimuli; the three "low acceptors", however, accepted only between 50% and 65% of the ambiguous critical stimuli.

To summarize, the results of the English-German study 2 reveal perceptual learning effects for the trained phoneme contrast in the language of training, the English f/s contrast, as well as for the untrained phoneme contrast in the untrained language of training, the German v/z contrast. The smaller perceptual learning effect on the English v/z contrast never reached statistical significance. Crucially, the German f/s contrast (trained phoneme contrast in an untrained language) reached significance in the analysis that focuses on the participants with an acceptance rate of at least 70% for the ambiguous critical [?fs] stimuli.

3.4 Language background in the English-German studies

In studies involving bilinguals, it is customary and important to discuss the language background of the participants, and – if possible and relevant – to obtain independent scores about the language proficiencies of the bilingual participants. Seemingly contradictory findings between experiments can often be explained when differences in the language background and/or language proficiency of the bilingual participants are taken into account. This section discusses information about the participants' language background, and the results of a separate L2 German language proficiency test in German conducted with each participant. Each participant in one of the English-German studies participated in a brief language perception proficiency test to gauge the participants' perception skills in their non-native language. Participants completed the perception proficiency test after the main perceptual learning study. This test was a same-different task with non-words embedded in white noise.

Each participant heard four blocks, each of which contained twelve similar-sounding non-word pairs. Half of the non-word pairs in each block did not differ phonologically, but were different recordings of the same non-words, e.g. /kyt/-/kyt/ or /kit/-/kit/ for non-native speakers of German; the other half of the non-word pairs differed in exactly one phoneme and presented a phoneme contrast that does not exist in the listener's native language; this is shown in Table 3.7 below. For example, native speakers of English taking the German language perception proficiency test heard /k**i**t/ - /k**y**t/ and judged whether the two syllables were the "same" or "different"; native speakers of German taking the English language perception proficiency test heard /á**w**a/ - /á**v**a/, among others.

Table 3.7: Same Different Task used for the L2 Language Perception Proficiency Test

Same-Different Task: Type of stimuli in each of four blocks	
Same (6)	Same – nonwords with sound A (3)
	Same – nonwords with sound B (3)
Different (6)	Different: nonwords with sound A followed by nonwords with sound B (3)
	Different: nonwords with sound B followed by nonwords with sound A (3)

All the English stimuli were recorded by a native speaker of English (the same person who also recorded the stimuli for the perceptual learning study); all German stimuli for this language were recorded by a native speaker of German (the author). For each stimuli pair that the listeners heard, they were instructed to press a button to indicate whether the two non-words they heard contained the “same” or “different” sounds. The ISI (inter-stimulus-interval) between the two stimuli in non-word pairs was set to a large value (1 second), to encourage phonological processing rather than more immediate and shallow phonetic-acoustic processing. The English non-word stimuli pairs and the German non-word stimuli pairs for the four blocks are shown in Table 3.8 below:

Table 3.8: Non-word stimuli pairs for the L2 language proficiency tests

L2 German perception proficiency test stimuli:	**L2 English perception proficiency test stimuli:**
/k**i**t/ vs. /k**y**t/	/á**w**a/ vs. /á**v**a/
/pná**ç**ə/ vs. /pná**x**ə/	/á**f**a/ vs. /á**θ**a/
/k**y**t/ vs. /k**ø**t/	/kl**ǽ**b/ vs. /kl**ɛ́**b/
/trɛ́**ç**ə/ vs. / trɛ́**ʃ**ə/	/wədǽ**s**/ vs. /wədǽ**z**/

For each person, we calculated the percent correct responses for all four contrasts.[14] In order to test whether the participants' perception proficiency score in their L2 German was correlated with how weak or strong their perceptual learning effects were, the overall percent correct response was then used as a continuous covariate in an ANCOVA analysis with the overall perceptual learning effect (pooled over all 3-4 contrasts), as well as each of the individual contrasts. In a separate analysis of the data, the results from the perception proficiency test were binned (visual binning in in SPSS) into two groups, a group of participants with higher scores and a group of participants with lower scores, for each of the two English-German studies. A second ANCOVA analysis was conducted with this categorical covariate and again the overall perceptual learning effect and each of the individual contrasts.

None of the ANCOVA analyses for the first or the second English-German study provided a significant effect. Details of the statistical data can be found in the appendix (Appendix 3.3 and 3.4). An improved L2 perception proficiency task could involve a task that relies less on an "acoustic-phonetic analysis task" as was the case in this same-different task, and more on "tasks involving lexical processes" (Diaz et al. 2012:680). Sebastian-Gallés & Baus (2005) and Diaz et al. (2012) have demonstrated that fluent early and late bilinguals achieve different results on tests designed to test phonological perception/processing; both studies found that bilinguals can score similarly to monolinguals on tests that focus on acoustic-phonetic analyses, but differ more from monolinguals on tests that involve the lexicon.

Additional language background information collected through a questionnaire (cf. Appendix 3.5) and the language perception proficiency test made it possible to test for correlations between the size of perceptual learning effects and specific aspects of participants' language background. We were able to analyze the following variables related to language background based on the questionnaires: Hours of German used per week, German perception proficiency score, proficiency in German (self-assessment score), age of acquisition of L2

[14] In the L1 English – L2 German study #1, participants heard stimuli that were randomly selected for each participant from a pool of recordings for each non-word. For the second English-German study, the non-word pairs were fixed to those stimuli pairs for which native German speakers scored the highest in a separate study conducted in Germany (Fall 2011). The analysis of the data for the first English-German study was limited to those non-word pairs that were used in the second English-German study (namely the ones that native German speakers scored on the highest); this ensured that the data points are based on the same recordings in each study; however, it reduced the number of pairs that were analyzed for the L2 German perception proficiency test for the participants in the first English-German study (the data points for the second English-German study were always 48 (12 non-word pairs for each of the four phoneme contrasts tested).

German, time in L2 environment, years of L2 German study, years of L2 German experience, and status in childhood as L1 English monolingual or bilingual (in English and another language in childhood).

The data for each of the two studies was analyzed in terms of language background information by means of ANCOVA's. For the English-German study 1, none of the ANCOVA's showed an overall influence of any of the individual language background factors onto the training condition results (as evidenced by the absence of significant interactions of the supposed covariates and the between-subject variable "condition" when analyzing all three continua at once). The individual results for the ANCOVA's conducted for the English-German study 1 can be found in Appendix 3.3.

For the English-German study 2, most of the ANCOVA's also did not show an overall influence of the individual language background factors onto the training condition results (as evidenced by the absence of significant interactions of the supposed covariates and the between-subject variable "condition" when analyzing all three continua at once). However, the overall training effect (the pooled mean perceptual learning effect) and the perceptual learning effects for the English f/s continuum, the English v/z continuum and the German f/s continuum are each larger in participants with a lower self-assessment proficiency score in L2 German. Moreover, the overall training effect for all four continua pooled (mean of all four continua), and for the English f/s continua are each larger in participants with fewer years of L2 German study. The perceptual learning effect for the German v/z continuum, however, was larger for participants with a higher score on the L2 German perception proficiency test. Further research is needed to determine why some of these language background factors were significant for specific continua in the second English-German study. The individual results for all the ANCOVA's conducted for the English-German study 2 (for all participants excluding the low acceptors), and supporting graphs and scatterplots can be found in Appendix 3.4. A summary of the language background variables tested and whether or not they correlated with the perceptual learning effect size is shown in Table 3.9 below.

Table 3.9: Summary of language background factors and whether they had an effect in the two L1 English – L2 German studies

Language Background Variable	L1 English – L2 German study #1 (SBU)	L1 English – L2 German study #2 (CAL)
Hours of German used per week		
German perception proficiency test scores		✓ (German v/z)
Proficiency in German (self-assessment score)		✓ (overall training effect; English f/s, English v/z, German f/s)
Age of acquisition of L2 German		
Time in L2 environment (L2-speaking country)		
Years of L2 German study		✓ (overall training effect, English f/s)
Years of L2 German experience		
Current age		
Status in childhood as L1 English monolingual or bilingual		

3.5 Differences between the English-German Studies

The two L1 English – L2 German studies reported in this chapter yielded different results. The first study showed significant perceptual learning effects only for the language (and contrast) of training, English f/s. The second study showed generalization of perceptual learning effects from the English accent with ambiguous [?fs] fricatives to the voiceless and voiced fricative contrasts in German, the f/s and v/z contrast, respectively. This difference in results between the two studies might be due to the difference in the language mode methodology alluded to above, as well as a specific aspect of the language background of the participants in the two studies. The English-German study 1 and English-German study 2 differed slightly in terms of the design of the study: While the English-German study 1 tested three phoneme contrasts (English f/s; German f/s and German v/z), the English-German study tested four phoneme contrasts (English f/s *and additionally English v/z*; German f/s and German v/z), a difference in the design that is unlikely to have had an impact on the difference in results between the studies. A likely more important difference in the design of the studies was that the English-German study 1 included all instructions in English (the reference word for the German continua were, of course, German words), while the English-German study 2 included instructions in English *and* in German: The listeners were provided with written instructions in English (as in study 1) *and,* following the English instructions, with additional written instructions translated into German for each of the phoneme categorization tasks involving German continua. In Experiment 2, listeners also completed a phoneme categorization practice part with English non-words (as in Experiment 1) *and* additionally with German non-words. The lack of clear evidence for cross-linguistic perceptual learning effects in the first study led to the inclusion of German instructions in the second English-German study; arguably, exposing listeners to both languages in the instructions might put them in a more bilingual language mode and possibly lead to (more) cross-linguistic perceptual learning effects, particularly in comparison to the first English-German study without bilingual instructions.

Moreover, whereas participants in the first English-German study participated in a quiet room within the Linguistics lab in the Linguistics Department at Stony Brook, the second English-German study was conducted in a quiet room within the German library of the German department[15] (at the University of California at Berkeley). On the way to the study, participants walked through the German department and part of the German library, and were surrounded by books in German while participating in the study. This situational context of the second study might have additionally created a strong German/German-esque setting, thus likely further increasing the bilingual language mode of the participants. (cf. Grosjean 1994, 2000; cf. Escudero & Boersma 2002; Escudero 2005; cf. Reis-Pereira (2009:113ff) for an overview.)

Finally, we analyzed language background factors to see whether they play a role in whether or not second language listeners show cross-linguistic generalization of perceptual learning effects. The participants' self-reported language background, which was collected through the language background questionnaire (cf. Appendix 3.5) in the two English-German studies, was rather similar. As noted above, we were able to compare participants' background

[15] Except for one participant, who participated in a quiet room in the general university library because the German library was closed.

based on the following variables: Hours of German used per week, German perception proficiency score (the same-different task described above), proficiency in German (self-assessment score), age of acquisition of L2 German, time in L2 environment, years of L2 German study, years of L2 German experience, and status in childhood as L1 English monolingual or bilingual (in English and another language in childhood). Except for hours of German usage per week, none of the tested language background questions showed a statistically significant difference between the first and the second English-German study. Most importantly, there is no statistically significant difference between the German perception proficiency scores of the first and the second English-German study, the variable that was predicted to most influence the size of the perceptual learning effects.

The comparison of the L2 German language background data for the participants in the first and the second English-German study, however, indicates that the participants in the first study typically used German for fewer hours per week than the participants in the second study (4.93 vs. 8.83 hs/wk, respectively), as shown in the histograms in Figure 3.8 and Figure 3.9 below.

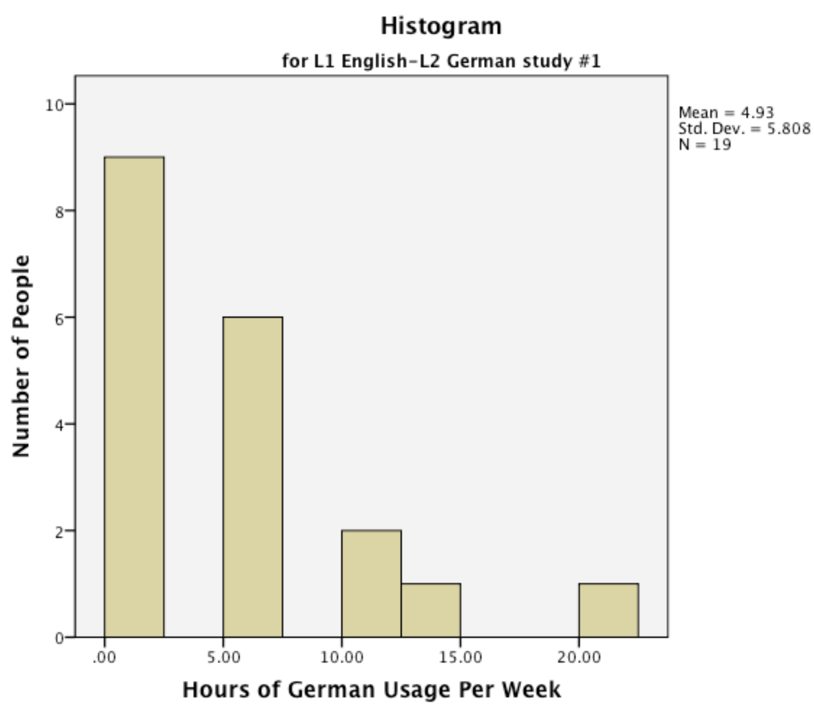

Figure 3.8: Histogram of the Hours of German per week for study #1 (Stony Brook).

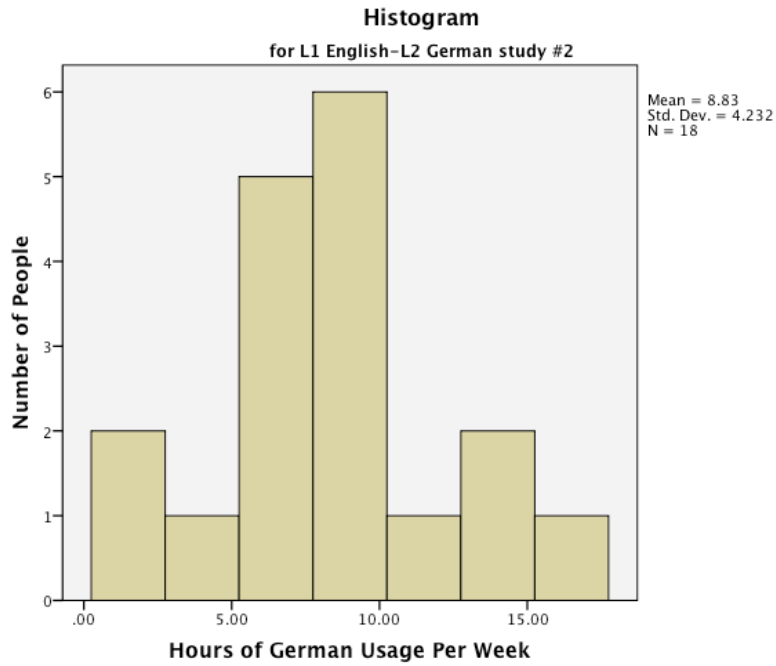

Figure 3.9: Histogram of the Hours of German per week for study # 2 (Cal)

Participants in the first English-German study (excluding 1 participant – SBU #17 – with +/- 3 s.d. from range, namely z-score of 3.91) used German for 4.93 hrs/week (range: 0.0-21.0 hrs/wk, median: 5.00 hrs/wk; s.d. 5.81; 95% CI intervals: 2.13-7.73, N=19[20-1]), while participants in the second English-German study used German for 8.83 hrs/week (range: 1.5-17.5 hrs/wk; median: 9.50 hrs/wk; s.d.: 4.23, 95% CI intervals: 6.73-10.94, N=18[21-3].) A 1-way ANOVA on the hours of German used per week by participants (excluding the one outlier) in the first study versus participants in the second English-German study shows that there is a statistically significant difference between the participants in the two groups with regard to amount of German used per week ($F(1,35)= 5.395$, p=.026*).

In conclusion, participants in the first and second L1 English – L2 German studies did not differ with regard to their L2 German perception proficiency score, their self-assessed proficiency levels in German, age of acquisition of L2 German, time spent in an L2 environment (L2-speaking country), number of years of studying L2 German, years of L2 experience, current age, or participants' status as monolinguals or bilinguals when growing up. The different results in the two English-German studies is therefore conjectured to be due to the difference in the amount of German each group of participants used per week, along with the more bilingual language mode in the second study. These findings are summarized in Table 3.10 below; detailed statistical results for these findings can be obtained in the Appendix 3.2.

Table 3.10: Summary of language background factors that differed/did not differ between the two L1 English – L2 German studies

Significant difference between studies	**No significant difference between studies**
Hours of German used per week	German perception proficiency test scores
	Proficiency in German (self-assessment score)
	Age of acquisition of L2 German
	Time in L2 environment (L2-speaking country)
	Years of L2 German study
	Years of L2 German experience
	Current age
	Status in childhood as L1 English monolingual or bilingual

3.6 Discussion

To recap, the first L1 English – L2 German study showed a significant perceptual learning effect in the language of training, English: listeners showed a perceptual learning effect for the trained /f/-/s/ contrast in English, but did not show significant perceptual learning effects in German, neither for the German /f/-/s/ continuum, nor for the German /v/-/z/ continuum. Participants in Experiment 2 showed a perceptual learning effect on the trained contrast /f/-/s/ in the language of training, English; surprisingly, the perceptual learning effect on the English /v/-/z/ contrast did not reach significance. Most importantly, the listeners also showed perceptual learning effects in the untrained language, German, both on the German /f/-/s/ and on the German /v/-/z/ contrast. These findings about significant vs. non-significant perceptual learning effects are summarized in Table 3.11 below.

Table 3.11: Summary Table of the two L1 English – L2 German studies, highlighting which of the continua show perceptual learning effects.

[Notes: *n.t.* = not tested; shaded cells with bold text indicate significant perceptual learning effects.]

	English f/s	English v/z	German f/s	German v/z
SBU E-Ger	*** st. sign. effect (p=.008); effect size: 6.6%**	*n.t.*	n.s.; [effect size: 4.8%]	n.s.; [effect size: 1.2%]
CAL (E-Ger)	*** st. sign. effect (p=.009) effect size: 10.8%**	n.s. (p=.309) Effect size: 4.9%	***st. sign. effect (p=.010); effect size: 15.1%**	*** st. sign. effect: p<.001 effect size: 20.5%**

We have suggested that experiment 2 facilitated a more bilingual (English-German) language mode, which might explain the difference in results between the two L1 English – L2 German studies reported in this chapter. The listeners were provided with written instructions in English *and* German for each German subpart, and the participants completed a practice part in English *and* in German. Moreover, the participants in Experiment 2 completed the study in a quiet room within the German library of a German department, while the participants in Study 1 participated in a quite room within the Linguistics lab in a Linguistics department. Additionally, the participants in Experiment 2 also used German for more hours each week at the time of the study compared to participants in Experiment 1 (participants self-reported 8.5 hours/week in Experiment 2 vs. 4.9 hours/week in Experiment 1; p=.035). This possible influence of language mode and language use is in line with the claim that "bilinguals may function more or less as monolinguals an/or native speakers depending upon language component, task type, context, and speaker variables" (Grosjean 1998, quoted in Mack 2003: 310). Table 3.12 below summarizes the language variables that were analyzed to test for differences in language background between the two L1 English- L2 German studies, and to test for covariates to the perceptual learning effect size within each of the two studies.

Table 3.12: Summary of language background factors and whether they had an effect in the two L1 English – L2 German studies

Language Background Variable	Difference between the two studies	L1 English – L2 German study #1 (SBU)	L1 English – L2 German study #2 (CAL)
Hours of German used per week	✓		
German perception proficiency test scores			✓ (German v/z)
Proficiency in German (self-assessment score)			✓ (overall training effect; English f/s, English v/z, German f/s)
Age of acquisition of L2 German			
Time in L2 environment (L2-speaking country)			
Years of L2 German study			✓ (overall training effect, English f/s)
Years of L2 German experience			
Current age			
Status in childhood as L1 English monolingual or bilingual			

The results of the two L1 English- L2 German studies suggest that the adjustments to unusual phoneme contrasts in a listener's L1 can influence a listener's L2. When listeners are in a rather monolingual language mode (L1 English – L2 German study #1), however, significant category retuning might be restricted to the language in which they were exposed to ambiguous fricatives. The results of both language-specific as well as cross-linguistic perceptual learning effects with different language modes and language backgrounds of the participants suggests that the representation of phonemes in second-language learners are neither completely separated or entirely merged, nor static and fixed representations. Rather, these findings suggest that phonemes common between languages might have representations and processing mechanisms that are dynamically associated with each other (cf. Kroll et al. (2012) for a review), with stronger interconnectedness if both languages are concurrently activated and used (cf. de Bot et al. 2005). Recent research in bilingual processing has found that language processing in many areas is dynamic in nature:

> The dynamic nature of bilingual language processing is supported by what we take to be the central observation in the recent research: bilinguals cannot switch off one of the two languages at will. When they listen to speech, read, or prepare to speak in only one of their two languages, information about the language not in use is also active and influences performance (e.g., Dijkstra, 2005; Kroll, Bobb, &Wodniecka, 2006; Marian & Spivey, 2003). Most critically, these cross-language interactions can be observed at virtually every level of language processing, including those grammatical structures that are shared across languages (e.g., Hartsuiker, Pickering, & Veltkamp, 2004). (Kroll et al. 2012:231f)

In research at the level of sounds, a large number of studies have concluded that the phonological and/or phonetic (sub-) systems of bilinguals and second language learners interact (cf. Barlow et al. 2013). In Flege's SLM (Speech Learning Model, Flege 1995b, 1999, 2002, 2003; Flege, Schirru, & MacKay 2003), the "L1 and L2 phonetic subsystems of a bilingual exist in a 'common phonological space' and so will necessarily influence each other" (Flege et al. 2003:469). Similarly, Best & Tyler's (2007) Perceptual Assimilation Model for L2 speech perception (PAM-L2) postulates that phonetically similar phonemes common to a second language learner's L1 and L2 are equated with each other.

Either language (L1 or L2) can influence the other, though age of acquisition and the amount of usage ("input" and "output", Barlow et al. (2013: 69)) in each language has an impact on the directionality of the interaction between the two languages (cf. Barlow et al. (2013:69)). Moreover, while bilingual children or children growing up bilingually are less likely to show cross-linguistic influence between their two languages, bilinguals with a dominant language are more likely to exhibit cross-linguistic influences of the dominant language onto the less dominant language (Barlow et al. 2013:69). The participants in the English-German studies presented in this chapter are mostly novice L2 German learners whose dominant language is English. It was therefore expected that the dominant L1 English would exhibit a strong effect upon the non-dominant L2 German in this context (compared to the L1 German – L2 English study discussed in the following chapter).

A large number of research studies have found that bilinguals who speak languages that differ in the VOT values of their stops produce VOT values that are intermediate between the VOT value of either language. On the other hand, some production research suggests that bilinguals can show native-like phonetic realization of sounds in both of their languages: English-Spanish bilinguals have been shown to produce VOT values similar to native speaker values for each language (Magloire & Green 1999); similarly, (L2-dominant) early Greek-English bilinguals have been shown to produce native-like VOT values in stop contrasts in both of their languages in unilingual language modes for Greek as well as for (L2) English, albeit only in word-initial position for VOT in L2 English (Antoniou et al. 2010). The authors conclude that "bilinguals can produce monolingual-like phonetic values in both languages in most phonetic contexts, yet show evidence of interlanguage interaction in certain other contexts, indicating that the truth lies somewhere in between phonetic merger between the L1–L2 and completely independent systems for each language" (Antoniou et al. 2010:640). The findings of the two English-German perceptual learning studies reported in this chapter might be taken to point to a similar conclusion: Novice second language learners' phonological representations and their interactions are not stable, but adjust dynamically and flexibly to the input, social circumstances, language activation and language mode.

Another interesting aspect of the findings of the second English-German study is that the perceptual learning effects were numerically larger in German than in English. The effect size for English f/s was 10.8% (and for English v/z 4.9%, though not quite significant), whereas the effect sizes for German f/s and German v/z were 15.1% and 20.5%, respectively. This very strong generalization of perceptual learning effects to the novice L2 German is argued to be due to the fact that that the language system of the L2 is still developing and more malleable (cf. Baker & Trofimovich 2005:22), and therefore more "susceptible" to the influence of perceptual learning effects than the less flexible and more "developed" L1 English sound system (cf. Baker & Trofimovich 2005:22). L2 German listeners with a higher score on the L2 German perception proficiency test – compared to L2 listeners with a lower score on the proficiency test – also showed a statistically significantly stronger effect on the German v-z continuum ($F(1,14)=6.544$, $p=.023$*; results for a binned categorical variable: $F(1,14)=8.384$, $p=.012$*), as shown in the next two figures below. Figure 3.10 below shows a scatterplot which illustrates that the difference between the two conditions – i.e. the gap between empty circles (Cond A) and black circles (Cond B) – becomes larger as the German Proficiency Perception Score increases.

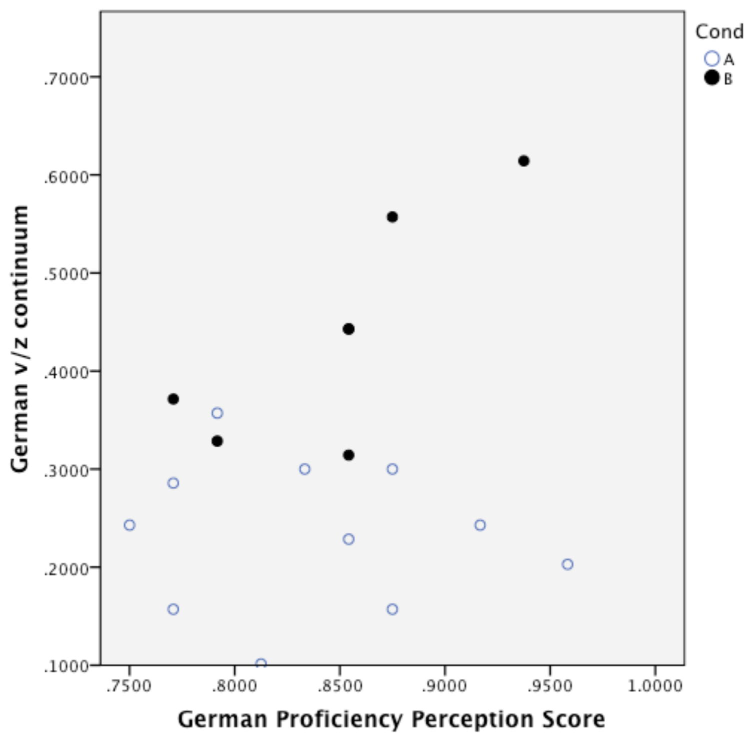

Figure 3.10: Scatterplot showing the relationship between Condition (the independent variable), the Percent /f/ responses on the German v/z continuum (the dependent variable), and the L2 German Proficiency Perception score (the continuous covariate).

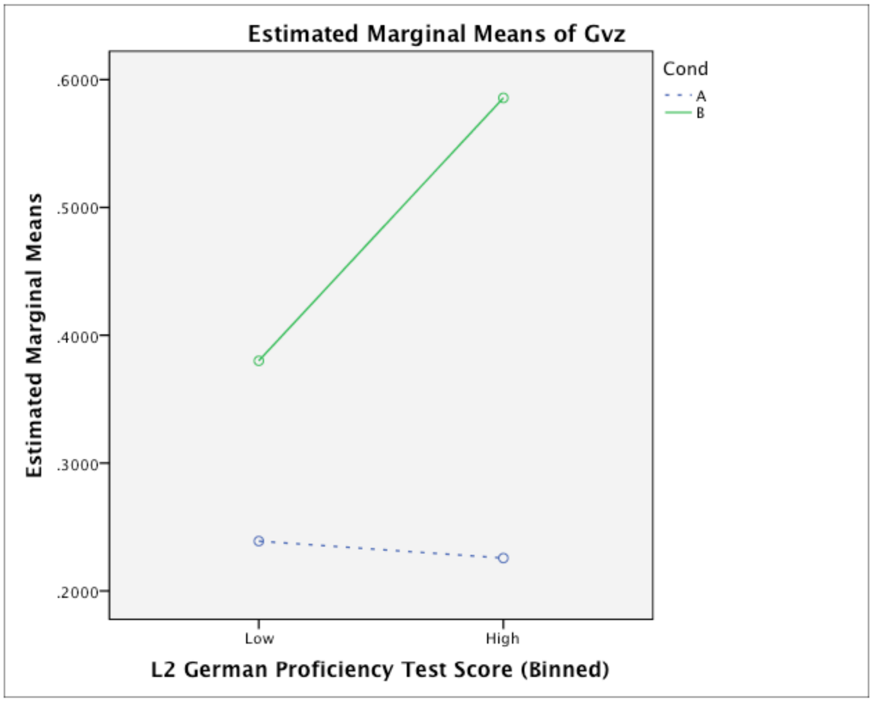

Figure 3.11: Perceptual learning effect (IV "Condition") for participants with low vs. high scores on the L2 German Proficiency Test (binned variable).

In this study, L2 listeners who are better at hearing L2 phoneme contrasts under noisy conditions – i.e. in the L2 German Perception Proficiency Test – are likely not only more apt at distinguishing new and (potentially) difficult L2 phoneme contrasts (especially in a lexical context, cf. discussion above), but also seem to have more malleable L2 phoneme contrast categories (at least for German /v-z/). Since the perception proficiency test did not test for phoneme distinctions in lexical content, but in nonsense words, the listeners were likely focused more on acoustic-phonetic and phonological details (even though they heard various recordings for each target phoneme, rather than the same token in each instance).

It seems plausible that the group of highly proficient L2 listeners had noticed (implicitly) that the German /v-z/ contrast differs phonetically, phonologically, and orthographically from the English /v-z/ contrast, thus creating novel and more malleable representations for this phoneme contrast. For example, listeners with higher perception scores in L2 German might have noticed allophonic differences in voiced fricatives (or voiced obstruents in general), namely the phonetic realization of the voiced fricatives in syllable-/word-final position. Moreover, although the German phoneme /v/ is typically described as a voiced, labiodental fricative, some investigations have documented that labiodental sounds often have weak frication, thus realizing /v/ as an approximant rather than a fricative (Iverson et al. 2008:1306; Scherer & Wollmann 1985), and is acoustically closer to labiodental approximant /ʋ/ that occurs in Dutch, based on canter of gravity values (Hamann & Sennema 2005b). Finally, it is now well-established that orthography can also influence phonological processing, and the phoneme /v/ is rendered differently in the orthography of English and German: whereas /v/ is usually expressed with the grapheme <v> in English, /v/ in German is expressed with the grapheme <w>, while German /f/ is expressed with the grapheme <v> (as in <Vogel> 'bird') or <f> (as in <Fahne> 'flag'). All of these differences between voiced fricatives, in particular /v/, between English and German might have led highly skilled L2 German listeners to have especially malleable representations for a voiced fricative contrast (/z-v/) in German, because it might have been interpreted as different from the voiced fricative contrast (/z-v/) in English.

Finally, we investigated whether the order of the presentations of the German fricative continua had an effect on the results. Note that the German fricative continua always occurred first, either German /f-s/ or German /v-z/, which were then followed by English /v-z/. The contrast that people were trained on in the lexical decision task, English /f-s/, was always presented last. It turns out that the effect for the German /v-z/ continuum was largest for those listeners who made judgments on the German /v-z/ continuum first, i.e. right after the lexical decision task, with an effect size of 28.76%. Listeners who heard the German /v-z/ continuum after the German /f-s/ continuum showed an effect size of merely 10.67%, indicating that the large effect on German v/z might be (partially) due to a timing/recency or ordering effect. This ordering effect is illustrated in the figure below.

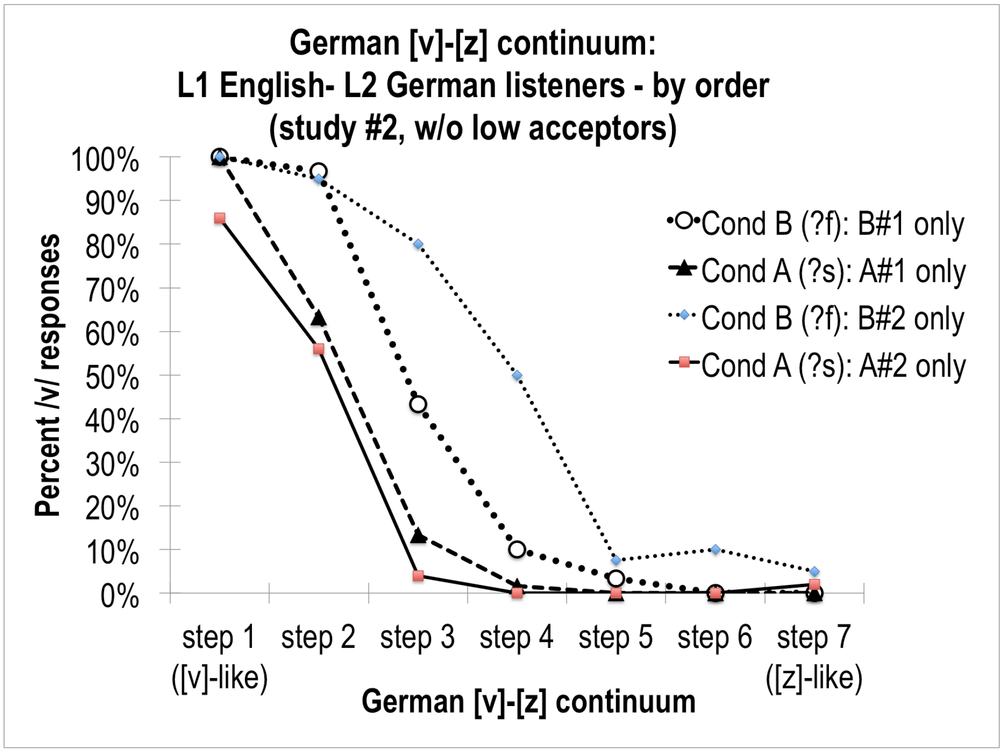

Figure 3.12: Order effect in the German /v-z/ continuum (L1 English – L2 German study #2)

Interestingly enough, seven out of eight listeners who had a "higher" score on the L2 Perception Proficiency Test also had the order of continua in which they heard the voiced fricatives (German /v-z/ continua) before they heard the voiceless ones (German /f-s/), referred to as order #2 above (and Condition A[?s]#2 and Condition B[?f]#2 in Figure 3.12 above). Most likely, this correlation was sheer coincidence. Figure 3.13 below illustrates the effects of the ordering of the German fricative continua onto the German /f-s/ continua in the second English-German study. There was no ordering effect on the German /f-s/ continuum, with listeners who heard the German /f-s/ continuum first showing a slightly smaller effect (15.08%) than listeners who heard the German /f-s/ continuum after the German /v-z/ continuum (15.29%).

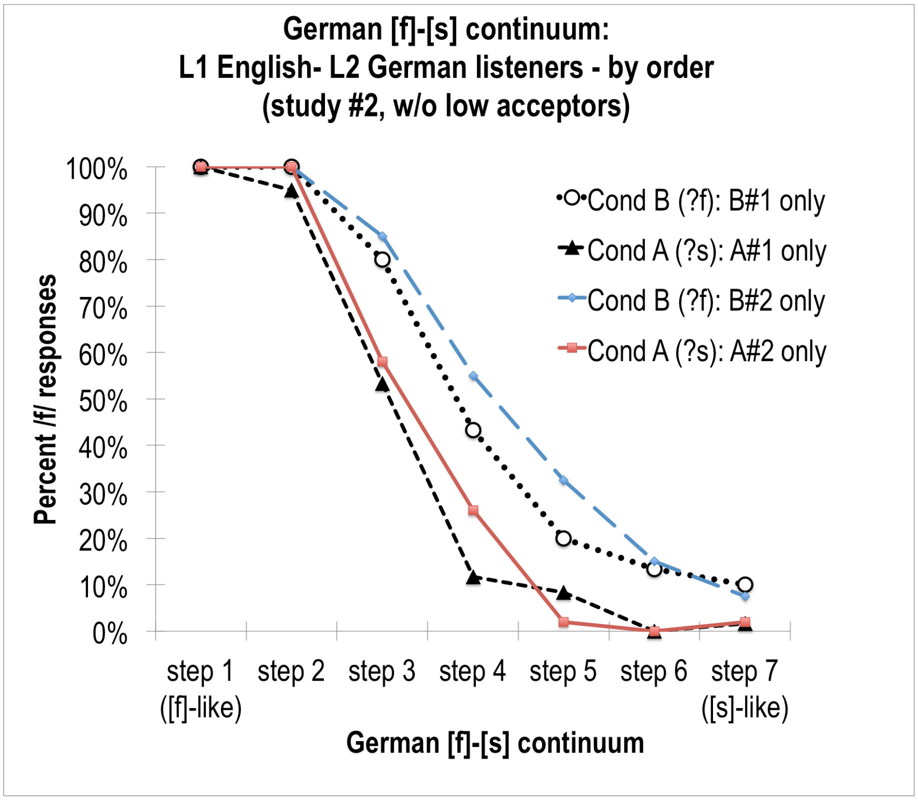

Figure 3.13: Order effect (n.s.) in the German /f-s/ continuum (L1 English – L2 German study 2)

Although it is not unambiguously clear why the German /v-z/ continuum showed such a large effect in the second English-German study, it appears it was in part driven by the very large effect when the German /v-z/ continuum was tested first, as well as by listeners' higher scores on the L2 German perception proficiency test. Novice L2 learners have been argued above to be especially susceptible to perceptual learning effects; those novice L2 learners with a keen ear for non-native phoneme distinctions, as evidenced by higher scores on the L2 German perception proficiency test, are likely more attuned to hearing and accommodating to small phonetic variation in their speech input, thus leading to stronger perceptual learning effects on the German /v-z/ continuum. Although it seems counterintuitive at first, perhaps perceptual learning effects sometimes become bigger when they generalize to a new case. For example, the results in Kraljic & Samuel (2006) show perceptual learning that was numerically slightly larger – though not significantly larger – on the phoneme contrast to which the effect was generalized, rather than on the trained phoneme contrast. (None of the other studies reported in this dissertation show an increase in the effect size of perceptual learning effects, when the learning generalizes to new phoneme contrasts). It seems likely that an effect size might be especially strong if it generalizes across a feature (here: voice) *and* a language, especially if the listeners generalize from their dominant, native language to their non-native language (cf. discussion above about stronger 'transfer'-effects from L1 to L2 especially in late bilinguals or L2 learners).

To sum up, the two L1 English – L2 German studies presented in this chapter show that cross-linguistic perceptual learning effects are possible, including generalization across both a phonological feature and across a language. The results further suggest that the generalization from a native to a novice second language leads to stronger effects because, as has been argued above, the L2 linguistic system is still more malleable and more susceptible to recent input. Finally, we speculated that the results might suggest that phoneme representations in novice second language learners are not static, and are not uniformly the same among all L1 English – L2 German speakers. Rather, a usage-based account and a dynamic view of bilingual language processing (e.g. Dynamics Systems Model of SLA; Kroll et al. 2012) might be able to explain why participants in L1 English – L2 German study #1 did not show statistically significant cross-linguistic perceptual learning effects: the infrequent use of German in the recent past and a rather unilingual language mode during the study might not have led to a strong activation of the German subsystem(s) and the interconnectedness between the English and German phoneme representations, and thus might not have led to a cross-linguistic perceptual learning effect. The participants in the L1 English – L2 German study #2, however, used German relatively frequently around the time of the study, and were arguably in a more bilingual language mode and setting. This bilingual mode might have provided activation of both languages and interconnections between the two phonological subsystems, which might have facilitated a strong cross-linguistic perceptual learning effect to their weaker non-native language.

4. Perceptual Learning in L1 German – L2 English speakers

This chapter reports on a perceptual learning study that was conducted with intermediate to advanced L2 English speakers (L1 German). This study tests whether non-native L2 English speakers show the same type of flexibility as native listeners when listening to a native English speaker with a novel accent in their L2, i.e. an unusual pronunciation of the f/s contrast. The study further set out to test how general this implicit learning in non-native listeners is, by testing whether it generalizes across the phonological voicing feature to the voiced fricative contrast v/z and across languages to the listeners' L1 German. All participants were trained on unusual pronunciations of the English f/s contrast, and were tested for perceptual learning effects on English /f-s/ and English /v-z/ continua, as well as on German /f-s/ and German /v-z/ continua. We also tested whether these effects are influenced by the listeners' proficiency level and language background in their L2. The L1 German – L2 English participants perceptually adapted to a novel accent in their L2 English, suggesting that the malleability of phoneme representations is not specific to native language grammars. Moreover, the listeners also generalized this perceptual learning effect from the English f/s contrast cross-linguistically to the German f/s contrast. However, the listeners did not generalize the perceptual learning effect to the voiced fricative contrast v/z in either English or German. We argue that perceptual adaptations in a second language do not seem to operate at the same abstract, featural level as in native language perception (cf. Chapter 2: English-only study; Kraljic & Samuel 2006). Rather, these findings suggest that non-native listeners adjust the representation of phoneme boundaries in their L2 at the level of individual phoneme contrasts and do not generalize these adaptation effects to phoneme contrasts that share relevant features.

4.1 Introduction

The previous chapter reported the results of two L1 English – L2 German studies that tested whether perceptual learning can lead to cross-linguistic effects when listeners are exposed to unusual pronunciations of a fricative place contrast within their native language, and tested on both their native language and their novice L2 language. The results of these studies provided evidence that the mental representations and mechanisms that are used for speakers' L1 and their novice L2 are neither statically "separate (independent)" or "shared (interdependent)" (Heredia & Brown 2012:272), but dynamically associated with each other (cf. De Bot et al. 2005; Cook 1995; Kroll et al. 2012). Differences in the language background and language mode of the participants in the two studies suggested that the representations of phonemes in these second language learners are more strongly connected with each other if the second language is used more frequently in the recent past and if participants have both languages relatively strongly activated at the time of the study.

The present chapter presents another study with second language learners; however, the L2 learners in this study differ from the populations in the previous chapter in that they are non-native speakers of English (L1 German- L2 English speakers), but proficient enough to complete the same lexical decision task which included some difficult vocabulary items in English. The purpose of this study is to determine how non-native German-English listeners deal with variation and idiosyncratic accents in their speech input, and to determine whether they also show evidence of integrated representations and generalization across phonological features and languages.

A few studies have investigated whether L2 learners show accommodation effects in L2 speech production when conversing with speakers of different accents. Beebe (1981), for example, found that L1 Chinese – L2 Thai learners produced more target-like vowels in Thai when they spoke with a native speaker of Thai, yet they produced more L2 Chinese accented vowels when they spoke with an L1 Chinese-L2 Thai speaker in Thai. Kim et al. (2011) found that L2 English speakers with medium proficiency – compared to L2 English speakers with either high or low proficiency – showed more adaptation effect when conversing with a native English speaker.

Another study also shows that at least some L2 learners converge with different native English accents when they speak in their L2 English. Zajac (2013) found in a pilot study that some, but not all, L1 Polish-L2 English speakers show evidence of accommodation in speech production when speaking with L1 English speakers of either an RP accent or a Canadian English accent. Accommodation was measured in terms of tapping and rhoticity. Only one out of four participants – the most fluent English speaker – showed accommodation effects on both variables (tapping and rhoticity). However, another participant who was almost as fluent in English did not show any signs of speech accommodation, and the least fluent participant produced less rhoticity when conversing with the speaker of RP. While Zajac's (2013) preliminary results could not determine which factors led to more accommodation in L2 speakers, the results strongly suggest that at least some L2 learners can adjust their speech output to become more similar to their interlocutors' native accent. The results also show that one participant's accommodation to Canadian English tapping increased over time while the same participant's accommodation to non-rhotic RP accent decreased over time, which Zajac (2013:236) takes to mean that the speech accommodation in non-native speakers might take place consciously or subconsciously for different aspects of speech.

The current study will test for similar adaptation effects in non-native speakers, though the focus will be on perceptual retuning effects in speech perception. The specific research goals for this study are as follows:

a) to test whether non-native listeners adjust L2 phoneme boundaries when they are exposed to unusual pronunciations in their second language. It seems plausible that L2 listeners might not adjust to unusual accents in their L2 as readily as L1 listeners, possibly because they might not perceive subtle differences in pronunciation or respond as automatically to unusual accents in an L2 (see Zajac 2013, for example). On the other hand, it also seems plausible that L2 listeners might adjust to novel pronunciations in their L2 more readily than L1 listeners, possibly because the phoneme representations in their L2 might be inherently less stable and more 'in flux' than native phoneme representations. The latter outcome would be similar to the results of the second L1 English – L2 German study in the previous chapter, in which the listeners showed stronger perceptual learning effects in the untrained L2 than in the language of training, their L1. Furthermore, non-native listeners' proficiency level in their L2 was hypothesized to play an important role in how readily listeners adjust their phoneme boundaries in their L2, with less proficient learners showing a different degree of adaptation in their perception. Again, theoretically it seems plausible that non-native listeners with a lower L2 proficiency might show smaller perceptual learning effects, because they might not be able to notice or respond to novel accents in their non-native language. On the other hand, the stronger results in the non-native language compared to the native language of the English-German study #2 (in the previous chapter) predict that less proficient non-native listeners will show stronger perceptual learning effects. Presumably, the phoneme representations of non-native listeners, especially less proficiency non-native listeners, are more malleable compared to those of more proficient non-native listeners or native listeners.

b) to test whether non-native speakers generalize perceptual learning effects cross-linguistically from their L2 to their L1, i.e. in the form of 'regressive transfer' rather than the 'progressive transfer' (L1 to L2) that was tested in the two L1 English – L2 German studies of the previous chapter. While the field of SLA has historically focused on the role of transfer from the native language onto the non-native language, more recent research has provided ample evidence for 'regressive transfer' from the L2 onto the L1 (e.g. Van Hell & Dijkstra 2002; Cook 2003; Chang 2010, 2011, 2012, 2013; Brown & Gullberg 2010; cf. Chapter 1: Introduction), as well as evidence for L1 language attrition (e.g. De Bot & Clyne 1994; Schmid 2004; Köpke et al. 2007; Montrul 2008, 2013). At the same time, the influence of the L2 onto the L1 has often been found to be smaller than effects of the L1 onto the L2. For example, in a silent reading study with English-French and French-English bilinguals that tested for interlingual homophone effects (Haigh & Jared 2007), English-French bilinguals who performed the lexical decision task in their native language (English) showed less influence from their French than French-English bilinguals who performed in their native language (French, experiments 4-6 vs. 1-3). Dijkstra (2005, cited in Haigh & Jared 2007:636) claimed that the amount of influence of the L2 onto the L1 (in reading) depends on the relative "strength" of the bilingual users' languages.

More pertinently, in a study on phonetic perceptual drift, novice L1 English – L2 Korean learners showed changes in the pronunciation of their native language English within their first few weeks of studying Korean in an intensive course at a South Korean university (Chang 2010). Specifically, the participants adjusted their VOT values and/or the onset of the fundamental frequency (*f*0) and raised their vowel space, thus rendering their L1 English speech more Korean-like in terms of these phonetic-acoustic features. This study provides evidence that regressive transfer from the non-native to the native language can happen rather quickly at the phonetic level, even for phonemes that are realized differently in terms of phonetic-acoustic cues in the two languages involved (in this case, VOT values in English and Korean).

c) to test whether non-native speakers generalize perceptual learning effects across sub-phonemic features and across languages. This question addresses whether perceptual learning in a non-native language is as broadly tuned as perceptual learning in a native language, which we have argued to target the level of representation where phonemes and phonetic features are mapped onto one another (Chapter 2). As the English-only study in Chapter 2 has shown, listeners who are exposed to a novel [?fs] accent in their native language generalize perceptual learning effects on the place contrast of voiceless fricatives (f/s) to voiced fricatives (v/z). The previous chapter (L1 English – L2 German study #2) has shown that cross-linguistic perceptual learning can generalize across both a sub-phonemic feature (*voice* on the fricative contrasts) *and* across language (from L1 English to L2 German). The current study is therefore designed to test whether cross-linguistic perceptual learning can generalize across sub-phonemic features when listeners perceptually adapt to a novel accent in their non-native language. Further, this study tests whether cross-linguistic perceptual learning also generalizes both across a sub-phonemic feature (from voiceless to voiced fricative contrasts) *and* across languages (from English to German) when listeners are exposed to unusual accents in their non-native language.

Overall, the goal of this study is to further probe the organization of the phonological systems in L1-L2 listeners from the perspective of *non-native* listeners. One important question is whether non-native listeners have the capability to retune phonetic boundaries following exposure to non-native speech with ambiguous phonemes. If so, the next question concerns whether perceptual learning in non-native listeners is as broadly tuned as in native listeners, leading to generalization of perceptual learning effects to untrained phoneme contrasts and the listeners' untrained language. Finally, this study will again address the question of whether phonetic retuning in non-native listeners is influenced by the listeners' L2 language proficiency, language mode, and language use factors. While there are an increasing number of studies on the perceptual tuning of native speakers and on modifications in speech production to accommodate to different accents of interlocutors, not much is known yet about the ability of L2 learners to accommodate – especially perceptually – to novel, native accents that differ from the L2 accent to which they are typically exposed.

The question remains whether non-native listeners have L2 perceptual representations that can adjust – implicitly – to novel native accents. The literature on perceptual tuning effects typically assumes that the effects reflect implicit learning results of which the listener is not aware. Tapping and rhoticity, the measures for speech accommodation in Zajac (2013), likely have more salient features compared to intermediate phoneme realizations, which are typically used for perceptual learning studies (including the studies reported in this thesis). Moreover, the studies reported above make use of existing native accents – both of which the L2 English listeners have likely been exposed to and possibly even taught about explicitly in an L2 classroom. Moreover, these existing native accents also differ on more variables than just tapping or rhoticity: for example, also on how vowels are realized.

Only one study to date has explicitly tested perceptual learning in non-native speakers and cross-linguistic effects of perceptual learning, though it leaves many of the questions open that are addressed in this dissertation. Reinisch, Weber & Mitterer (2013) tested L1 Dutch and L1 German-L2 Dutch speakers living in The Netherlands on unusual pronunciations of the /f-s/ contrast in native Dutch (experiment 1 and 2, respectively). In an additional study (experiment 3), Reinisch and colleagues trained L1 Dutch-L2 English speakers on a non-native (L1 Dutch)

English accent with additionally modified /f/ or /s/ fricatives, and tested for cross-linguistic effects on /f-s/ continua in both their L2 English and their L1 Dutch. The findings showed a) that non-native listeners living in an L2 environment can adjust a phoneme contrast in their L2 (experiment 2), and b) that perceptual learning can have cross-linguistic effects when L1 Dutch-L2 English listeners listened to an L1 Dutch-L2 English accent with modified /f/ or /s/ fricatives, and were then tested on /f-s/ continua in L1 Dutch and L2 English (experiment 3). However, these findings still leave open the question of whether non-native speakers who do not live in an L2 environment (unlike the participants in experiment 2 in Reinisch et al. 2013) can adjust perceptually to unusual pronunciations in their L2. Moreover, the results of the cross-linguistic English-Dutch effects are based on L1 Dutch-L2 English participants listening to a non-native English accent with modified fricatives. This cross-linguistic effect might be due to listeners ascribing the unusual pronunciations of the English fricatives to the speaker's native language (Dutch) or the speaker's general idiosyncratic and non-native speech pattern, and therefore generalizing the effects from English to Dutch.

In sum, we still need to determine a) whether perceptual learning is possible in second language learners who are not immersed in an L2 environment and are not using their L2 on a daily basis with native speakers, b) whether perceptual learning shows cross-linguistic effects when L2 learners listen to a novel, native accent, and c) whether perceptual learning can generalize to untrained phoneme contrasts in non-native listeners' L2 as well as L1. The study reported in this chapter is designed to answer these questions.

In particular, the current study specifically addresses whether non-native listeners' "perceptual representations developed for L2 phonetic segments resemble those of native speakers of the target L2" (Flege 2003:326) with regard to their flexibility and ability to implicitly align with novel native accents. Sebastian-Galles et al. (2006:1277) conducted a study with Spanish-Catalan and Catalan-Spanish bilinguals; Spanish-Catalan but not Catalan-Spanish bilinguals failed to reject Catalan non-words that differ from real Catalan words by one vowel contrast that is notoriously difficult for L1 Spanish listeners to distinguish. However, both groups of participants showed N400 effects when processing the Catalan non-words, which Sebastian-Galles et al. (2006) interpret to mean that Catalan-Spanish speakers had included a phonological variant in their lexical representation that reflects typical Spanish-accented pronunciations of Catalan words. They conclude: "first-language representations seem to be more dynamic in their capacity of adapting and incorporating new information". Specifically, "substantial plasticity is observed for *first language* phonological representations. Indeed, it is commonly reported that natives are able to easily adapt their phonological system to difficult situations (such as distorted speech, background noise) and dialects; this ability seems to be less present in the second language" (Sebastian-Galles et al. 2006:1288). The current study sets out to test whether L2 phoneme representations are indeed less malleable than L1 phoneme representations.

If L2 leaners can adjust their phoneme representations to novel L2 accents, this leads to the question of whether these changes in a listener's L2 grammar can also affect the representations of this listener's L1. A growing body of research has revealed that bilinguals and second language learners exhibit not only influence of the L1 onto the L2, but also influence of the L2 onto the L1 (e.g. Cook 2003; Chang 2010, 2011, 2012, 2013). The evidence even includes findings of cross-linguistic L2 to L1 influence in co-speech gesturing, which shows that, for example, L1 English-L2 Dutch speakers gesture differently than monolingual English speakers (Gullberg 2009,2011; cf. Gullberg (2013:425-428) for an overview).

An extension of the question of whether non-native listeners' representations are as dynamic and adaptive as native phonological representations is the question of which level of sound representations is adjusted: Do perceptual accommodations to novel native accents in non-native speakers target abstract, higher-order features, or do they apply at a surface level, such as at a specific phoneme or phoneme contrast? If perceptual learning in a listener's non-native language also happens at the level of features (Kraljic & Samuel 2006) or the level at which phonological features and phonetic cues/features are mapped onto one another – as argued in Chapter 2 – then perceptual learning on the place contrast of the voiceless fricatives /f/-/s/ should generalize to the place contrast of voiced fricatives /v/-/z/ in non-native listeners as well. However, if perceptual learning in non-native listeners does not apply at the level of abstract features (or their mapping onto phonetic cues) but rather at the level of individual phonemes or phoneme contrasts, then the perceptual learning effect of the /f-s/ contrast should not generalize to the voiced fricative contrast /v-z/ in either of the listeners' languages.

In fact, there is evidence that L2 listeners show learning effects that are – at least in certain conditions – restricted to only those L2 sounds or sound contrasts on which the L2 learners were specifically trained. Laboratory studies on "speech perception training" for vowels (Nishi & Kewley-Port 2007:1496) in non-native listeners have shown that L2 English listeners trained on English vowels only showed improvement for the specific vowels they were trained on, but no improvement for untrained vowels (Akahane-Yamada et al. 1997; Sperbeck et al. 2005; cited in Nishi & Kewley-Port 2007:1497). Similarly, while L1 Japanese-L2 English listeners trained on the three most difficult English vowels generalized their identification improvements to untrained words containing these trained vowels, they did not generalize learning to untrained vowel phonemes (Nishi & Kewley-Port 2007:1506).[16]

The question of whether or not non-native listeners generalize learning to untrained phones might also depend on the type of sounds, features or cues involved in the phonemes or phoneme contrasts on which listeners are trained, and even the phonetic context in which these sounds occur. While the studies reviewed above were concerned with non-native speakers learning vowels, L2 listeners have been shown to generalize improvements on the voicing contrast in bilabial stops – a difference in VOT values – to stops at an untrained place of articulation, alveolar stops (McClaskey, Pisoni & Carrell 1983, cited in Nishi & Kewley-Port 2007:1497; McReynolds & Bennet 1972, cited in Nishi & Kewley-Port 2008:1480). Further, Thomson (2012) found that phonetic training helped L1 Mandarin-L2 English learners improve their ability to identify English vowels, yet only in certain phonetic contexts. The learners were trained on vowels that were preceded by bilabial stops; they generalized their identification improvement to syllables in which the vowels occurred after velar stops but not to syllables in which the vowels occurred after alveolar fricatives. While Thomson (2012:1250) does not have a clear explanation for these results, it appears that the improvement for vowel identification can be limited to specific phonetic contexts, depending on the training material.

There are of course some notable differences between studies on accommodation in speech production and perception training in L2 learners. Nishi & Kewley-Port (2007), for

[16] The authors consider that this might be the result of different "learning mechanisms" when training on specific vowels (citing Goldstone 1998): "their improvement may not represent perceptual learning but rather a strategy shift in one of the underlying mechanisms to attend only to the task-relevant cues that resulted in higher performance" (Nishi & Kewley-Port 2007:1506). They conclude that "learners may learn to ignore cues that are related to other categories" (Nishi & Kewley-Port 2007:1508) when they are focused on making improvements on a subset of L2 vowels. The same authors conducted another study to test a hybrid approach (training on a subset and on the entire set) (Nishi & Kewley-Port 2008); they suggest that "the cognitive mechanisms involved during training might be different" (Nishi & Kewley-Port 2008:1490) in the case of the subset training and training with the entire vowel set. They indicate that "Juslin, Olsson, and Olsson (2003) explored the difference between categorization and multiple-cue judgment tasks, and found that people shifted strategies from simple memorization of exemplars to abstraction of acoustic cues when the criterion changed from a binary to a continuous, probabilistic variable." (Nishi & Kewley-Port 2008:1490).

example, compared their training participants to other L2 learners who were exposed to English in the US in a naturalistic setting – the control group – and who did not show significant improvements in their vowel perception over the course of three months. Nishi & Kewley-Port (2007:1505) interpret this as a sign that mere exposure to the target vowels does have the same effects as targeted training sessions in the identification or discrimination of L2 vowel sounds. Perceptual learning studies, on the other hand, merely expose listeners to unusual, "accented" speech and measure small, implicit (and presumably temporary) adaptation effects that result from this exposure to a novel accent, without listeners being aware of the goal of the study. Moreover, in our case, listeners are exposed to unusual realization of a phoneme contrast involving consonants, and a phoneme contrast that is common to both their L1 and their L2. Unlike a lot of the second language acquisition literature, this study does not address how non-native learners acquire novel L2 phonemes or novel L2 phoneme contrasts, but whether unusual accents in the listeners' L2 can lead to implicit adjustments of phoneme boundaries for phonological contrasts that exist in both the listeners' L2 and L1.

4.2 Methodology

4.2.1 Participants

Fifty-two L1 German – L2 English listeners in Germany participated in the study[17]. The participants were recruited from the Freie Universität in Berlin, Germany, provided written consent before participating, were at least 18 years old, and were paid for participating in this study. No participant reported any current hearing impairments, language or learning disabilities. All participants classified German as their native language (eleven reported having grown up at least partially with two languages during infancy and/or childhood, but all classified German as their native and dominant language; the other languages participants grew up with were: French (1), Turkish (1), Russian (1), Spanish (2), Portuguese (2), English (1), Greek (1), Italian (2).) Most participants grew up in Germany or a German-speaking country (one each in a German-speaking area in Italy and Switzerland); three participants moved to Germany around age 1 (one

[17] The results reported are based on the data of 51 (N=52-1) participants, because one participant accepted fewer than 50% of the critical items in the Lexical Decision task. See below for details.

participant each from Turkey, Cuba, and Russia); one participant lived in Brazil for the first two years after birth, then in Chile for three years before moving to Germany; one participant moved to Belgium for 1.5 years at the age of 1.5 years old and to Italy for several years at the age of 7; one person lived in Hong Kong from 6-7 years old (attending a German-French school); one participant lived in Brazil from 5-10 years old; one participant lived in the US from 6-12 years old; and one participant lived in Ireland from 13-17 years old.

All participants also learned English as a second language, on average at age 8.88 years old [median: 10, range: 5-12 (excluding one English-German bilingual who reported having grown up with both English and German from infancy, but having been raised in Germany and being German-dominant)].

4.2.2 Assessment of knowledge of critical stimuli

The goal of this study was to determine whether non-native listeners of English can be shown to adjust their phoneme category boundaries in their non-native language and, if so, whether this effect carries over into their native language German, and to the untrained voiced fricative contrast /v-z/. As discussed in the introduction, the language background, proficiency level and experience with English are very important in this study, because perceptual learning is not likely to take place if the participants are exposed to unusual pronunciations of fricatives within lexical items that they do not know or if they do not accept the critical stimuli, especially those with ambiguous fricatives.

We therefore gave the L2 English listeners an additional "vocabulary questionnaire" after they completed the perceptual learning study on the computer (cf. Appendix 4.1 and 4.2). This test consisted of a list of words, namely the forty critical items, presented in writing. Each participant was asked to indicate how sure they were about using each word in a written sentence by indicated their answer on a scale from 0 (completely unsure) to 5 (definitely sure). These questionnaires were scored by adding up the answer choices, with a maximum possible score of 200 points. Overall, the non-native listeners judged that they knew most of the words pretty well (when they saw them in writing), and achieved an average score of 169.55/200 (SD 22.45) (84.77%), ranging from 95 to 198; the median was 174.0, and two modes of 175.0 and 187.0.

The distribution of the individual scores is shown in Figure 4.1 below. (Each vocabulary item was worth up to 5 points, for a total maximum score of 200.)

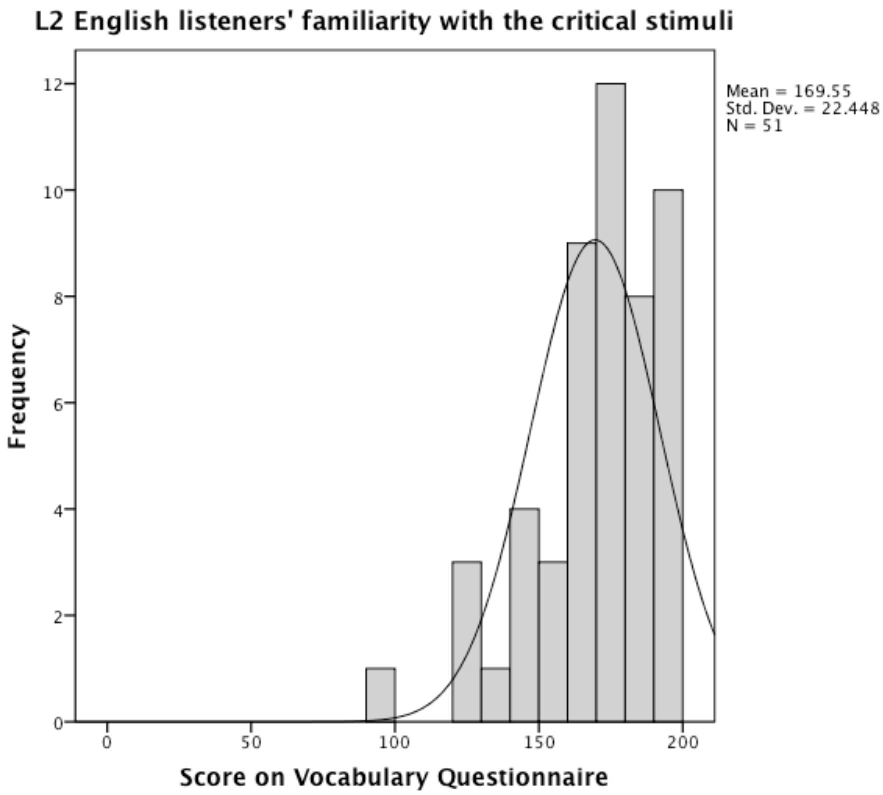

Figure 4.1: Histogram of the L1 German – L2 English participants' scores on the L2 English vocabulary questionnaire about the forty critical stimuli of the Perceptual Learning task.

In addition to the vocabulary questionnaire, all except one of the L2 English listeners accepted more than half of the modified, critical stimuli during the lexical decision task of the main perceptual learning experiment. The one L2 English listener who accepted less then half of the modified critical stimuli was excluded from all further analyses. The distribution of the number of ambiguous stimuli that individual L2 English listeners accepted is presented in Figure 4.2 below.

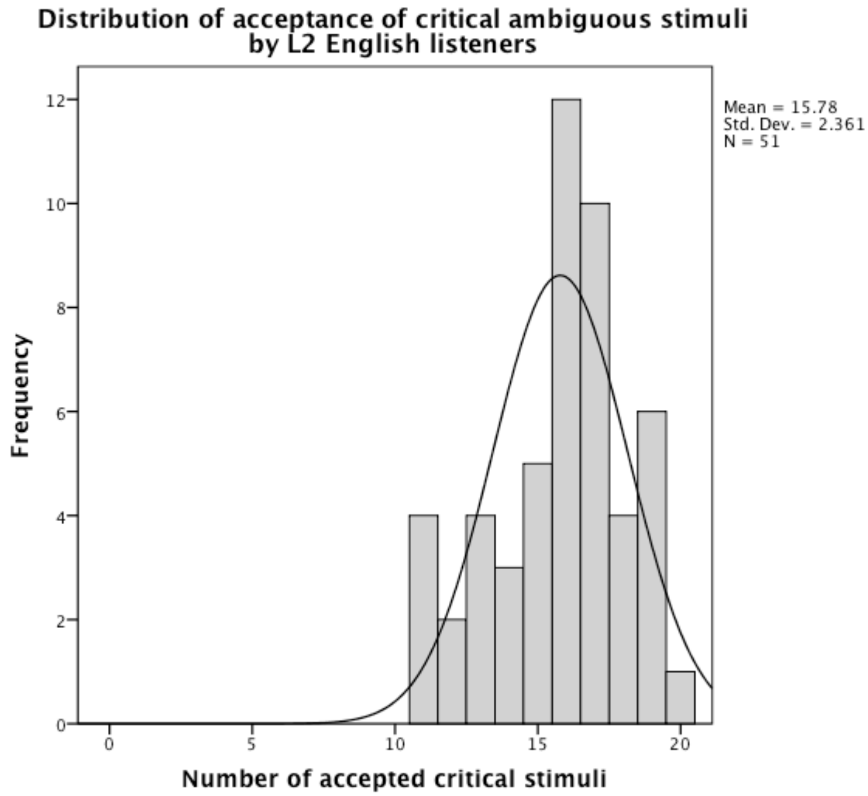

Figure 4.2: Histogram of L1 German – L2 English participants' number of accepted ambiguous critical stimuli during the Lexical Decision task (Note: N=51; The one low acceptor is excluded from further analysis.).

On average, the remaining 51 L1 German – L2 English listeners accepted 15.8 out of the 20 ambiguous critical stimuli (78.9%, SD 2.36), ranging from 11/20 (55%) to 20/20 (100%). This suggests that the non-native listeners largely knew and were able to recognize the modified stimuli. The analyzed 51 L1 German – L2 English participants' knowledge of English vocabulary was therefore assumed to be sufficiently large to allow for perceptual learning effects to take place.

4.2.3 Design, materials, and procedure

The design of the L1 German – L2 English study is identical to the design of the L1 English – L2 German study #2 presented in Chapter 3, which follows the design in Norris et al. (2003). In short, participants first completed the same auditory lexical decision task in English as in the previous studies, including 20 English words with either ambiguous /f/ or /s/ fricatives, depending on the condition – (cf. section 2.1.2 above for a discussion of the design of the Lexical Decision task). As in the other studies, this Lexical Decision task was followed by a phoneme categorization task, which included continua of voiceless and voiced fricatives contrasting in place (/f-s/ and /v-z/, respectively) in both English and German, i.e. a four-part auditory phoneme categorization task just like in the English-German study #2 reported in Chapter 3 (cf. 3.2.2.2 Design and Procedure for English-German study 2). This L1 German- L2 English study also used the same recordings and sound files for the English stimuli as in the English-only study from Chapter 2 (cf. section 2.2.2.2 Stimulus selection and 2.2.2.4 Stimulus construction for details), and the same recordings and sound files for the German stimuli as in the English-German studies (L1 English – L2 German study #2) from Chapter 3 (cf. section 3.2.2 for details).

The general procedure also followed the previous studies: participants were randomly assigned to one of the two training conditions and were not told that the study contained ambiguous sounds. As in all other studies, the experiment was presented through SuperLab software on a MacBook Pro laptop, and participants listened to the stimuli at a comfortable (low) volume through Sennheiser HD 555 headphones. As usual, participants conducted two practice parts: One for the Lexical Decision task and one for the phoneme identification task. Just like in the L1 English – L2 German study #2, a bilingual language mode was facilitated by conducting the practice part for the phoneme identification task in both English and German and by including additional German instructions after the English instruction for the German continua portions (cf. 3.2.2.2 Design and Procedure for English-German study 2). The experiment took place in a quiet room in the building of the JFK-Institute of the Freie Universität in Berlin. This locale likely also heightened the participants' bilingual language mode by increasing their activation of English (and associations with the US culture), as they were surrounded by posters, media and pictures relating to the English language and the US culture. Table 4.1 below presents an overall summary of the experimental design for the L1 German – L2 English study, including stimuli details about both phases of the study.

Table 4.1: Summary of the experimental design for the L1 German – L2 English study (Berlin)

Phase	Stimuli	Language	Cond. A ([?s])	Cond. B ([?f])	Order	*
Phase 1: Lexical Decision Task	**critical stimuli**	English	20 [?s] words	20 [?f] words	randomized *(except first three filler words)*	*same*
			20 [f] words	20 [s] words		
	fillers		60 filler words	60 filler words		
			100 non-words	100 non-words		
Phase 2: Phoneme Categorization	**7-step continua**	German	/f-s/ continuum	/f-s/ continuum	counter-balanced	*same*
		German	/v-z/ continuum	/v-z/ continuum		*same*
		English	/v-z/ continuum	/v-z/ continuum	always next-to-last	*same*
		English	/f-s/ continuum	/f-s/ continuum	always last	*same*

* Same as English-German study #2 (from Chapter 3)?

4.3 Results

4.3.1 Lexical decision results

The participants[18] in this study performed relatively well overall in the lexical decision task, which they completed in their L2 English, with a mean overall accuracy rate of 85.35% (SD 5.39%), ranging from 71.00% to 94.50%. The table below illustrates the mean accuracy for all words (critical stimuli and filler words): 84.6% (SD 7.1%); individual scores ranged from

[18] The analysis excludes the one L1 German – L2 English participant who accepted fewer than half (9 out of 20) of the ambiguous critical stimuli during the Lexical Decision Task. This one participant was excluded from all further analyses.

63.00% to 98.00%. Table 4.2 below further shows that participants were also faster to correctly accept words (355.7 ms, SD 124.0) than to correctly reject non-words (541.3 ms, SD 210.2).

Table 4.2: German-English study: Lexical Decision (all stimuli)

	words	**non-words**
% correct*	84.6%	86.1%
(SD)	(7.1%)	(8.3%)
RT (in ms)**	355.7	541.3
(SD)	(124.0)	(210.2)

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

** Mean accuracy (i.e. "yes" responses to critical items)*

The next table, Table 4.3 below, illustrates the mean overall accuracy for unmodified and modified critical stimuli for the L1 German – L2 English listeners, excluding the one participant who accepted fewer than half of the modified critical stimuli (cf. above). Overall, the response time (RT) and acceptance data suggest that the critical stimuli sounded acceptable to the non-native listeners as well, although non-native listeners also accepted slightly fewer of the ambiguous critical stimuli and responded slightly more slowly to the ambiguous critical stimuli compared to the natural critical stimuli. Participants accepted fewer (12.86%) of the ambiguous [?f] compared to natural [f] items (75.6% compared to 88.46%, respectively) and responded about 76.9 more slowly to ambiguous [?f] items. Participants also accepted slightly fewer (2.28%) of the ambiguous [?s] items compared to natural [s] items (82.12% compared to 84.40%, respectively) and responded slightly more slowly to ambiguous [?s] items (38.9 ms slower). (Note again, though, that each participants heard either only ambiguous [?f] and natural [s] words OR ambiguous [?s] words and natural [f] words, just like in the English-only study and the English-German studies reported in the previous chapters.)

Table 4.3: German-English study: Lexical Decision (critical stimuli)

	Natural stimuli		Ambiguous stimuli	
	[f]	**[s]**	**[?f]**	**[?s]**
% correct* (SD)	88.46% (10.08%)	84.40% (8.94%)	75.60% (11.30%)	82.12% (11.59%)
RT (in ms)** (SD)	347.8 (151.9)	336.7 (141.0)	424.7 (160.6)	375.5 (129.3)

** Mean accuracy (i.e. "yes" responses to critical items)*

*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

I carried out separate RM ANOVAs on the lexical decision results, one on acceptance rates of the critical stimuli and one on RT data. The analyses were performed separately using participants (F1) or the critical stimuli (F2) items as repeated measures. Additional factors in the ANOVAs were 1) training condition, i.e. whether participants heard ambiguous [?f] and unambiguous [s], or ambiguous [?s] and unambiguous [f] pronunciations, a between-participants but within-items factor, and 2) stimulus type, i.e. whether the critical stimuli contain an /f/ or an /s/ phoneme, a between-item but within-participants factor.

The data show that the difference in acceptance rate was larger for /f/-stimuli ([f] stimuli: 88.5%, [?f] stimuli: 75.6%) than it was for /s/-stimuli ([s] stimuli: 84.4%, [?s] stimuli: 82.1%) as indicated by a significant interaction of stimulus type and condition for both F1 and F2 analyses (F1(1,49)=35.889, p<.001; F2(1,38)=7.142, p=.011). Participants also generally accepted (slightly) more natural stimuli ([f] stimuli: 88.5%, [s] stimuli: 84.4%) than ambiguous ones ([?f] stimuli: 75.6%, [?s] stimuli: 82.1%), a (marginally, in the case of the F-1 analysis) significant main effect of condition (F1(1,49)=3.932, p=.053; F2(1,38)=14.645, p<.001). Neither the by-subject nor the by-item analysis showed a significant effect of stimulus type on acceptance rates (F1(1,49)=.942, p=.337; F2(1,38)=.031, p=.861).

The RT data show a significant interaction between condition and stimulus type in the by-subject analysis only (F1(1,49)=20.088,p<.001; F2(1,38)=.920,p=.344). There is a main effect of condition in the by-item analysis (F2(1,38)= 7.466, p=.009) but not in the by-subject analysis (F1(1,49)=.240,p=.626). The main effect of stimulus type is significant in the by-subject analysis (F1(1,49)=5.451,p=.024) but not in the by-item analysis (F2(1,38)=.286, p=.596).

4.3.2 Results for voiceless fricatives (in English and German)

As in previous chapters, the analysis of the responses to the fricative continua is based on the average percent labial (%f and %v, respectively) calculated for each continuum in each language for each participant. The data was analyzed after the exclusion of one participant who accepted fewer then half of the ambiguous critical stimuli (cf. above, section XXX).[19] A RM ANOVA was performed on percent labial (%f, %v) responses to the continua. The analysis included two within-participant repeated measures factors: language (English, German) and phoneme contrast (f/s contrast, v/z contrast). The analysis further included the between-subject factor of condition (training condition with ambiguous [?f] words or ambiguous [?s] words). (As in the literature, RT is not analyzed, and F2 (by item) analysis is not conducted on a single continuum with seven steps.)

There was no significant three-way interaction between language, phoneme contrast, and condition (F(1,49)=.146; *p*=.704). There was, however, a significant interaction between language and phoneme contrast (F(1,49)=76.156; *p* <.001), which shows that the difference in average percent labial responses between English and German is higher for the v/z contrast (19.89%: English average: 52.37%; German average: 32.48%) than for the f/s contrast (1.38%: English average: 41.02%; German average: 42.40%). There was also a significant interaction between phoneme contrast and condition (F(1,49)=8.277; p =.066), indicating that the difference between conditions is larger for the f/s contrast (average for English and German f/s: 8.05% [=9.71%+6.40%/2]) than for the v/z contrast (average for English and German v/z: -0.05%

[19] Two participants each showed one ceiling effect in one of the four conditions by categorizing all of the fricatives of a continuum as *one* particular sound. Following Tabachnick & Fidell's (2007:77) suggestion, these outlying cases were replaced with the "next most extreme score" (plus one/minus unit in the ANOVA's) for this continuum from a participant in the same condition. This way, the data of these two participants is not lost, while the effect of their one-sided responses (though in the predicted direction!) is mitigated. Specifically, one participant in condition B ([?f]) categorized all of the English f/s continuum as /f/. This 100% labial (f) score was replaced with the next highest % labial score plus 1: 68.57%+1=69.57%. Further, one participant in condition A ([?s]) categorized all of the German v/z continuum stimuli as /z/. This 0% labial (v) score was replaced with the next lowest score minus 1: 21.43%-1%=20.43%.

[=0.68%+ -0.77%/2]). There was also a main effect of language ($F(1,49)=51.460$, $p <.001$), showing that the overall percent labial responses for English were statistically significantly larger (46.7%) than the overall percent labial responses for German (37.5%). Finally, there was also a main effect of condition ($F(1,49)=4.260$, $p=.044$), which shows that overall, participants had higher percent labial responses in condition B ([?f]) (44.1%) than in condition A ([?s]) (40.1%), although this effect is dominated by the interaction between condition and language, discussed above. (Lastly, there was no statistically significant interaction of language and condition ($F(1,49)=.848$, $p=.362$), nor a statistically significant effect of the variable phoneme contrast ($F(1,49)=.203$, $p=.654$).

The most crucial results are the tests for perceptual learning effects, i.e. statistically significant effects of condition (condition A [?s] vs. condition B [?f]) for each of the four individual contrasts (English f/s, English v/z, German f/s, German v/z); These results are obtained through simple main effects by means of pairwise comparisons, which tests the effect of the factor condition (i.e. the difference in percent labial responses between the two conditions) at each of the levels of the factors language (English, German) and phoneme contrast (f/s, v/z). Participants showed a perceptual learning effect on the English f/s continuum, the language and phoneme contrast of the training portion of the study. Participants in the ambiguous [?f] condition categorized more items on the English /f-s/ continuum as the labial fricative /f/ (45.97%) than participants in the ambiguous [?s] condition (36.26%), a significant simple main effect ($F(1,49)=9.472$, $p=.003$) obtained through a planned pairwise comparison. The difference in percent /f/ responses between the two conditions amounts to 9.71%. Figure 4.3 below illustrates this successful perceptual learning on the English f-s continuum for non-native English listeners in the L1 German – L2 English study.

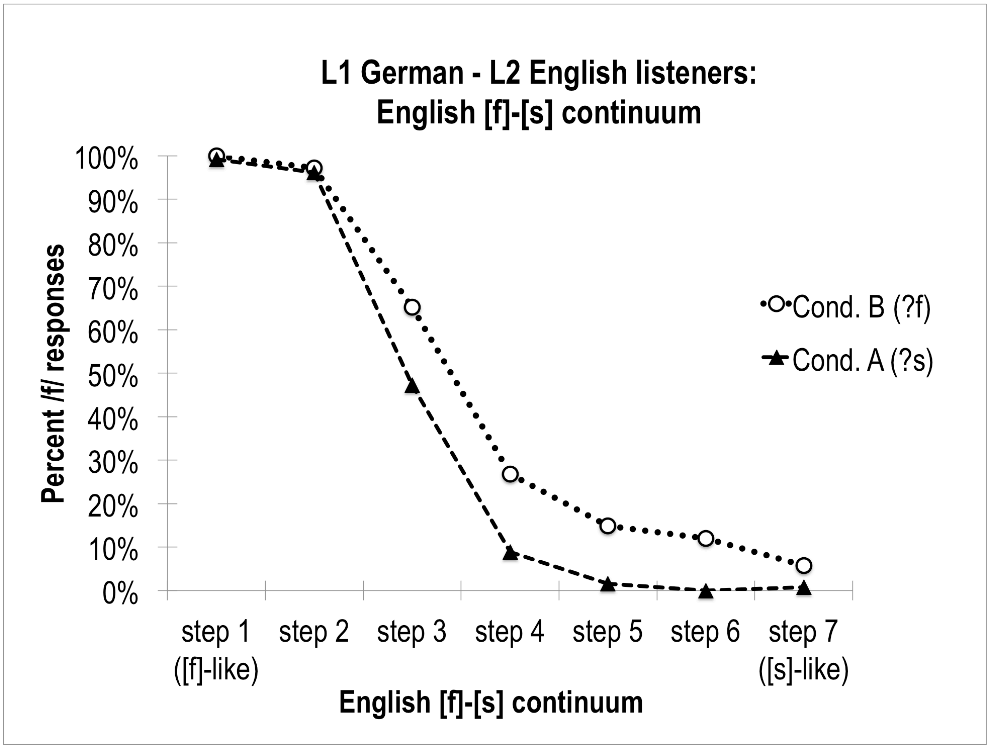

Figure 4.3: Study with L1 German – L2 English listeners (without one low acceptor): Perceptual learning on the English f-s continuum.

Participants also showed a perceptual learning effect on the German f/s continuum, the phoneme contrast of training in an untrained language: Participants in the ambiguous [?f] condition categorized more items on the /f-s/ continuum as the labial consonant /f/ (45.66%) than participants in the ambiguous [?s] condition (39.27%), again a significant simple main effect ($F(1,49)=4.687$, $p=.035$) obtained through a planned pairwise comparison. The difference in the percent /f/ responses between the two conditions amounts to 6.40%. Figure 4.4 below illustrates successful perceptual learning on the German f-s continuum in L1 German- L2 English listeners after listening to a novel, native accent in L2 English.

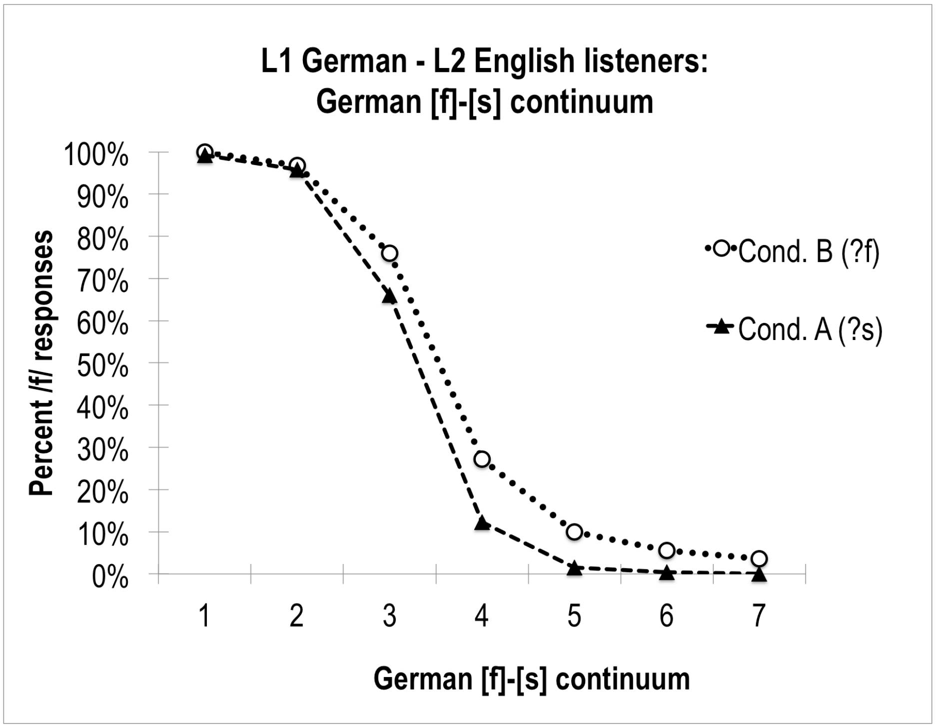

Figure 4.4: Study with L1 German – L2 English listeners (without one low acceptor): Perceptual learning on the German f-s continuum.

To summarize, L1 German – L2 English learners as a group showed a significant perceptual learning effect on the voiceless fricative contrast /f-s/ in English as well as in German. First, this shows that non-native listeners who are not dominant in their L2 and do not live in an L2 environment can perceptually adjust the phoneme boundaries L2 in their L2 when listening to a novel native accent in their L2 – at least for the phoneme contrasts and the bilingual population tested in this study. Second, non-native listeners as a group also generalize the perceptual learning effect across languages to the trained phoneme contrast f/s in their native language.

4.3.3 Results for voiced fricatives (in English and German)

This section discusses whether non-native listeners who have learned to perceptually adapt to unusual pronunciations of voiceless fricatives in their L2 English also generalize this effect to voiced fricatives, both in the language of training, their L2 English, and in the untrained language, their L1 German. Participants did not show a perceptual learning effect on the English v/z continuum, the untrained phoneme contrast in the language of training: Participants in the ambiguous [?f] condition categorized only slightly more items on the /v-z/ continuum as the labial fricative /v/ (52.72%) than participants in the ambiguous [?s] condition (52.04%), not a significant simple main effect of condition at the level of English v/z ($F(1,49)=.039$, $p=.844$), obtained through a planned pairwise comparison. The difference in the percent /v/ responses between the two conditions is merely 0.68%. Figure 4.5 below illustrates the lack of a significant perceptual learning effect on the English /v-z/ continuum in L1 German- L2 English listeners after listening to a novel native [?fs] accent in L2 English.

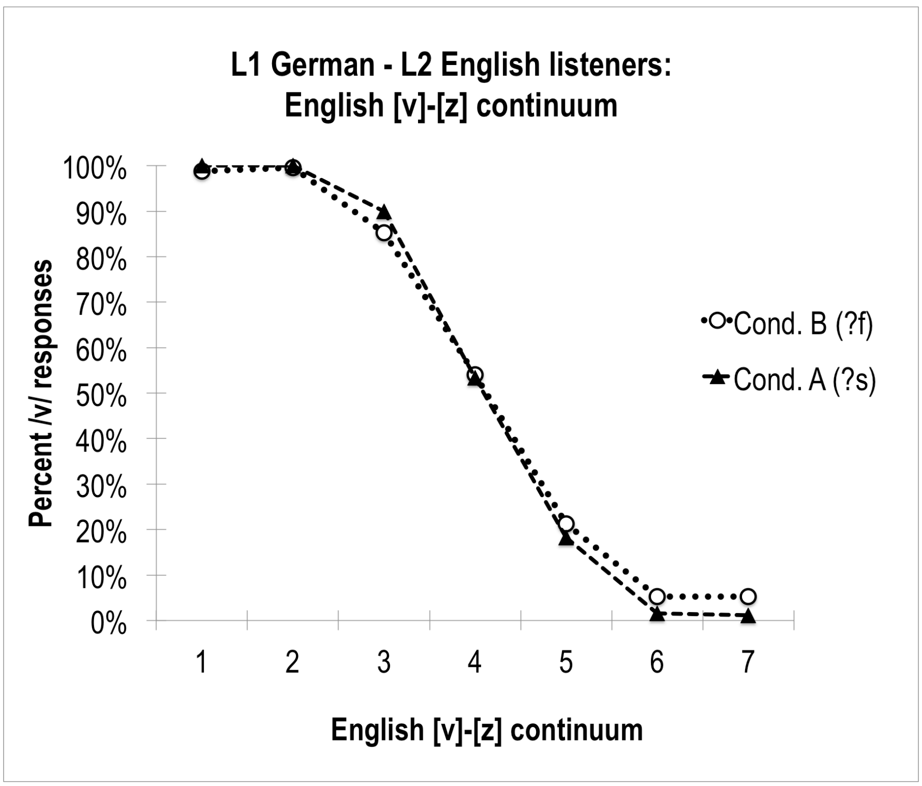

Figure 4.5: Study with L1 German – L2 English listeners (without one low acceptor): No perceptual learning on the English v-z continuum.

Participants also did not show a perceptual learning effect on the German v/z continuum, the untrained phoneme contrast in the untrained language: Participants in the ambiguous [?f] condition categorized slightly fewer items on the /v-z/ continuum as the labial fricative /v/ (32.08%) than participants in the ambiguous [?s] condition (32.86%), not a significant simple main effect of condition at the level of German v/z (F(1,49)=.116, *p*=.734), obtained through a planned pairwise comparison. The difference in the percent /v/ responses between the two conditions is -0.77%, indicating that the small difference is a difference in the unexpected direction. Figure 4.6 below illustrates the lack of a significant perceptual learning effect on the German /v-z/ continuum in L1 German – L2 English listeners after listening to a novel native [?fs] accent in L2 English.

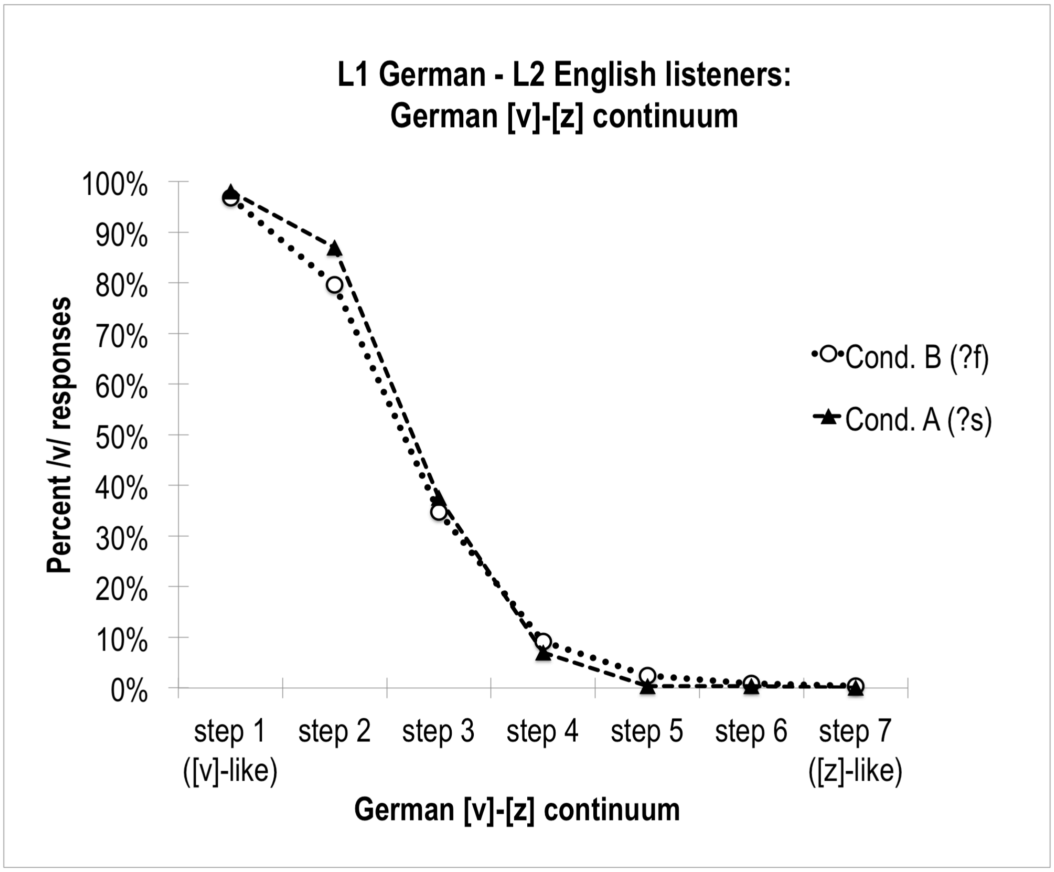

Figure 4.6: Study with L1 German – L2 English listeners (without one low acceptor): No perceptual learning on the German v-z continuum.

In sum, the L1 German – L2 English participants showed perceptual learning effects after having been exposed to a novel, native accent in their non-native language (English). They perceptually adjusted to the unusual f/s contrast in English – the accent to which they were exposed – and generalized this perceptual learning effect cross-linguistically to the f/s contrast in their L1 German. At the same time, however, these non-native English listeners did not generalize the perceptual learning effects to the voiced fricative contrast in either English or German. These findings suggest that non-native learners adjust the representation of phoneme contrasts in their L2 at the level of individual phonemes or phoneme contrasts. They do not generalize these adaptation effects to phoneme contrasts in their L2 or L1 that share the relevant feature of the phoneme contrast, namely place of articulation for the f/s and v/z contrasts.

The contrast between /s/ and /z/ in German is important to the current study, since it allows us to test for generalization of perceptual learning effects from a voiceless fricative contrast to a voiced one. However, many Southern German dialects do not realize /z/ as a voiced fricative [z], and only exhibit [s] in their phonetic inventory. The data collection for the L1 German – L2 English study therefore took place in Berlin, Germany, to avoid recruiting too many participants whose native German dialect does not have the voiced alveolar sibilant [z] as part of their native inventory. The participants in Berlin were asked to produce words like "Sonne" ('sun') or "Hase" ('rabbit'), which are pronounced with [z] in Standard German, but are frequently devoiced in Southern German dialects. Since students in Berlin frequently did not grow up in Berlin, this measure was taken to determine whether or not the participants actually produced /z/ as [z] in their native German dialect.

As a matter of fact, eight participants (four in each condition) did not produce a voiced [z] for underlying /z/ in Standard German. Since this difference in the native phoneme inventory might lead to a difference in the perceptual learning effects, we performed additional analyses for which we excluded the eight participants with a [z]-less dialect. The overall conclusions of these statistical analyses are the same as previously: Participants showed perceptual learning effects on the voiceless fricative contrast f/s in both English and German, but did not show any perceptual learning effects on the v/z contrast in either English or German. Numerically, the effect sizes of the perceptual learning effects on the English f/s contrast and the German f/s contrast were slightly larger when the [z]-less speakers were excluded. For the English f/s contrast, participants in condition A ([?s]) judged 36.36% of the stimuli on the f/s continuum as /f/, while participants in condition B ([?f]) judged 48.53% of the stimuli as /f/, a statistically significant effect size of 12.2% ($F(1,41)=12.676$, $p=.001$). For the German f/s contrast, participants in condition A ([?s]) judged 39.46% of the stimuli as /f/, while participants in condition B ([?f]) judged 46.72% of the stimuli as /f/, as statistically significant effect size of 7.3% ($F(1,41)=4.706$, $p=.036$). At the same time, the English v/z continuum did not lead to a significant perceptual learning effect with an effect size of 0.18% (Condition A [?s]: 53.26% vs. Condition B [?f]: 53.44%; $F(1,41)=.002$, $p=.962$). Similarly, the German v/z continuum did not lead to a significant perceptual learning effect with an effect size of -1.58% (Condition A [?s]: 32.97% vs. Condition B [?f]: 31.39%; $F(1,41)=.389$, $p=.536$). While the statistical results do not differ when the small number of [z]-less speakers is included, this comparison enables the following speculation, which requires additional research with more participants before definite conclusions can be reached. The larger effect sizes for the group of participants who speak a dialect of German that distinguishes /s/ and /z/ phonemes might suggest that hearing a more restricted range of variability for a specific phoneme, and thus by implication for a specific phoneme contrast, makes listeners more adaptive to novel accents. Speakers whose native dialect does not phonemically distinguish /s/ and /z/, on the other hand, show slightly weaker perceptual adaptation effects to unusual pronunciations involving these phonemes, likely because they are exposed to a larger range of variability for at least one of the relevant phonemes. In other words, it is speculated that the slightly weaker perceptual learning effect in listeners with a Southern German accent might be due to the listeners being accustomed to a relatively broad range of pronunciation options for one of the relevant phones of the relevant phoneme contrast.

4.3.4 L2 English language background assessment

4.3.4.1 L2 English perception proficiency test

This section discusses the L2 English Perception Proficiency Test that was conducted with each L1 German – L2 English participant. Each person participated in a brief language perception proficiency test to gauge their perception skills in their non-native language, just like L1 English – L2 German learners participated in a brief L2 German Perception Proficiency Test described in the previous chapter (section XXX). The L2 English participants also completed the perception proficiency test after the main perceptual learning study. Just like in the L2 German study, this L2 perception proficiency test was a same-different task with non-words embedded in white noise.

Each participant heard four blocks, each of which contained twelve similar-sounding non-word pairs. Half of the non-word pairs in each block did not differ phonologically, but were different recordings of the same non-words, e.g. /klæb/-/klæb/ or /klɛb/-/klɛb/ for non-native speakers of English. The other half of the non-word pairs differed in exactly one phoneme and presented a phoneme contrast that does not exist in the listener's native language, at least not in the particular phonetic position, for example /wədæ**z**/-/wədæs/. This is shown in Table 4.4 below. For example, native speakers of German taking the English language perception proficiency test heard /klæb/ - /klɛb/ and judged whether the two non-words were the "same" or "different".

Table 4.4: Same Different Task used for the L2 Language Perception Proficiency Tests (*This table corresponds to Table 3.7, repeated here for ease of reference.)*

<table>
<tr><th colspan="2">Same-Different Task:
Type of stimuli in each of four blocks</th></tr>
<tr><td rowspan="2">Same (6)</td><td>Same – nonwords with sound A (3)</td></tr>
<tr><td>Same – nonwords with sound B (3)</td></tr>
<tr><td rowspan="2">Different (6)</td><td>Different: nonwords with sound A followed by nonwords with sound B (3)</td></tr>
<tr><td>Different: nonwords with sound B followed by nonwords with sound A (3)</td></tr>
</table>

All the German stimuli were recorded by a native speaker of German (the author). For each stimuli pair that the listeners heard, they were instructed to press a button to indicate whether the two non-words they heard contained the "same" or "different" sounds. The ISI (inter-stimulus-interval) between the two stimuli in each non-word pair was set to a large value (1s), to encourage phonological processing rather than more immediate and shallow phonetic-acoustic processing. The English non-word stimuli pairs for the four blocks are shown in Table 4.5 below (This table corresponds to part of Table 3.7, repeated here for ease of reference):

Table 4.5: Non-word stimuli pairs for the L2 English language proficiency tests. Each pair differs in one phoneme, which is marked in bold.

L2 English perception proficiency test stimuli:
/á**w**a/ vs. /á**v**a/
/á**f**a/ vs. /á**θ**a/
/kl**ǽ**b/ vs. /kl**έ**b/
/wədǽ**s**/ vs. /wədǽ**z**/

For each person, we calculated the percent correct responses for all four contrasts. The L2 English listeners answered correctly 77.49% on average (SD 8.22%), i.e. they scored midway between chance and perfect; individual scores ranged from 62.50% to 97.92%. The distribution of the data is presented in Figure 4.7 below.

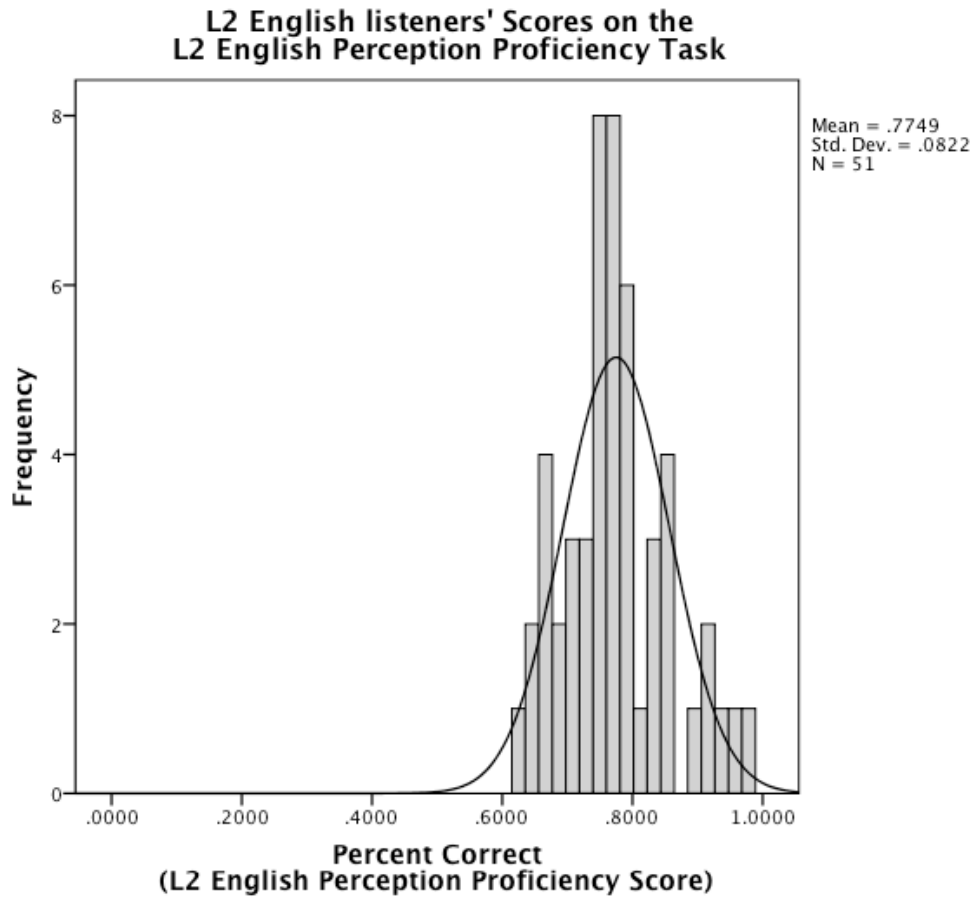

Figure 4.7: Histogram of the L2 English Perception Proficiency Results (51 L1 German – L2 English listeners).

There was some variability in terms of RT: on average, participants took 725.79 ms (SD 219.15); the scatterplot of the individual results is presented in Figure 4.8 below:

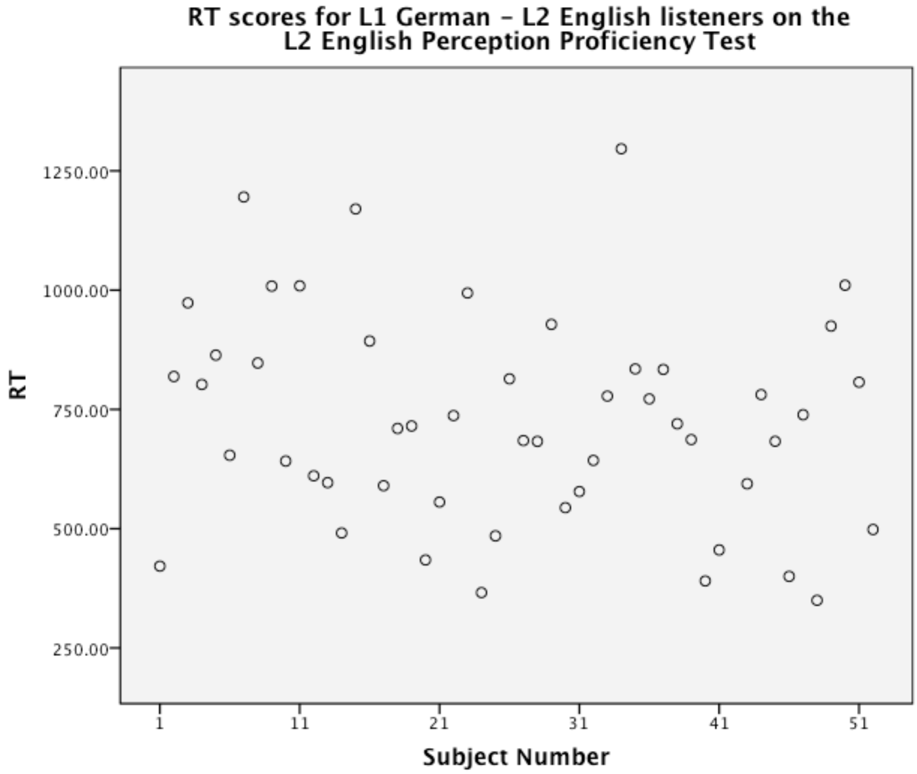

Figure 4.8: Scatterplot of the RT on the L2 English Perception Proficiency Test (51 L1 German – L2 English listeners).

More important than RT are the average scores for each participant on the L2 English perception proficiency test and whether they correlated with the perceptual learning results. Indeed, when dividing the participants into two groups by binning the data into the two most evenly-sized groups based on their scores on the L2 English proficiency test, those participants who scored lower (75% or less correct [<=36/48], N=23) showed a numerically larger perceptual learning effect on the English f/s continuum than those participants who scored higher on the perception test (higher than 75% [37+/48], N=28). Specifically, the group of L2 English learners with higher L2 English scores showed an effect of 7.45%, whereas the group with lower L2 English scores showed a perceptual learning effect of 12.57%, a difference of 5.12%[20] (which is,

[20] When the groups are divided such that "high proficiency" learners could not have more than 10 mistakes on the L2 perception proficiency test (i.e. at least 38/48 points, which is 79.17%), then the group of participants with the higher scores (79.17% or higher, N=20) shows a perceptual learning effect of 12.54%, while the group of participants with lower scores (less than 79.17%, N=31) shows a perceptual learning effect of merely 5.39%, a difference of 7.15% between the two proficiency groups.

however, not significant – see below). The finding that the L2 English speakers with somewhat lower proficiency levels adapted more than the L2 English speakers with higher proficiency levels is in line with the findings in Kim et al. (2011), where L2 speakers with mid-range proficiency levels adapted more than either L2 speakers with high or low proficiency levels. (Note that all of the L1 German – L2 English participants in our study had at least intermediate English proficiency in order to be able to complete the lexical decision task in English in the first phase of the study.) Figure 4.9 below presents the perceptual learning effects on the English f/s continuum split up for the L2 English learners with lower (the two colored lines) and higher (the two black lines) L2 English scores.

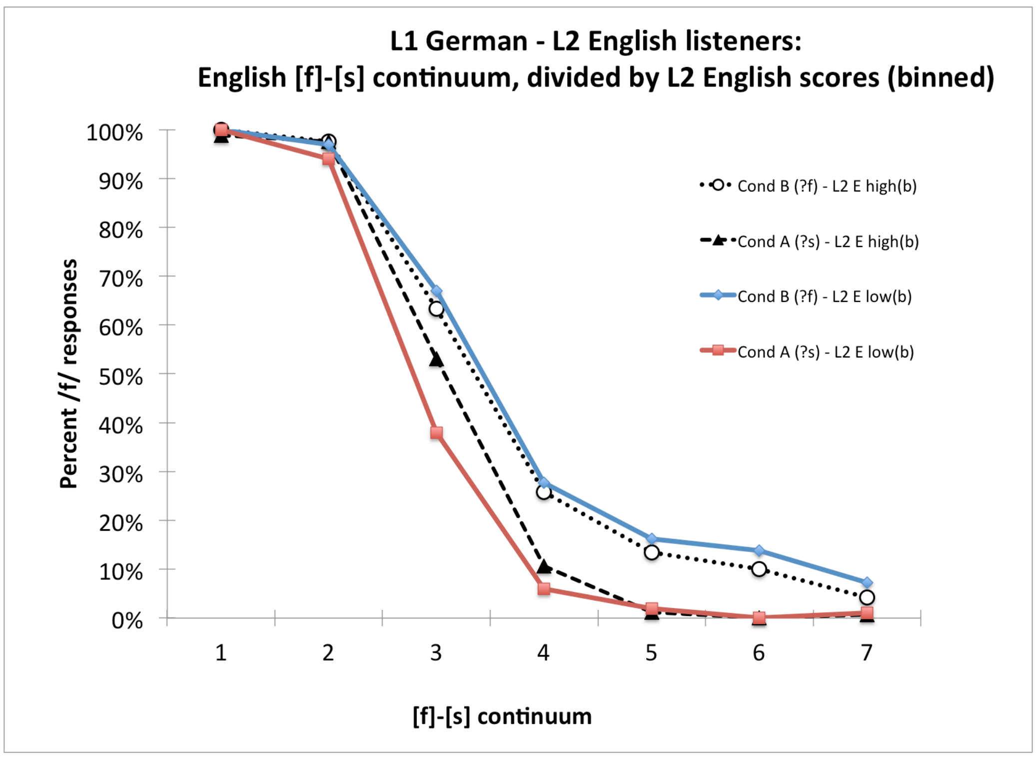

Figure 4.9: Perceptual Learning effect on the English f/s continuum in the L1 German – L2 English study for participants with low vs. high scores on the L2 English German Proficiency Test.

This difference in effect size between the high and low L2 English listeners can also be seen in the following figure. The percent /f/ responses between the two conditions is larger for the low L2 English learners (on the left side) compared to the high L2 English learners (on the right side).

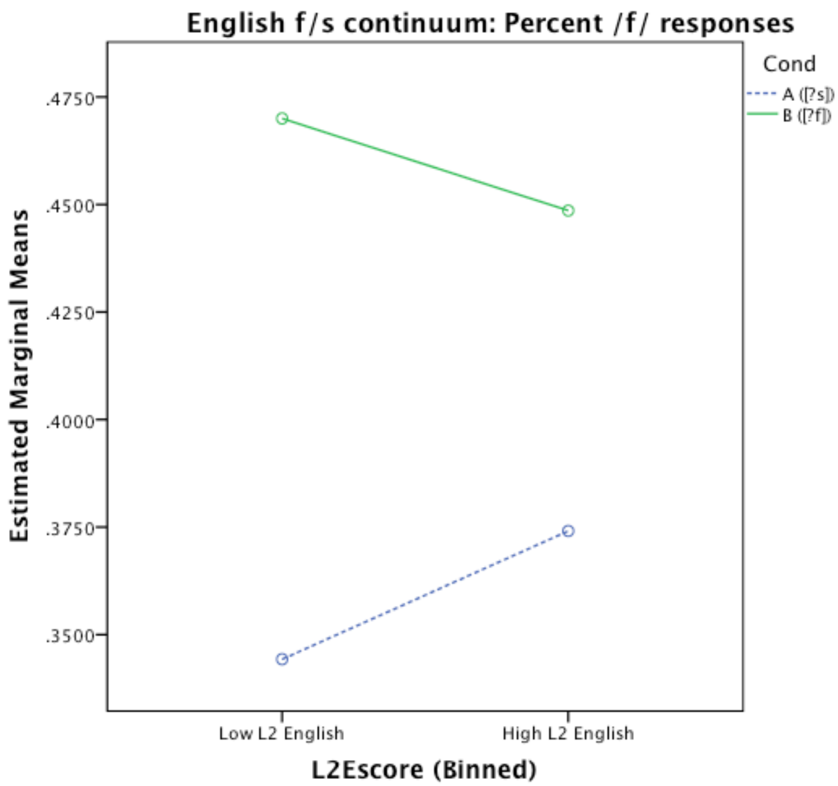

Figure 4.10: Difference in perceptual learning effect size for participants with low vs. high scores on the L2 English perception proficiency test.

In a RM ANCOVA analysis in which the participants' score on the L2 English proficiency test was entered as a continuous or categorical (binned into two groups) between-subject co-variable, there were no statistically significant effects, as shown in 1a) and 1b), respectively, below:

1) English perception proficiency score:

   a. Continuous variable: Condition * English Proficiency Score for:

- all four continua: F(1,47)=.051, *p* = .822 (n.s.)
- English f/s continuum: F(1,47)=1.035, *p* = .314 (n.s.)
- English v/z continuum (F(1,47)=1.552, *p*= .219 (n.s.)
- German f/s continuum: F(1,47)=.046, *p* = .832 (n.s.)
- German v/z continuum: F(1,47)=.355, *p* = .554 (n.s.)

b. Two groups (binned categorical variable): Condition * English Proficiency Score for:

- all four continua: F(1,47)=.454, *p* = .504 (n.s.)
- English f/s continuum: F(1,47)=.624, *p* = .434 (n.s.)
- English v/z continuum (F(1,47)=3.610, *p*= .064 (n.s.)
- German f/s continuum: F(1,47)=1.684, *p* = .201 (n.s.)
- German v/z continuum: F(1,47)=.573, *p* = .453 (n.s.)

However, in a RM ANCOVA analysis with L2 English Perception Proficiency Score (entered as a categorical, between-subject co-variable) within just the group of participants with lower L2 English scores, there was a statistically significant effect of condition at the level of the English f/s phoneme contrast (F(1,47)=6.848, *p* = .012). The equivalent RM ANCOVA analysis with L2 English Perception Proficiency Score within just the group of participants with higher L2 English scores did not lead to a statistically significant effect (F(1,47)=2.916, *p* =.094), nor did the RM ANCOVA analyses at any of the other phoneme contrasts (English v/z, German f/s, German v/z). This shows that we have evidence that there is a statistically significant perceptual learning effect on the English f/s continuum for the group of participants who scored lower on the L2 English perception proficiency test, but not for the group of participants who scored higher on the L2 English perception proficiency test.

4.3.4.2 L1 German- L2 English learners' further language background

As discussed at the beginning of this chapter, in this study with non-native listeners, their knowledge of the crucial L2 English words – or lack thereof – and their ability to accept unusual pronunciations of these words is likely to have an influence on whether or not the listeners show perceptual learning effects. First, we tested whether the number of accepted ambiguous critical stimuli correlated with perceptual learning effects in non-native speakers. Indeed, a RM ANCOVA analysis revealed a simple effect at the level of the f/s contrast in English through a statistically significant interaction of condition and the acceptance of ambiguous stimuli (entered as a continuous variable) ($F(1,47)=4.354$, $p=.042$). As Figure 4.11 below shows, this indicates that higher acceptance rates of the ambiguous critical stimuli – online, during the study – is correlated with higher perceptual learning effects: As the number of accepted ambiguous stimuli increases, the percent /f/ responses increases for condition B ([?f]) and somewhat decreases for condition A ([?s]). In other words, when the number of accepted stimuli is smaller, the percentages of /f/ responses for both conditions are close together (both around 40%), while the percent /f/ responses are further apart when the number of accepted stimuli is higher.

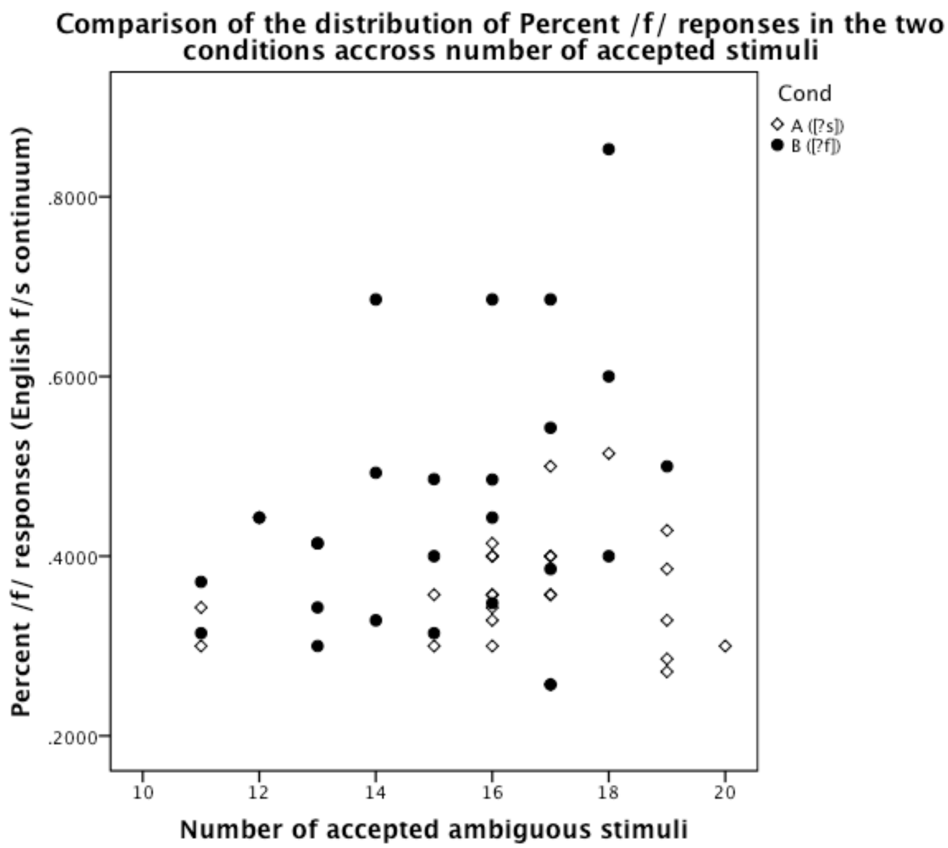

Figure 4.11: Correlation between number of accepted ambiguous stimuli and Percent /f/ responses on the English f/s continuum for the two conditions (L1 German – L2 English study).

The Figure above captures the notion that the perceptual learning effect was bigger when people knew and accepted more of the ambiguous, critical stimuli during the online lexical decision task. This type of result was predicted, since listeners cannot be expected to learn from words with which they are not familiar. This positive outcome of the covariate analysis of the number of accepted ambiguous stimuli and the perceptual learning effect size was found only for the f/s continuum in English, as shown in 2) below:

2) Number of Accepted Ambiguous Stimuli: Condition * Accepted Ambiguous Stimuli interaction for:
   - all four continua: $F(1,47)=1.142$, $p=.291$ (n.s.)
   - English f/s continuum: $F(1,47)=4.354$, $p=.042$*
   - English v/z continuum: $F(1,47)=.003$, $p=.954$ (n.s.)
   - German f/s continuum: $F(1,47)=.108$, $p=.744$ (n.s.)
   - German v/z continuum: $F(1,47)=.155$, $p=.695$ (n.s.)

Further, no significant effect was found between any of the four continua (English f/s, English v/z, German f/s, German v/z) and the score of the vocabulary questionnaire – an offline measure of word knowledge – based on all forty critical stimuli. The relevant statistics are shown in 3) below:

3) Score on Vocabulary Questionnaire: Condition * Vocabulary Score interaction for:
   - all four continua: $F(1,47)=.314$, $p=.578$ (n.s.)
   - English f/s continuum: $F(1,47)=.002$, $p=.961$ (n.s.)
   - English v/z continuum: $F(1,47)=.056$, $p=.814$ (n.s.)
   - German f/s continuum: $F(1,47)=.284$, $p=.596$ (n.s.)
   - German v/z continuum: $F(1,47)=2.925$, $p=.094$ (n.s.)

Just like in the study with L1 English – L2 German participants, the additional language background information collected through a questionnaire and the data from the language perception proficiency test made it possible to test for correlations between the size of the perceptual learning effects and specific aspects of participants' language background by means of ANCOVA analyses (Analysis Of Co-Variance). We were again able to analyze the following variables related to language background based on the questionnaires: 4) Status in childhood as

L1 German monolingual or bilingual (in German and another language in childhood); 5) Hours of English used per week; 6) Self-assessed proficiency in English; 7) Age of acquisition of L2 English; 8) Time spent in an English-speaking country; 9) Time spent in English-speaking situations; 10) Years of L2 English study; 11) Years of L2 English experience; (and 12) Current age). However, none of the additional ANCOVA analyses for the L1 German – L2 English study provided any statistically significant effects. Details of the statistical results and distributions of the participants' results on the individual language background variables can be found below:

4) Monolingual vs. bilingual status as a child: Condition * MonoBil interaction for:
   - all four continua: $F(1,47)=.330$, $p=.568$ (n.s.)
   - English f/s continuum: $F(1,47)=.054$, $p=.817$ (n.s.)
   - English v/z continuum: $F(1,47)=.105$, $p=.747$ (n.s.)
   - German f/s continuum: $F(1,47)=1.176$, $p=.284$ (n.s.)
   - German v/z continuum: $F(1,47)=.490$, $p=.487$ (n.s.)

5) Hours of English used per week: Condition * Hours of English interaction for:
   - all four continua: $F(1,47)=1.896$, $p=.175$ (n.s.)
   - English f/s continuum: $F(1,47)=1.049$, $p=.311$ (n.s.)
   - English v/z continuum: $F(1,47)=.822$, $p=.369$ (n.s.)
   - German f/s continuum: $F(1,47)=2.240$, $p=.141$ (n.s.)
   - German v/z continuum: $F(1,47)=.000$, $p=.994$ (n.s.)

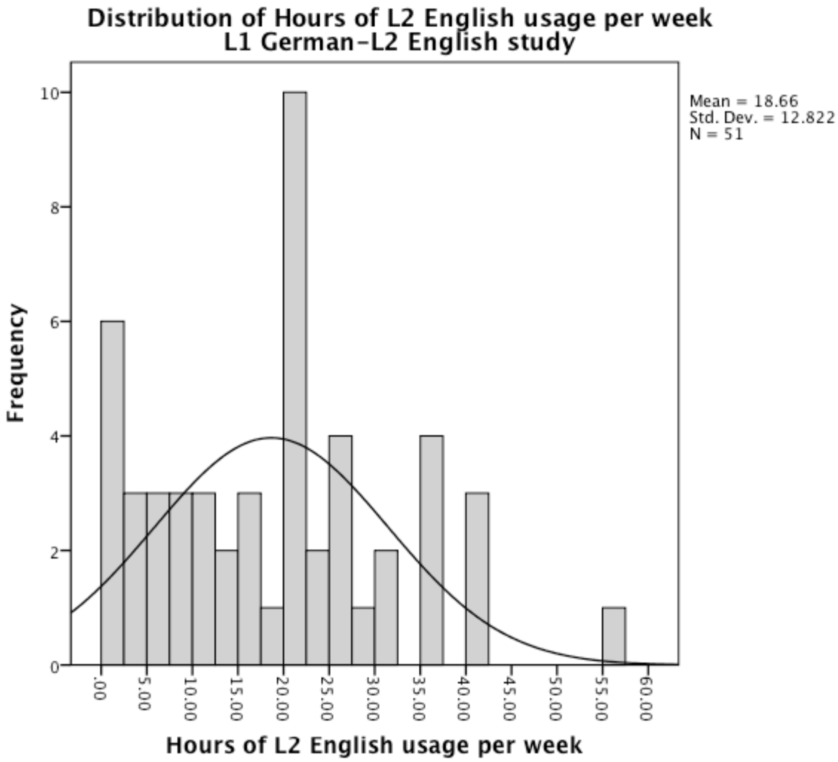

Figure 4.12: Histogram of the Hours of L2 English used by L1 German – L2 English participants.

6) Self-assessment proficiency score (in L2 English): Condition * L2 English self-assessment score interaction for:

- all four continua: F(1,47)=.274, *p*=.603 (n.s.)
- English f/s continuum: F(1,47)=.143, *p*=.707 (n.s.)
- English v/z continuum: F(1,47)=.484, *p*=.490 (n.s.)
- German f/s continuum: F(1,47)=1.928, *p*=.172 (n.s.)
- German v/z continuum: F(1,47)=.248, *p*=.621 (n.s.)

Note: the histogram in Figure 4.13 below is based on numbers where participants' self-assessed proficiency scores were transformed into the following numerical values: 1= Beginner, 1.5=Beginner-to-Intermediate; 2= Intermediate, 2.5=Intermediate-to-Advanced; 3=Advanced, 3.5=Advanced-to-Native Speaker; 4=Native Speaker.

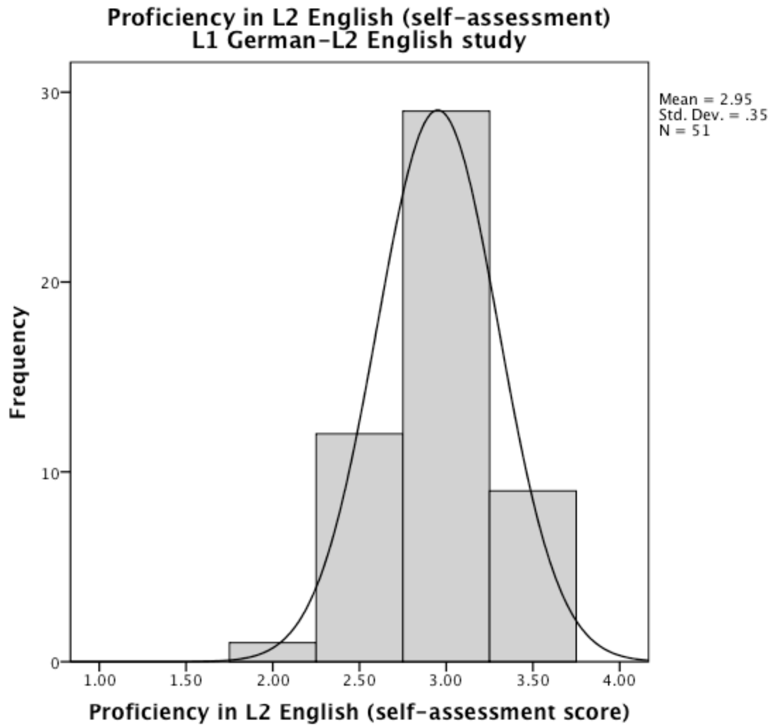

Figure 4.13: Histogram of the distribution of L1 German – L2 English participants' self-assessed L2 English proficiency level

7) Age of Acquisition of L2 English: Condition * Age of Acquisition interaction for:

- all four continua: F(1,47)=.952, *p*=.334 (n.s.)
- English f/s continuum: F(1,47)=2.935, *p*=.093 (n.s.)
- English v/z continuum: F(1,47)=.079, *p*=.779 (n.s.)
- German f/s continuum: F(1,47)=1.684, *p*=.201 (n.s.)
- German v/z continuum: F(1,47)=1.206, *p*=.278 (n.s.)

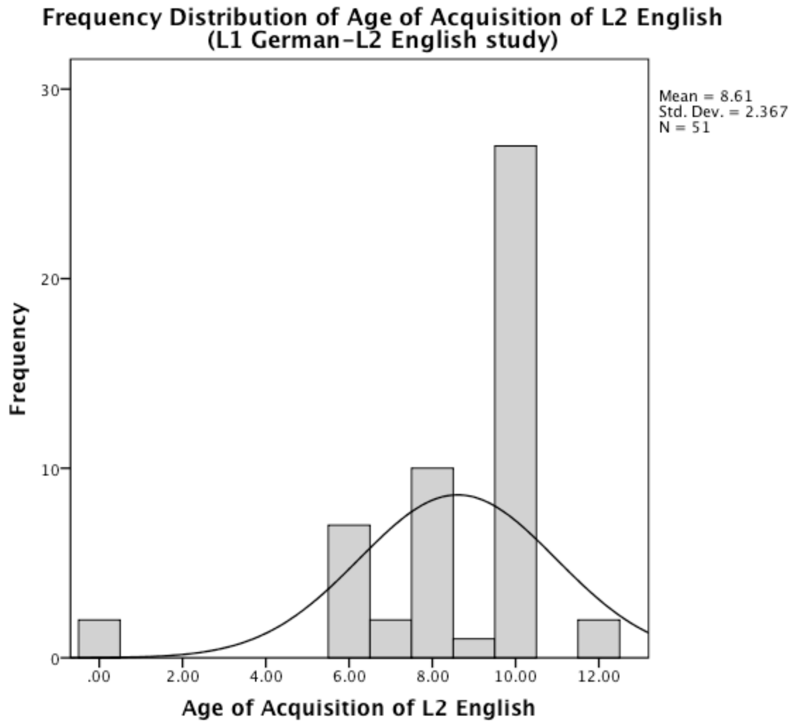

Figure 4.14: Histogram of the distribution of Age of Acquisition of L2 English in the L1 German – L2 English participants.

8) Time in English-Speaking Country: Condition * Time L2 Country interaction for:

- all four continua: $F(1,47)=1.864$, $p=.179$ (n.s.)
- English f/s continuum: $F(1,47)=.406$, $p=.527$ (n.s.)
- English v/z continuum: $F(1,47)=.178$, $p=.675$ (n.s.)
- German f/s continuum: $F(1,47)=1.436$, $p=.237$ (n.s.)
- German v/z continuum: $F(1,47)=2.312$, $p=.135$ (n.s.)

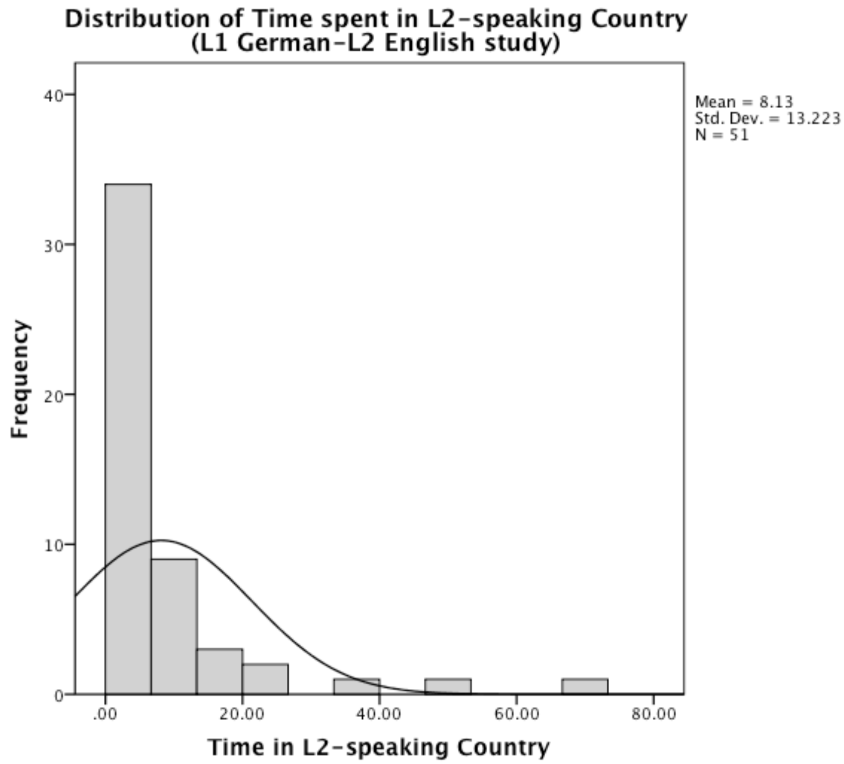

Figure 4.15: Histogram of the number of months L1 German – L2 English study participants spent in an English-speaking country.

9) Time in English-Speaking Situations: Condition * L2 English situation interaction for:

- all four continua: F(1,47)=2.050, *p*=.159 (n.s.)
- English f/s continuum: F(1,47)=.766, *p*=.386 (n.s.)
- English v/z continuum: F(1,47)=.254, *p*=.617 (n.s.)
- German f/s continuum: F(1,47)=1.362, *p*=.249 (n.s.)
- German v/z continuum: F(1,47)=1.806, *p*=.185 (n.s.)

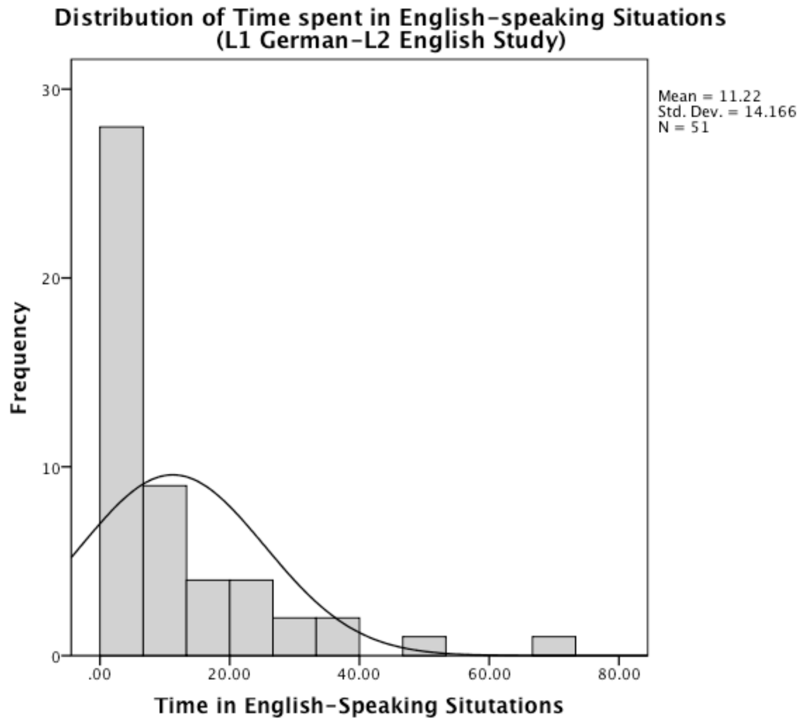

Figure 4.16: Histogram of the number of months L1 German – L2 English study participants spent involved in English-speaking situations.[21]

10) Years of L2 English Study: Condition * Years L2 English study interaction for:

- all four continua: $F(1,47)=.000$, $p=.987$ (n.s.)
- English f/s continuum: $F(1,47)=.019$, $p=.891$ (n.s.)
- English v/z continuum: $F(1,47)=.756$, $p=.389$ (n.s.)
- German f/s continuum: $F(1,47)=.000$, $p=.987$ (n.s.)
- German v/z continuum: $F(1,47)=2.468$, $p=.123$ (n.s.)

[21] "Time involved in an English-speaking situation" does not necessarily mean time spent in an English-speaking country, but could refer to, for example, working in an English-speaking hostel, or participating in an English-speaking drama club.

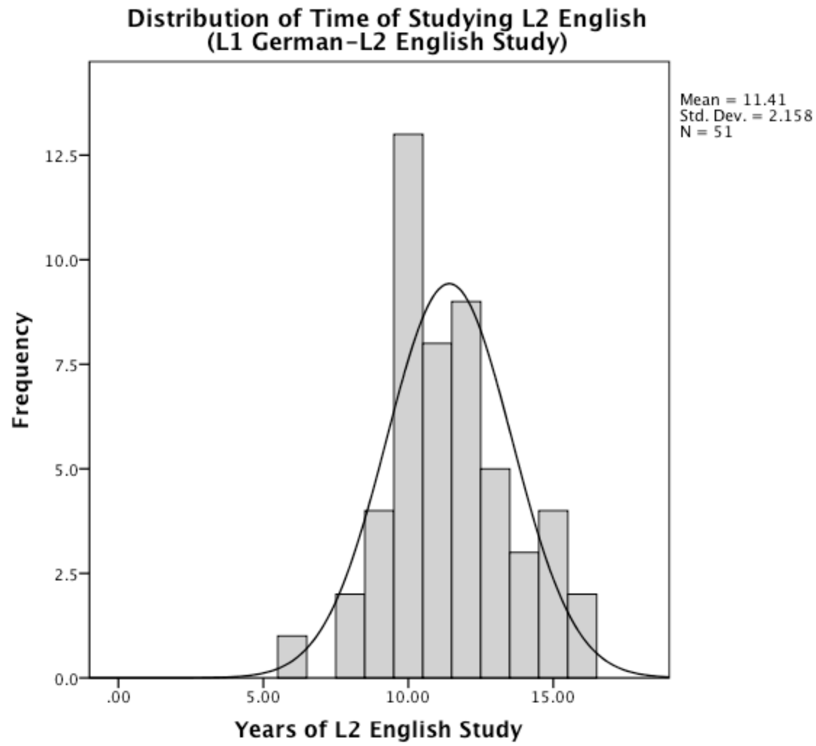

Figure 4.17: Histogram of the number of years L1 German – L2 English study participants studied English.

11) Years of L2 English Experience: Condition * Years L2 English experience interaction for:

- all four continua: $F(1,47)=.029$, $p=.867$ (n.s.)
- English f/s continuum: $F(1,47)=.105$, $p=.748$ (n.s.)
- English v/z continuum: $F(1,47)=1.304$, $p=.259$ (n.s.)
- German f/s continuum: $F(1,47)=.016$, $p=.900$ (n.s.)
- German v/z continuum: $F(1,47)=.270$, $p=.606$ (n.s.)

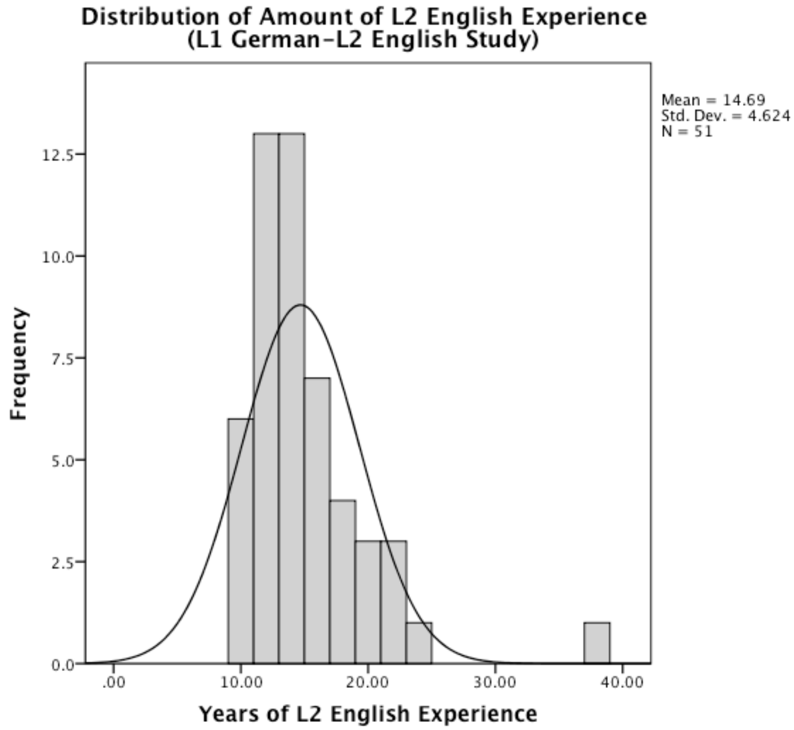

Figure 4.18: Histogram of the number of years L1 German – L2 English study participants had experience using English (whether or not they formally studied English at the time).

12) (Current age): Condition * Current Age interaction for:

- all four continua: F(1,47)=.113, *p*=.739 (n.s.)
- English f/s continuum: F(1,47)=.705, *p*=.405 (n.s.)
- English v/z continuum: F(1,47)=3.041, *p*=.088 (n.s.)
- German f/s continuum: F(1,47)=.098, *p*=.755 (n.s.)
- German v/z continuum: F(1,47)=.018, *p*=.894 (n.s.)

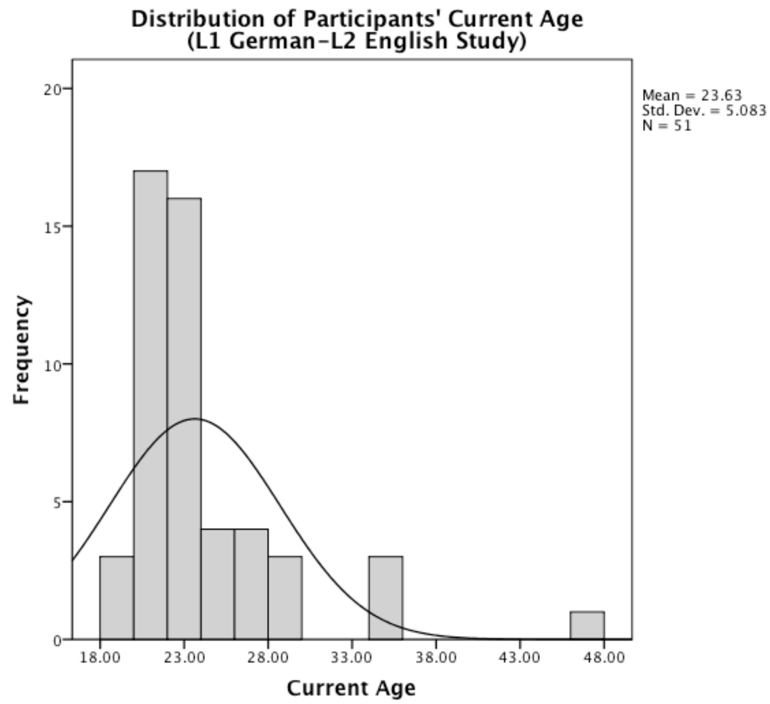

Figure 4.19: Histogram of the distribution of the age of L1 German – L2 English participants (at the time of the study).

In sum, the L2 English language background of the L1 German – L2 English study participants was assessed on a dozen factors, ranging from age of acquisition of L2 English, time spent in an English-speaking environment, to self-rated proficiency assessments, and an L2 English perception proficiency test. While there might not have been a large enough variability in the sample for some of the variables to function as useful predictors – for example for the variable age of acquisition – we have evidence for the relevance of at least two aspects of the participants' language background for perceptual learning results. Those L2 English learners with a lower proficiency score on the L2 English perception proficiency test show a stronger perceptual learning effect size on the f/s continuum in English than L2 English learners with a higher L2 English proficiency score. Moreover, bigger perceptual learning effect sizes on the English f/s continuum correlated with larger numbers of accepted ambiguous stimuli during the lexical decision task. Thus, larger perceptual learning effects in the trained English f/s contrast were correlated with those non-native listeners who had more difficulties identifying and distinguishing L2 phoneme contrasts embedded in white noise in a Same/Different task, and with participants who accepted more unusual pronunciations of the critical stimuli. This suggests that perceptual learning in a non-native language might show the strongest effects in L2 learners who have not yet mastered non-native phoneme contrasts (in non-words and in adverse listening conditions) and who are willing to accept novel pronunciations of a specific phoneme when listening to their L2.

4.4 Discussion and conclusion

This L1 German – L2 English study was conducted to test whether non-native listeners could perceptually retune phoneme categories in their L2 just like native listeners retune phoneme categories in their L1, to test whether non-native listeners would generalize perceptual learning effects cross-linguistically from their L2 to their L1 ("regressive transfer"), and to test whether non-native listeners adjust their phoneme categories at the level of individual phonemes and phoneme contrasts or at an abstract, phonological featural level.

The perceptual learning effects for this L1 German – L2 English study are summarized in Table 4.6. Overall, the participants showed statistically significant perceptual learning effects on the f/s contrast in both English and German, but did not show any perceptual learning effects on the v/z contrast in either English or German.

Table 4.6: Summary Table of the L1 German – L2 English study, highlighting which of the continua show perceptual learning effects.

[Shaded cells with bold text indicate statistically significant perceptual learning effects.]

	English f/s	**English v/z**	**German f/s**	**German v/z**
L1 German – L2 English Study	*** sign. effect (*p*=.003)** **effect size: 9.7%**	n.s. (*p*=.844) effect size: **0.7**%	*** sign. effect (*p*=.035);** **effect size: 6.4%**	n.s. (*p*= .734) effect size: **-0.8**%

I will now interpret these results in light of the questions that the study set out to address, as described in the introduction to this chapter. The L1 German – L2 English participants perceptually adjusted their f/s phoneme contrast in English after listening to a novel, native English accent with either ambiguous /f/ or ambiguous /s/ phonemes. These results show that non-native listeners can perceptually retune phoneme contrasts in their L2, at least for phoneme contrasts that exist in both their L2 and their L1. Moreover, the group of participants with lower scores on a Same/Different L2 phoneme perception task had a slightly larger effect size for the perceptual learning on the trained English f/s contrast. At the same time, the size of perceptual learning effects correlated with the number of accepted ambiguous critical stimuli during the Lexical Decision Task. These findings makes sense in light of other research that has argued that L2 grammars, or interlanguages, are more "in flux" and not stable (e.g. Broselow, Chen & Wang 1998:274,277). Yet, these results do not seem to corroborate Sebastian-Galles et al.'s (2006:1288) conclusion that L1 phoneme representations are generally more malleable than L2 representations. However, this apparent contradiction stems from a difference in the type of L2 representations discussed in the L1 German – L2 English study presented here, and in Sebastian-Galles et al.'s (2006) ERP study, which measured event-related brain potentials. In Sebastian-Galles et al. (2006), Spanish-Catalan bilinguals had more difficulties than L1 Catalan speakers in rejecting Catalan non-words that differed from existing Catalan words by one phoneme contrast which is known to be very difficult for L1 Spanish listeners to distinguish. Catalan-dominant bilinguals, on the other hand, were able to correctly reject the Catalan non-words with the Catalan-unique phonological contrast. They did, however, show an N400 effect for both real Catalan words and Catalan non-words with the one changed phoneme, presumably because Catalan-dominant bilinguals live in an environment with many L1 Spanish-L2 Catalan speakers and have a lexical representation for the Spanish-accented Catalan words ("phonological variants", Sebastian-Galles et al. 2006:1289) that do not distinguish the relevant vowel contrast. Sebastian-Galles et al.'s (2006) claim about the reduced plasticity in L2 representations can still be held with regard to non-native L2 phoneme contrasts. However, the results of the current study shows that phoneme contrasts that L2 listeners also distinguish in their L1, such as the f/s contrast in English and German, can dynamically adjust to new information, such as a novel accent.

Further, the L1 German – L2 English participants also perceptually adjusted their f/s phoneme contrast in their L1 German after listening to a novel, native English accent with either ambiguous /f/ or ambiguous /s/ phonemes. These results show that non-native listeners can generalize their perceptual learning effects cross-linguistically from their non-native to their native language. There is, a priori, no particular theoretical reason to expect that phonetic retuning of phoneme contrasts in one language would have to affect both languages in bilingual listeners. For example, Maye et al. (2008) have shown that listeners' adaptation to an artificial accent with lowered front vowels resulted in listeners accepting both unaccented as well as accented phonetic forms (although only accents with *lowered* vowels, not with *raised* vowels). Maye and colleagues (2008:556) conclude: "our finding […] suggests that it is possible to develop and maintain more than one mapping (e.g. bilingual, bi-dialectal, native- vs. foreign-accented etc.), thus suggesting that adaptation effects could be specific to a particular dialect, accent, or language and not affect other dialects, accents, or languages that the speaker-listener uses." Most importantly, the L1 English – L2 German studies from the previous chapter have already illustrated that the generalization of perceptual learning-effects across languages from L1 to L2 seems to be correlated with recency of L2 use and how bilingual a participant's language mode is.

The SLA literature has typically focused on the effects of the first language onto the second and has only more recently started to focus on the testable influence of the L2 onto the L1, including but not limited to attrition effects in L1. Cook (2003) already points out that VOT production values can exhibit a strong influence of a speaker's L2, although it might not be noticeable without phonetic training or tools. Cook (2003:12) describes "differences in the first language of L2 users for plosive consonants such as /p/ and /b/ or /k/ and /g/ across pairs of languages such as Spanish/English (Zampini & Green, 2001), French/English (Flege, 1987), and Hebrew/English (Obler, 1982), which are essentially undetectable in normal language use." In other words, the VOT values in bilinguals' L1 production are often found to be influenced by the speakers' L2 use, even for languages that employ quite different ranges of the VOT continuum (such as English vs. Spanish, French).

Jarvis (2003:81f) also argues that the L2 can influence the L1, and that a bilingual's languages are dynamically interlinked. Jarvis (2003) comes to the conclusion that language knowledge and representations in second language learners are dynamic and interconnected in nature, much like the analysis provided for the results of the L1 English – L2 German study in the previous chapter: "[W]e conclude that the end state of second language acquisition, if there is one, is probably best characterized as a dynamic and partially integrated multicompetence of all languages that the learner knows, rather than as separate, rigid, steady-state grammars for the L1 and L2(s) (e.g. Cook (1991, 1992); Jarvis (1998); Pavlenko (1999)).”

While these researchers have provided ample evidence for L2-to-L1 effects, it has also been found (e.g. Haigh & Jared 2007) that the influence of the L2 onto the L1 is not as strong as the influence of the L1 onto the L2. In the current study, the perceptual learning effect on the listeners’ trained L2 English f/s contrast is numerically slightly larger (9.7%) than the perceptual learning effect on the listeners’ untrained L1 German f/s contrast (6.4%). As Table 4.7 below recaps from the previous chapter, in the second L1 English – L2 German study, the perceptual learning effect on the listeners’ trained L1 English f/s contrast is numerically smaller (10.8%) than both the perceptual learning effect on the listeners’ untrained L2 German f/s contrast (15.1%) and the untrained L2 German v/z contrast (20.5%). To the extend that L2 sounds are presumably less stable and fixed than L1 sounds, bigger shifts might be expected for perceptual learning for L2 sounds.

Table 4.7: Summary Table of the L1 German – L2 English study compared to the second L1 English – L2 German study with cross-linguistic perceptual learning effects (Chapter 3).

[Shaded cells with bold text indicate statistically significant perceptual learning effects]

	English f/s	**English v/z**	**German f/s**	**German v/z**
L1 German – L2 English Study	*** sign. effect (*p*=.003) effect size: 9.7%**	n.s. (*p*=.844) effect size: **0.7**%	*** sign. effect (*p*=.035); effect size: 6.4%**	n.s. (*p*= .734) effect size: **-0.8**%
L1 English – L2 German Study #2	*** sign. effect (*p*=.009) effect size: 10.8%**	n.s. (*p*=.309) effect size: **4.9%**	*** sign. effect (*p*=.010); effect size: 15.1%**	*** sign. effect: *p*<.001 effect size: 20.5%**

Although the L1 German – L2 English listeners in this study adjusted their f/s phoneme contrasts in both English and German, they did not adjust the phoneme contrasts for the voiced fricative contrasts v/z in German or English, which are distinguished by the same place of articulation features. These results suggest that perceptual learning in a non-native language appears to take place at the level of individual phonemes and phoneme contrasts, rather than at an abstract level of features that distinguishes other phoneme contrasts as well. Some other research in the area of second language acquisition and adult vs. infant learning studies show similar restrictions to specific phonemes when adult and/or second language learners are implicitly or explicitly trained on novel sounds or sound contrast. Nishi & Kewley-Port (2007) have shown that L1 Japanese – L2 English listeners do not generalize training from difficult English vowels to untrained English vowels. They further review other studies in which L2 English listeners did not generalize training on English vowels to untrained vowels (Akahane-Yamada et al. 1997; Sperbeck et al. 2005; cited in Nishi & Kewley-Port 2007:1497).

Maye and colleagues have found differences in how learning and generalization seem to take place in infant vs. adult listeners. Maye and colleagues (Maye & Weiss 2003; Maye et al. 2008) found that infants exposed to a voicing contrast at one place of articulation by means of a bimodal distribution learned to discriminate the voicing contrast not only at this place of articulation but also generalized this learning to a different place of articulation (experiment 2). This finding stands in contrast to a study with adults (Maye & Gerken 2001) in which adults do not generalize learning to an untrained place of articulation (in a slightly different methodological set-up) (Maye et al. 2008:129, Maye & Weiss 2003). Maye et al. (2008:130) suggest that the likely conclusion is that "[i]nfants appear to extract the featural properties of the input speech, while adult learning may be restricted to the segmental level."

Peperkamp & Dupoux (2007) also found that adult learners in an artificial language experiment did not generalize learning from trained consonants to untrained consonants. In their study, Peperkamp & Dupoux (2007, study 1) trained native French speakers on an artificial grammar in which either stops or fricatives showed allophonic voicing word-medially, whereas French has phonemic voicing in obstruents in all positions. After a training phase for allophonic voicing on two places of articulation (on either fricatives or stops), the test phase included the trained places and an additional untrained place of articulation. Peperkamp & Dupoux (2007:16) concluded that the adults in their artificial grammar study learned the allophony rule on a "segment-by-segment basis. This is a bit surprising, as the phonological system of natural languages is typically organized around natural classes."

A possible explanation for the phoneme contrast-specific learning in non-native listeners is that the acquisition, representation, and especially the retuning of phoneme boundaries by adults in an artificial language environment or in a second language is not as likely to generalize to untrained phonemes compared to adaptations in a listener's native language when they are infants. One exception to this (provisional) generalization seems to be when (adult) listeners are exposed to "salutatory alternations" (White 2014). Here, listeners hear phonological alternations between two sounds, e.g. [t] and [ð], although their language inventory includes a sound, [d], which "is more similar to each of the two alternating sounds than the two alternating sounds are to each other" (White 2014:98). The study finds that listeners generalized the learning pattern from dissimilar sounds to more similar sounds – "which had been 'leaped over' (White, 2014)" – but listeners rarely generalized the learning pattern in the opposite direction, from similar sounds to more dissimilar sounds.

The results of the study presented in this chapter appear to be a case where second language learning and fine-tuning effects are segment-specific and dependent on positive evidence. The adult, non-native listeners do not generalize adjustments to the interpretation of place of articulation in voiceless fricatives to other members of the natural class – i.e. voiced fricatives – without positive evidence that other members of this natural class are also affected. The findings of the study reported here, combined with other research on second language learning and artificial language learning in adults vs. infants, suggest that adult/L2 language learners assume the most conservative and restrictive grammar that can account for the patterns discovered in their language input. In other words, based on positive evidence in their language input, adult, non-native listeners postulate grammatical rules that affect the fewest possible elements (phones), rather than grammatical rules that affect the entire natural class of the specific sounds that exhibited a new pattern.

These findings suggest that learning of new phoneme contrasts or learning in an L2-like or artificial language-type environment can be specific to segments, while perceptual learning in L1 in adults has been shown to generalize across features to untrained phoneme contrasts (Chapter 2: English-Only Study, Chapter 3: L1 English – L2 Germany Study #2; Kraljic & Samuel 2006). Although more research needs to be done to further explain this finding, one possible explanation might be that L2 representations are more 'explicit' – as opposed to 'implicit' – in nature, whereby the terms 'explicit' and 'implicit' seem to refer to the level of detail and abstractness contained within the representations. Bialystok (2001), for example, has proposed that bilingual children use more explicit language representations compared to monolingual children who use more implicit language representations (cf. Murphy & Pine

2003:142ff). Another possibility might be that second language learners rely less on phonological features or phonetic cues but are rather driven by orthographic representations (cf. Peperkamp & Dupoux 2007:17). The Roman script used for English (and German) likely reinforces the idea that individual sounds, which correspond to graphemes and phonemes, are the building blocks of language, and might reduce the saliency of sub-phonemic features for second language learners, whose language processing is known to generally be less automatic.

To conclude, the L1 German – L2 English participants showed perceptual learning effects on the trained English f/s contrast as well as on the untrained German f/s contrast, but did not generalize retuning effects across the feature voice to either the English v/z contrast or the German v/z contrast. The results of this L1 German – L2 English study suggest that non-native listeners can adjust their phoneme boundaries to adapt to systematic variation in their L2 English input, presumably because the L2 representations are still malleable and 'in flux', especially for less proficient L2 listeners. Moreover, non-native listeners who adapt to a novel, native accent in their L2 can generalize these perceptual learning effects from their L2 to their L1 ('regressive transfer'), although the cross-linguistic effects are not as strong as the L1-to-L2 cross-linguistic effects in the second L1 English – L2 German study of the previous chapter.

Taken together, the experiments in this and the previous chapter support a model of sound representations in bilinguals in which the phonemes of the two languages are dynamically associated with each other and can influence each other in a bi-directional way if both languages are sufficiently activated. Finally, the current study did not provide evidence that non-native listeners generalize perceptual learning effects on the place contrast in voiceless fricatives (the f/s contrast) across voice to the voiced fricatives (the v/z contrast) in either English or German. This finding suggests that perceptual adaptation effects in non-native listeners do not seem to operate at an abstract, featural level as has been proposed for native perceptual learning effects (Kraljic & Samuel 2006, cf. also the English-only study reported in Chapter 2). Rather, perceptual retuning in non-native listeners seem to take place at the level of individual segments and phoneme contrasts, and non-native listeners do not generalize these adaptation effects to similar phoneme contrasts that share relevant features in either their L2 or L1. Thus, the phonetic-phonological processes involved in adapting to novel, native accents in a non-native language are generally similar in character to the processes involved in adapting to accents in a listener’s native language. Yet, perceptual retuning in a listener’s non-native language also differs from perceptual retuning in a listener’s native language in some important specifics.

5. Discussion and Conclusion

5.1 Summary of Perceptual Learning Studies

This dissertation set out to test perceptual learning in monolingual, novice and advanced second-language learners to determine how diverse listener populations perceptually adjust to a novel accent when listening to (native) English. In all studies, listeners participated in a 'classic' perceptual learning paradigm (Norris et al. 2003) in which they first completed a lexical decision task during which they were exposed to a novel accent, and then completed a categorical perception task with three to four parts during which they categorized ambiguous phonemes on a continuum. All listeners were exposed to the same novel accent, namely words pronounced by a native American English speaker whose [f] or [s] fricative phonemes were replaced by ambiguous [?fs] mixtures, in conditions A and B, respectively. The other fricatives and all other phonemes were not modified. Perceptual learning is reasoned to take place when the overall number of percent labial responses (%f, %v, or %p) is statistically significantly different between the two conditions for each relevant labial-alveolar continuum.

One central question in this dissertation is the level at which linguistic perceptual learning effects take place. Based on the results of the English-only study (Chapter 2), it has been argued that perceptual retuning takes place at the level where phonetic details are processed in terms of phonological features. Another central question is whether perceptual learning carries across languages in second language learners who are listening in either their native or their non-native language, and if so, under what circumstances. Two English-German studies (Chapter 3) and the German-English study (Chapter 4) showed that perceptual learning can carry across languages if the phonemes and phoneme contrasts under consideration are realized similarly in both languages, and that perceptual learning is affected by the language mode and language background of the listeners.

Before discussing the comprehensive results, a summary of the perceptual learning studies discussed in the previous chapters is provided in the table below:

Table 5.1: Summary Table of all Perceptual Learning studies (as reported in Chapters 2-4).
(Note: the shaded cells highlight statistically significant perceptual learning effects.
Blank cells = not tested; B = [?f] training; A= [?s] training.)

	English f/s	English v/z	German f/s	German v/z	English p/t	N
English-only	* st. sign. effect (p=.001); effect size: 11.9% B: 42.6%; A: 30.7%	* st. sign. effect (p=.014); effect size: 11.6% B: 54.9%; A: 43.3%			n.s.; size: 1.5% B: 45.4%; A: 44.0%	N=21 [22-1] (1 low accept. <50% ?f acc.)]
English-German #1	* st. sign. effect (p=.008); effect size: 6.6% B: 44.6%; A: 38.0%		n.s.; effect size: 4.8% B: 45.2%; A: 40.4%	n.s.; effect size: 1.2% B: 33.6%; A: 32.4%		N=20 [no low accpt.s]
English-German #2	* st. sign. effect (p=.009) effect size: 10.8% B: 47.0%; A: 36.2%	n.s. (p=.309) effect size: 4.9% B: 59.2%; A: 54.3%	* st. sign. effect (p=.010); effect size: 15.1% B: 54.8%; A: 39.7%	* st. sign. effect: p<.001 effect size: 20.5% B: 43.9%; A: 23.4%		N=18 [21-3] (3 low accept. <75%)
German-English	* st. sign. effect (p=.003); effect size: 9.7% B: 46.0%; A: 36.3%	n.s.; effect size: 0.7% B: 52.7%; A: 52.0%	* st. sign. effect (p=.035); effect size: 6.4% B: 45.7%; A: 39.3%	n.s.; effect size: -0.8% B: 32.1%; A: 32.9%		N=51 [52-1] (1 low accept. <50% accpt.)

As Table 5.1 above shows, all four perceptual learning experiments led to a perceptual learning effect on the trained /f/-/s/ phoneme contrast in English, the language in which all listeners were exposed to the unusual accent (cf. first column in Table 5.1 above). This shows that the participants in all experiments perceptually accommodated to the novel accents by adjusting the phoneme boundary of the relevant phonemes, /f/ and /s/, in the language in which they were exposed to the novel accent. This study is the first to demonstrate that native and non-native listeners adjust to modified pronunciations of the /f/-/s/ contrast in word-medial position in a native but novel (modified) English accent.

The English-only study, which tested English monolinguals on three English phoneme contrasts (Chapter 2), showed that the generalizability of perceptual learning effects has limits. While listeners extended the adaptation effects on the place of articulation on voiceless fricatives (/f/-/s/) to voiced fricatives (/v/-/z/), they did not extend it to the same place of articulation in (voiceless) stops (/p/-/t/). This finding has been argued to show that perceptual learning effects do not target purely abstract phonological features, but rather, that they adjust how phonological features are signaled by phonetic features. Although the stops and fricatives share the highest level of abstract phonological features for place of articulation, namely *CORONAL* and *LABIAL*, the phonetic realization and relevant acoustic cues (for speech perception) of these shared phonological features differ strongly for the different manner classes (stops vs. fricatives). While the place of articulation in fricatives is primarily cued by differences in formant structures (frequency level, intensity), the place of articulation in stops is primarily cued by differences in the formant transitions going into and coming out of the stop closure phase. The results of the English-only perceptual learning study therefore suggest that perceptual retuning does not target an abstract phonological feature or feature contrast without reference to their phonetic realizations in different segment types. Rather, the results of the English-only perceptual learning study suggest that perceptual learning consists of specific retuning effects that modify how a certain type of phoneme contrast – here, place of articulation in fricatives – is signaled by acoustic cues in the speech stream.

The English-only study therefore provides evidence that generalization of perceptual learning effects does take place at an abstract level, albeit not limit-less or blindly, but with reference to the phonetic realization of these features in different segment contexts. This finding is in line with several studies that tested for generalization of perceptual learning across speakers. While Kraljic & Samuel (2006, 2007) report that perceptual learning on ambiguously voiced stop contrasts in English generalizes across speakers – as well as across place – perceptual learning on fricatives with ambiguous places of articulation have revealed results that are more variegated. Eisner & McQueen (2005) and Kraljic & Samuel (2005, 2007) found that perceptual learning on such fricatives is speaker-specific unless the fricatives are acoustically similar between the speakers (Kraljic & Samuel 2005; cf. Reinisch & Holt 2014:541). Reinisch & Holt (2014) further illustrated that perceptual learning on fricatives can lead to cross-speaker generalizations or speaker-specific results for the *same* speaker voice used to test for generalization, depending on how the fricative continua are constructed for the generalization speaker's voice. When the fricative continua of the exposure and generalization voices were not "sampled across [a similar] perceptual space" (Reinisch & Holt 2014:552), perceptual learning showed speaker-specific results (Reinisch & Holt 2014, Experiment 2). However, when the fricative continua for the two speakers are created such that the listeners perceive them as sounding more similar to each other, cross-speaker generalization occurred (Reinisch & Holt 2014, Experiment 3).

Further evidence for the role of similarity in supporting speech adaptation effects is provided by the discussion of divergent results in Pardo (2006) and Kim et al. (2011) on phonetic convergence in spoken dialogues of interlocutors with the same native language but different accents (Kim et al. 2011:139f). While Pardo's (2006) study finds phonetic convergence between dialogue partners whose dialects differ slightly, Kim et al.'s (2011) study finds no phonetic convergence between dialogue partners who seem to have larger dialectal differences. Taken together, the results of studies on phonetic convergence in speech production with different dialect partners (Pardo 2006; Kim et al. 2011), the results of multiple studies testing for cross-speaker generalization of perceptual learning effects (Eisner & McQueen 2005; Kraljic & Samuel 2005, 2007; Reinisch & Holt 2014), and the results of the cross-manner perceptual learning study in English in the present work (Chapter 2), suggest that adaptation in general is more likely if speakers have similar speech patterns. In particular, the findings suggest that perceptual learning generalizes from the exposure cases to novel cases if they have sufficient acoustic details in common.

This finding also relates to the intricate results of the three bilingual perceptual learning studies reported in Chapters 3 and 4. In particular, the outcome of the English-only study led to the question of whether perceptual learning effects generalize across languages (and across phonological features) as long as two languages share the same phoneme contrasts and distinctive phonological features. A common phonological feature, and in particular a common feature with comparable phonetic realizations across two languages, does not a priori guarantee that adaptation effects which occur in one language automatically carry over into the listeners' other language. Indeed, as the results suggest, language background, language use, and language mode all appear to play a role in whether or not, as well as how strongly perceptual adaptation effects generalize cross-linguistically.

The results of the two L1 English – L2 German studies (Chapter 3) suggest that perceptual learning effects can generalize beyond the specific language in which listeners were exposed to a novel accent, here English. The first English-German study did not show any cross-linguistic perceptual learning effects – i.e. no statistically significant results for either the German f/s-continuum or the German v/z-continuum, though it is possible that more power in the study would have led to a significant result for (at least) the German f/s continuum. The second-English German study, on the other hand, showed large perceptual learning effects on both German fricative continua. The different outcomes of these two studies correlated with differences in the study set-up and language background of the two groups of participants. The volunteers in the fist English-German study used German for significantly fewer hours per week compared to the volunteers in the second English-German study. Moreover, the volunteers in the second English-German study participated in a study environment that promoted a bilingual language mode during the study and around the time of the study.

These diverging results between the two L1 English – L2 German studies (Chapter 3) can be explained with a model of bilingual phonological representations in which the phonemes that are common between two languages – such as /f/ and /s/ in the case of English and German – have separate phonological representations in each language, which are, at the same time, dynamically associated with each other. This means that although the representations are separate, the representations are cognitively associated with each other and can influence each other during cognitive processes, such as speech perception. First, a model with separate phonological representations for each language in bilinguals would be able to explain why perceptual learning effects do not automatically affect both languages of a bilingual or second-language listener (Chapter 3, English-German study #1). Moreover, in cases where bilingual listeners do show cross-linguistic perceptual learning effects, the effect sizes are not identical for corresponding phoneme contrasts in English and German, again suggesting that each language has its own phonological representations for phonemes, including phonemes that are common to both languages. Therefore, the results of the two English-German studies suggest that each language has its own, separate phoneme representations, including for phonemes that are common to both languages of a bilingual or second language learner. These results can be captured with a model that has been proposed by Léwy & Grosjean (1997) in their Bilingual Model of Lexical Access (BIMOLA), as shown in Figure 5.1 below.

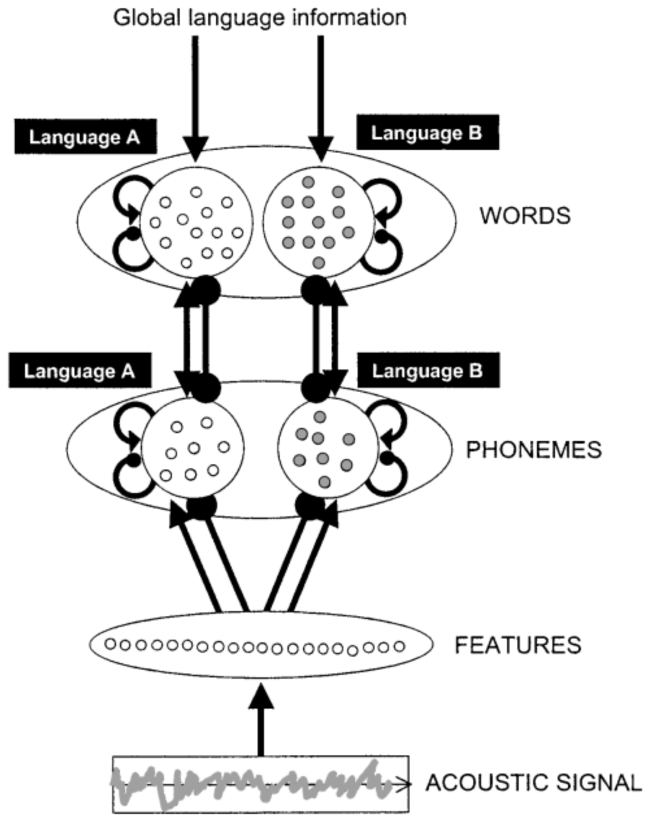

Figure 10.2 The Bilingual Model of Lexical Access (BIMOLA) (Léwy & Grosjean, 1997) bilingual speech perception.

Figure 5.1: The BIMOLA (Bilingual Model of Lexical Access) by Léwy & Grosjean (1997): Each language has separate phoneme representations, as shown at the "PHONEMES" level.

Next, the second English-German study illustrates that boundary shifts in English can lead to boundary shifts in German (Chapter 3). These results suggest that the separate language-specific representations of phonemes common to two languages are linked or dynamically associated with each other. Thus, when second-language learners use their second language more regularly, and when participants are in a more bilingual language mode, the interconnections between shared phonemes appear to be strengthened and facilitate cross-linguistic perceptual learning effects. Figure 5.2 below is an adaptation of the original BIMOLA model seen above in Figure 5.1. The modified figures include an arrow illustrating dynamic interconnectedness, i.e. connectedness between phoneme representations that varies with language mode and language use. The two versions of the adapted BIMOLA model represent two states of dynamically interrelated phoneme representations. The adapted model in a) has a thinner arrow, symbolizing a weaker interconnectedness between representations for a common phoneme; the adapted model in b) displays a thicker arrow, symbolizing a stronger interconnectedness between representations for a common phoneme between the two languages. Thus, the non-cross-linguistic perceptual learning effects in the first English-German study can be accounted for with the weaker interconnectedness between phoneme representations, as illustrated in model a) in Figure 5.2. Cross-linguistic perceptual learning effects in the second English-German study can be accounted for with a stronger cross-linguistic interconnectedness between phoneme representations, as illustrated in model b) in Figure 5.2 below.

a) b)

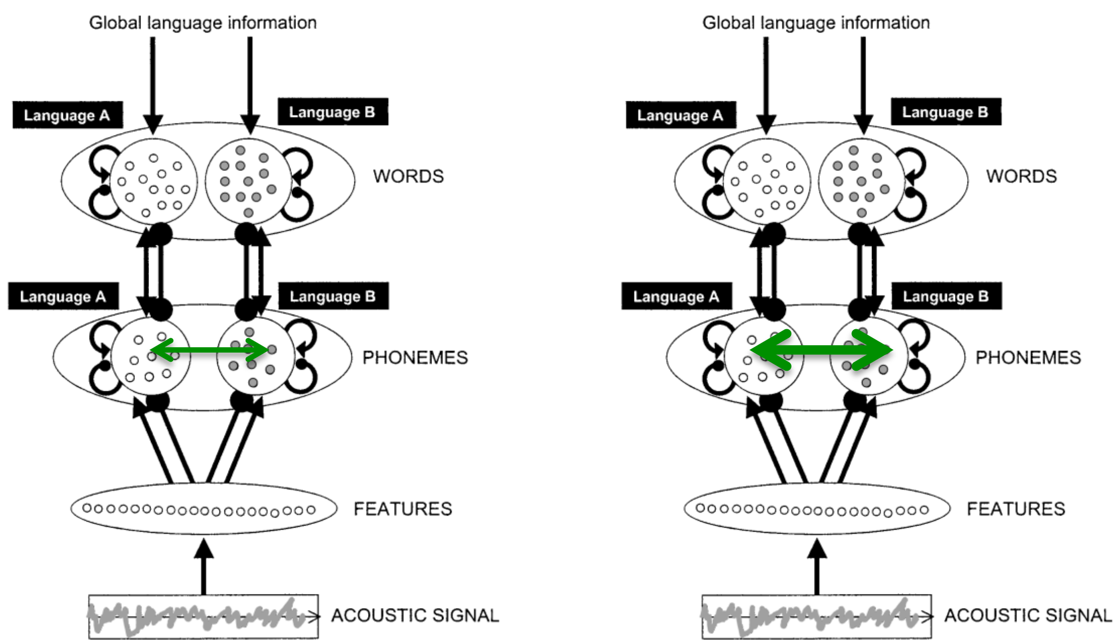

Figure 5.2: An adaptation of BIMOLA (Bilingual Model of Lexical Access) by Léwy & Grosjean (1997): Phoneme representations that are common to both languages are separate but dynamically interrelated. A more bilingual language mode and more language use of both languages leads to stronger interconnectedness (figure a vs. b).

A similar claim about "a connection between phonologically corresponding categories across languages" (Chang 2012:252) was put forth in the well-known study on phonetic drift in an L1 Portuguese-L2 English bilingual (Sancier & Fowler 1997). In a longitudinal case study, they demonstrated that the VOT values of both the bilingual's native Portuguese and non-native English became shorter after spending several months immersed in an L1 environment in Brazil, and longer after spending several months immersed in an L2 environment in the US. Similarly, in a production study on L1 Quichua-L2 Spanish bilinguals with varying ages of L2 acquisition, Guion (2003) shows that those bilinguals who were able to produce distinct L2 Spanish vowels in addition to their L1 Quichua vowels also changed the pronunciation of their three L1 Quichua vowels and produced all of them higher. Guion (2003:121) suggests that "the two systems are merged or at least interrelated". Guion (2003) later rules out the possibility that Quichua vowels are raised because they are merged with Spanish vowels, because the low vowel /a/ would not raise to merge with a Spanish vowel (Guion 2003:122). She concludes that the entire Quichua vowel system raised to allow for sufficient discrimination between the L1 Quichua and the L2 Spanish vowel system in a bilingual sound system (Guion 2003:122f).

Additionally, Chang (2011, 2012, 2013) finds that adult novice L1 English-L2 Korean learners immersed in an L2 environment show "phonetic drift", immediate phonetic influence from the novel L2 being acquired onto their L1, both for vowels and consonants. Based on these results, Chang (2011, 2012, 2013) also argues for "cross-language linkages" (Chang 2012:249). Chang (2012:255) hypothesized L2-to-L1 influence in novice L2 learners, because he assumed that the SLM's (Speech Learning Model, Flege 1995b) equivalence classification takes effect automatically and from the start of L2 speech learning. However, the SLM is a model for speech learning in second language learners with many years of experience, and it is not clear that the mechanisms proposed as part of the SLM should necessarily be available immediately and automatically for novice L2 learners. Moreover, Chang (2011, 2012) argues that phonetic drift in the L1 English vowel space in L1 English-L2 Korean learners is not the result of associations between individual vowels, but rather a shift at the global level of the entire vowel system, similarly to the argument provided by Guion (2003) (cf. Chang 2011:430; Chang 2012:255). For consonants, however, Chang (2012) argues that the phonetic drift in English stops, measured in terms of VOT values and $f0$, is a case of "cross-language linkage" at the level of specific Korean and English segments as well as at the level of "stop type generally" (Chang 2012:261).

Further, Chang (2013) showed that experienced and inexperienced second language learners enrolled in the same language program differed in how much phonetic drift they showed in their native language. Second language learners without any prior experience with the target non-native language showed stronger effects from the non-native onto the native language – a "novelty effect" – compared to second language learners with some prior experience with the non-native target language. Chang (2012:249) suggests that connections between languages start as soon as someone learns a novel, non-native language, and argues that both the native and non-native phonological systems are "dynamic systems undergoing continuous change" (Chang 2013:520). Overall, the results of the two L1 English – L2 German studies (Chapter 3), combined with the research findings on perceptual drift in general (Guion 2003) and on specific phonetic aspects of consonants (e.g. Sancier & Fowler 1997; Chang 2011, 2012, 2013), strongly indicate that phonemes common to two languages do not exist in isolation from each other, but are interconnected with each other.

Lastly, the German-English study (Chapter 4) shows that a) perceptual learning effects also take effect in intermediate-to-advanced L1 German – L2 English listeners who are exposed to an unusual pronunciation of a phoneme contrast (which is known to them from their native language) in their non-native language, and b) that perceptual learning also has cross-linguistic effects in non-native listeners. Regarding the first point, participants with lower scores on the L2 English Perception Proficiency Test showed stronger accommodation effects on the English f/s continuum, suggesting that L2 learners with more 'room for improvement' also show stronger perceptual learning effects. It appears as if the L2 grammars of these intermediate listeners were sensitive to the discrepancies between the L2 learners' grammar and the target grammar, maintaining a status of 'being in flux' (Broselow et al. 1998:274) and consequently adapting more strongly to a novel accent in L2. Kim et al. (2011) tested for phonetic convergence in spontaneous dialogues between dyads of different language backgrounds, including dyads consisting of non-native and native speakers of English. Preliminary analyses of eight non-native–native dyads showed that non-native speakers with an intermediate level of a foreign accent in English converged more with the speech of their native interlocutor than non-native speakers with either a weak or strong foreign accent (Kim et al. 2011:140). The intermediate-to-advanced L1 German – L2 English listeners with lower scores on the L2 English perception proficiency test are likely in a similar mid-range in terms of (phonological) language proficiency as the mid-range non-native speakers in Kim et al.'s (2011) study. Together, the results of Kim et al.'s (2011) study and the German-English study in Chapter 4 suggest that non-native language users might be more adaptive when they are at an intermediate level of language proficiency. This mid-range proficiency level appears to be the most fertile ground for L2 English learners' fine-tuning of their speech production and perception skills, and is based on the interlocutor and the speech input they are exposed to at a given time, respectively.

The finding that L2 listeners adapt to accents in their non-native language and, in particular, that they are more adaptive when their L2 perception proficiency is lower could be explained by one of the tenets in Flege's (1995b, 2003) Speech Learning Model (SLM). The SLM proposes that adult L2 learners (and, in particular, non-beginner L2 learners) "retain the capacity which infants and children make use of in acquiring their L1" (Escudero 2007b:123). In other words, it appears as if L2 adult learners make use of similar mechanisms as (children and adult) native listeners to fine-tune phonological representations, such as statistical learning, which has been proposed and used in Kuhl's (1991, 2000) Native Language Magnet (NLM model). Mechanisms such as statistical learning, which play a role in the acquisition of L1 phonological contrasts, have been argued to still be active principles in adult second language learners who have not yet reached the stage of fossilization (i.e. their endpoint in L2 language development), and who still have room for improvement – i.e., the groups of L2 learners for which Flege (1995b, 2003) proposed his SLM. The evidence collected here does not entail that adult L2 learners should be able to perceive (non-native) L2 contrasts equally well as native listeners, nor that L2 learners should be able to learn how to perceive non-native phonological distinctions, in line with Kuhl's NLM model and Flege's SLM. As a matter of fact, perceptual learning in L2 listeners on phoneme contrasts common to the listeners' two languages is not identical to perceptual learning in native listeners when testing for the generalizability of perceptual learning effects, as described below.

Regarding the second important finding of the L1 German- L2 English study (Chapter 4), advanced L1 German – L2 English listeners differ from English native listeners in terms of how far they generalize the perceptual learning effects. The listeners generalized perceptual learning effects across languages from the English f/s contrast to the German f/s contrast, but did not generalize the perceptual learning effects to the voiced fricative contrast /v/-/z/, in either the language of training, their L2 English, or the untrained language, their L1 German. The monolingual English listeners (Chapter 2), on the other hand, generalized the perceptual learning effect from the voiceless ambiguous fricatives (/f/-/s/ contrast) to the voiced ambiguous fricatives (/v/-/z/ contrast) within English. These results suggest that speech accommodation effects for consonants proceed differently in non-native speakers than in native listeners.

Overall, several experimental studies suggest that native listeners generalize learning to other *consonants* that share the relevant feature(s). In another adaptation study with vowels, Maye et al. (2008) report that exposing native English listeners to an English accent with lowered front *vowels* does not lead listeners to generalize this lowering to back vowels (Experiment 2). This finding in Maye et al. (2008) can be explained with the fact that vowel height in front and back vowels makes use of different acoustic-phonetic cues, including additional features such as lip-rounding for /u/, for example. These results in studies involving vowels might also be due to the fact that listeners are better able to distinguish fine-grained acoustic details in vowel realizations and vowel continua, which lead to categorical perception that is less steep – less categorical – in vowels compared to consonants (Pisoni 1973; quoted in Cutler 2012:8f). Aside from these studies on vowel adaptation, native listeners have been shown repeatedly to make use of sub-phonemic features and to generalize learning/imitation between consonants that share the relevant features, from studies on perceptual learning (Kraljic & Samuel 2006), to phonetic imitation (Nielsen 2011), and phonological learning (Goldrick 2004).

L2 listeners, on the other hand, seem to change their grammatical systems in small steps, only adapting the particular phoneme contrast that was pronounced in a novel way. As a learning strategy in continued second language acquisition, L2 learners seem to fine-tune their L2 grammars by making the smallest possible modifications given the input. If different rules or grammar adjustments can capture the same novel pattern discovered in the input, learners take the modifying rule that requires the least amount of modification or adjustment while still being able to account for the newly recognized patterns in the input. In this case, the L2 listeners hear pronunciations that seem to have shifted the boundary between the two phonemes /f/ and /s/ (in English). The grammar adjusts the phoneme boundary of /f/ and /s/ in English accordingly, but not of the related phoneme boundary /v/-/z/, since the input did not provide any evidence for a modified boundary in voiced fricatives. In fact, no voiced fricatives occurred at all in the input.

An alternative explanation as to why non-native listeners did not generalize perceptual learning effects across both languages and features might be found in cognitive limitations of second language processing. Second language speech production and perception is known to afford a high "attentional demand [which] adds processing load" (Costa et al. 2008:540), which might block alignment between speakers and listeners (under Pickering & Garrod's (2004) interactive-alignment account) and reduce the overall level of automaticity in speech production, speech perception, or dialogue. It seems possible that non-native listening requires so much attention and working memory that the resources available for linguistic processing, including phonetic fine-tuning and adaptation effects, are reduced. Consequently, perceptual learning might take place at a phonemic level, where only the particular phoneme contrast is adjusted that was modified in the artificial accent to which the non-native listeners were exposed. Such individual adjustments at the phoneme- rather than feature-level might also explain why it is more difficult for non-native than native listeners to get accustomed to a novel dialect in the non-native listener's L2 (Cutler 2012:407). Further research is necessary to establish whether this reduced generalization of perceptual learning effects in non-native listeners results from purely linguistic or cognitive constraints, or perhaps from a combination of the two.

Although cross-linguistic generalization in the L1 German – L2 English study (Chapter 4) was more limited, the German-English listeners did generalize the perceptual retuning effect to the German f/s contrast, although the unusual accent was provided in only English. This phenomenon can also be captured by the model of bilingual phonological representations presented above in the adapted BIMOLA model in Figure 5.2. To recap, the phonemes of each language have separate but dynamically associated phonological representations for phonemes that are common between the two languages. The adapted BIMOLA model predicts that participants generalize the adjustment of the boundary for /f/ and /s/ in English to the boundary for /f/ and /s/ in German. This is predicted because the participants in this study were intermediate-to-advanced L2 English users who, on average, used English on a regular basis and were in a bilingual language mode during the course of the study. Furthermore, the L2 English listeners did not adjust the phoneme boundary of the English v/z contrast, most likely because their retuning strategies are conservative and restrictive in nature. Finally, the v/z boundary in German was not perceptually modified because it is dynamically associated with the (unmodified) v/z boundary in English, as predicted by the adapted BIMOLA model.

A comparison of the results of the (second) English-German study and the German-English study suggests that generalization from an L2 to an L1 ('regressive transfer') is not as strong as generalizations from an L1 to an L2 (‘progressive transfer’). In the second L1 English – L2 German study, the effect sizes for participants’ perceptual learning effect were largest in their L2 German: the effect size was 10.8% for L1 English f/s, 15.1% for L2 German f/s, and 20.5% for L2 German v/z. In the L1 German – L2 English study, on the other hand, perceptual learning in listeners’ L2 English led to perceptual learning on the trained L2 English f/s contrast (9.7%), and a slightly weaker perceptual learning effect on the f/s contrast in the listeners’ L1 German f/s (6.4%). These differences in the strength of transfer effects might be due to either a) L2 representations being weaker, less mature, not as replete with exemplars, and thus more easily affected compared to more mature L1 representations with lots of exemplars and experiences, or b) associations between the representations that include directionality and allow for more influence from L1 onto L2 than from L2 onto L1. Further research will be needed to tease these two possible explanations apart.

Another area of further inquiry should address why the second L1 English – L2 German study found such a large perceptual learning effect size in the L2 German v/z case. As described above, larger perceptual learning effects in German v/z correlate with listeners’ higher scores on the L2 German perception proficiency test as well as the order of presentation (when German v/z was tested right after the lexical decision task). It is not possible, however, to determine whether both factors contributed in isolation to this effect and increased the perceptual learning effect size, or whether these two factors coincidentally co-varied with the larger perceptual learning effect size on German v/z. Reinisch & Holt’s (2014:549) study on testing for cross-speaker generalization of perceptual learning effects in fricatives, which showed overall speaker-*specific* results, demonstrates that the first block of the continuum responses did in fact show perceptual learning transferring from the exposure voice to the generalization voice. Reinisch & Holt (2014:549) suggest that the listeners first apply perceptual learning “non-selectively” until they have gathered enough input from the novel speaker to recognize that the fricative realizations between the two speakers do not match, at which point perceptual learning is no longer generalized to the new voice. The initial “non-selective” or general perceptual learning effect, however, was only apparent in the very first block of the testing phase and was not larger than for the training voice. Our German v/z perceptual learning effect, however, was present for the entire testing block and was substantially larger than the effect on the trained English f/s contrast. Therefore, it does not seem likely that the German v/z continuum had such a large perceptual learning effect because this continuum was presented first after the lexical decision exposure phase.

The more convincing explanation is that the perceptual learning effect in German v/z was larger than in German f/s, which in turn was also larger than English f/s, because the German phonemes /v/ and /z/ have been recognized by the L2 German listeners with superior L2 German perceptual skills as different from English /v/ and /z/. Factors that might have contributed to L2 German learners recognizing a difference between English /v/ and German /v/ are the non-English-like [v] realizations in many German speakers, the different orthographic representations of [v] in German compared to English, and possibly the voiceless allophonic realization of voiced obstruents – including /v/ – word-finally. In addition, L2 representations are likely weaker and more malleable than L1 representations because listeners have not yet garnered as much experience and as many L2 exemplars. It makes sense that particularly those listeners who are good at discriminating non-native phoneme contrasts are also more likely to notice differences in the pronunciation of similar and equivalent sounds in the two languages. Perhaps, a stronger L2 perceptual capacity is also an indication of a generally stronger awareness of differences between the languages, including differences in the orthographic representation between the two languages.

The results reported in this dissertation and the results reported in Bradlow & Bent (2008) and in Baese-Berk et al. (2013) suggest that perceptual learning is not a general increase in tolerance for unusually pronounced segments – e.g., a foreign accent – but rather an adjustment of particular features and/or phoneme contrasts for which listeners have heard various non-canonical examples. Indeed, the non-native [?fs] accents to which listeners were exposed in the experiments reported in this dissertation are based on mixtures of [f] and [s] segments that are based on specific recordings for each individual word. Similarly, Baese-Berk et al. (2013) suggest that high-variability training with multiple foreign accents (Thai, Korean, Hindi, Romanian, Mandarin) encouraged listeners to hone in on the systematic variability common to all five speakers' accented English.

5.2 Future research

The research findings presented in this study concentrated on the fricative contrasts /f/-/s/ and /v/-/z/, and, to some extent, the stops /p/-/t/, in English and German. The overall study focused on two languages with common phonemes and very similar phonetic/acoustic details of the phonemes under consideration. An earlier study with Hindi-English bilinguals did not find cross-linguistic perceptual learning effects from the voicing contrast on English velar stops to the four-way voicing contrast in stops at various places of articulation in Hindi (Schuhmann 2012). In order to increase the likelihood of finding cross-linguistic perceptual learning results in the current study, languages were selected in which the relevant phonemes have comparable phonetic realizations. To the same effect, sequential bilinguals were tested, both novel L2 learners who were exposed to an unusual accent in their native language, and advanced L2 learners who were exposed to an unusual accent in their non-native language.

In order to determine whether listeners can in fact generalize across multiple, variably produced segments – while systematically controlling the variability of the relevant segments – one avenue for future research would be to design a perceptual learning study with an artificial novel accent in which multiple features on multiple phoneme contrasts are manipulated. Reinisch & Holt (2014) conducted a study in which they exposed native English listeners to one ambiguous phoneme contrast within a foreign accent – Dutch-accented English.

One of the next steps in this research program is to test for cross-linguistic perceptual learning effects on phonemes that are common between two languages but differ (more) saliently in terms of the phonetic realization of the relevant phonetic-acoustic cues. For example, a future study could train English-Spanish or English-(Canadian) French bilingual listeners on a modified /f/-/s/ contrast in one of the languages, and then test whether the perceptual learning effects carry over into the other language. The relevant difference to the English-German/German-English studies reported in this thesis is that certain dialects of Spanish and French typically realize /s/ as dental (Maddieson 1984:267, though see Martínez-Celdrán, Fernández-Planas & Carrera-Sabaté 2003:257*f* for Castilian Spanish, who classify only stops as dental, specifically laminal denti-alveolar (Martínez-Celdrán et al. 2003:257)), while most English speakers typically realize /s/ as alveolar (Ladefoged & Maddieson 1996:154). In the case of English-French bilinguals, a future study could even test for generalization across languages and across features with different phonetic-acoustic details, as French has both voiceless and voiced dental/denti-alveolar sibilant fricatives, /s/ and /z/, respectively (Maddieson 1984:267). English-French listeners could be trained on a modified /f/-/s/ contrast in English, and then tested on an /f/-/s/ continuum in English and in French, as well as on the voiced fricative continuum /v/-/z/ in English and in French. This study would determine whether perceptual learning can generalize across languages and across features, even when the distinctive feature under consideration, here the alveolar place of articulation, is realized with distinct CORONAL places of articulation in the two languages.

Along the same lines, a future study could test a language combination in which certain shared phoneme contrasts have even more dissimilar, and possibly more salient and noticeable phonetic realizations in the two languages, such as different ranges on the VOT continuum for voicing contrasts in stops. While English and German use the phonetic feature [+/- spread glottis] to indicate a stop contrast between a long-lag /p/ and a short-lag /b/ (and between voiced elements, some prevoicing), Spanish, French and Dutch, for example, use ('true') phonetic voicing to distinguish a similar stop contrast, namely short-lag /p/ and pre-voiced /b/. For instance, English-Spanish or English-French listeners could be trained on a modified place of articulation contrast in stops in English, such as a modified /p/-/t/ contrast, and then tested for perceptual learning effects on the place contrast in English as well as in Spanish /p/-/t/. As Spanish stops are also often classified as being predominantly dental (Maddieson 1984:267) or laminal denti-alveolar (Martínez-Celdrán et al. 2003:255,257) in certain dialects, this would test whether perceptual learning can generalize to another language in which the phoneme contrast is realized phonetically differently in terms of the place of articulation and in terms of the phonetic features that are used in the two languages to distinguish voiced and voiceless phonemes – measurable in terms of VOT values. Moreover, a future study with English-French or German-French bilinguals could test whether the trained /p/-/t/ contrast in English/German would also lead to perceptual adjustments on the voiced stop contrast that differs in the same place of articulation features, namely /b/-/d/, in English as well as French (Spanish does not have /b/ or /d/ in enough contexts to make effective comparisons to English; cf. Maddieson 1984:267; cf. Martínez-Celdrán et al. 2003:257). Results of such a study would provide additional evidence about whether bilinguals have interconnected phonemes and/or features for processing similar speech sounds in their two languages.

Another similar avenue for future research would be to train English-French or German-French speakers on the laryngeal contrast directly, which is realized as a [+/-voice] contrast in French, but as a [+/- spread glottis] contrast in English and German. The listeners could be provided with modified versions of the /p/-/b/ contrast in English or German, for example, and then tested for generalization of perceptual learning effects to the French /p/-/b/ contrast. This would determine whether perceptual retuning of a distinctive feature that is realized using different aspects of the same acoustic cues continuum in the two languages, namely the phonological feature *voice*, which is realized with different VOT values in English/German and French, can generalize from one type of voicing language to another. To test for cross-linguistic and cross-featural effects, this study could further test for perceptual learning effects on the alveolar stop contrast /t-d/ in English and the dental stop contrast /t-d/ in French.

Finally, another planned research study in this research program focuses on the types of bilinguals involved, rather than the type of phonological contrasts tested. It has been demonstrated that sequential or late bilinguals often show more effects of the two languages influencing each other – either by showing evidence of merged categories with values intermediate between those of monolinguals of either language, or by exaggerating the differences between sounds in the two languages (cf. Werker, Byers-Heinlein & Fennell 2009:3653; Flege et al. 2003; Fowler et al. 2008) – compared to simultaneous or early bilinguals, who often discriminate all sound contrasts in both of their languages like monolinguals of each language (cf. Werker et al. 2009:3653; Pallier et al. 1997, 2001; Sundara & Polka 2008). This finding leads to the hypothesis that perceptual learning should lead to cross-linguistic effects with highly proficient, *sequential* bilinguals, but not with highly proficient, *simultaneous* bilinguals, who seem to be able to store and process the phonological contrasts in their two languages separately.

We have shown that perceptual learning can generalize across voice in fricatives with an ambiguous place of articulation, namely from the /f/-/s/ to the /v/-/z/ contrast in monolingual English listeners. However, listeners do not extend perceptual learning to stops with the same place of articulation (/p/-/t/). We have argued extensively that this limitation in perceptual learning across manner contrasts must be founded in the difference in how place features are signaled in fricative contrasts as opposed to stop contrasts. A growing body of research suggests that perceptual learning is more likely to generalize to new contexts – such as new contrasts or new voices – the more similar the training and the testing cases are.

The majority of the research in this thesis tests for perceptual learning in novice and intermediate-to-advanced second language listeners, to establish how L1 English – L2 German and L1 German – L2 English learners mentally represent and adjust the fricatives /f,s,v,z/, which are common to English and German. The thesis demonstrated that non-native listeners are adept at learning to perceptually adjust to aberrant phonetic details in their second language – at least for a phoneme contrast common to both of their languages. However, perceptual learning in a second language is not as broadly tuned as perceptual adaptations in a listener's first language: adjustments to phoneme boundaries are specific to those phoneme contrasts for which non-native listeners heard unusual pronunciations, but they do not generalize this adaptation effect to other phoneme contrasts that share the same relevant feature. The data further revealed that perceptual learning effects can generalize from the listeners' native language to their novel non-native language. All of the results from the bilingual experiments can be explained with a model of bilingual listeners in which phonemes that are common to two languages have separate but dynamically associated phonological representations. Separate representations explain why adjustments do not automatically affect both languages, as was the case in the first L1 English-L2 German study. When listeners are in a bilingual language mode and use their novel L2 on a more regular basis – as was the case in the second L1 English- L2 German study – the interconnections between shared phonemes seem to strengthen, thus facilitating cross-linguistic effects.

To conclude, this dissertation used the methodology of perceptual learning to further our understanding of the adaptive nature of phoneme categorization in speech perception and more specifically, the representation and processing of phonemes that are common to two languages. The findings showed that perceptual learning is a retuning process that adjusts how phonological features are signaled by phonetic cues, and that for second language learners, this process proceeds on a phoneme by phoneme basis, rather than readily generalizing different sounds contrasting in the same feature. In addition, the results indicate that the phonological representations of phonemes that are common to the two languages in bilingual English-German listeners have separate but dynamically interrelated of phonological representations. Much interesting work remains to be done on the nature of these representations and the way that they are associated in the minds of bilinguals.

Appendix 1: Filler words

	Filler words	Syllables	frequency[22]
1	award	2	47.4
2	glamour	2	38.5
3	annoying	3	44.5
4	barbecue	3	37.8
5	bewilder	3	27.4
6	blueberry	3	42.6
7	brewery	3	34.7
8	bullying	3	37.5
9	cabaret	3	29.3
10	calendar	3	50.2
11	columbine	3	26.9
12	corridor	3	49
13	crocodile	3	46.4
14	decorum	3	31.9
15	dishwasher	3	40.7
16	engineer	3	52.1
17	gallery	3	47.1
18	gardenia	3	32.3
19	gondola	3	36.1
20	gullible	3	33.3
21	Halloween	3	45.7
22	library	3	57.8
23	melodic	3	23.9
24	memory	3	56.9
25	millionaire	3	40.9
26	myriad	3	40.6
27	neighborhood	3	56.7
28	rekindle	3	34.6
30	acclimation	4	1

[22] FSI for *acclimation* is listed as "1", because the word was not included in Zeno et al. (1995).

	Filler words	Syllables	frequency[23]
31	ammunition	4	46
32	amygdala	4	31.5
33	armadillo	4	36.6
34	bulimia	4	28.7
35	calibration	4	28.3
36	Colombian	4	34.2
37	colonial	4	55.2
38	combination	4	55.9
39	condemnation	4	40.4
40	dandelion	4	44.7
41	education	4	60.9
42	embroidery	4	37.3
43	generation	4	54.9
44	inhibition	4	34
45	legendary	4	44.7
46	linoleum	4	41.1
47	macaroni	4	41.4
48	mahogany	4	45
49	malaria	4	46
50	missionary	4	43.8
51	recognition	4	51.6
52	reliable	4	50.4
53	theologian	4	34.3
54	accommodation	5	41.6
55	biological	5	52.3
56	biomedical	5	37
57	collaboration	5	35.5
58	communication	5	57.2
59	humiliation	5	43.3
60	theological	5	41.5
	Average:	**3.60**	**41.53**

[23] FSI for *acclimation* is listed as “1”, because the word was not included in Zeno et al. (1995).

Appendix 2: Filler non-words

	Filler non-words	syllables
1	abown	2
2	adreng	2
3	ammain	2
4	amolb	2
5	dibale	2
6	ibsholb	2
7	ibume	2
8	lesholde	2
9	shiloo	2
10	shorray	2
11	umbroyl	2
12	unench	2
13	aloygen	3
14	engrabel	3
15	anligon	3
16	anulshon	3
17	bayneeder	3
18	blimodik	3
19	bloyminder	3
20	bollidic	3
21	brindoming	3
22	dranulare	3
23	dreenaly	3
24	galakeed	3
25	glenmodish	3
26	grudebane	3
27	harudom	3
28	heroilum	3
29	hindeydi	3
30	imbloyel	3

	Filler non-words	syllables
31	imkellish	3
32	kalaben	3
33	keymodish	3
34	kiloka	3
35	kinudare	3
36	klarimbo	3
37	kolango	3
38	komerdy	3
39	kormouber	3
40	krimodel	3
41	lidimen	3
42	lorashen	3
43	loydermy	3
44	luddermel	3
45	lundoykel	3
46	makalood	3
47	makoday	3
48	modashi	3
49	neligoon	3
50	niathim	3
51	nibdilling	3
52	nodabare	3
53	ombeelad	3
54	onishon	3
55	oymalere	3
56	romashal	3
57	roymakeer	3
58	rulordgem	3
59	runodeng	3
60	umbering	3

	Filler non-words	syllables
61	umkeeben	3
62	undower	3
63	uyarmel	3
64	wakabore	3
65	woldemier	3
66	abolimky	4
67	abonishen	4
68	ankla.na.ble	4
69	bedulachi	4
70	beliderel	4
71	benemallow	4
72	berleshelang	4
73	berthuenou	4
74	boreguma	4
75	dihormely	4
76	ebrishium	4
77	emodooler	4
78	gabomilly	4
79	gellimoder	4
80	gerodingel	4
81	gomindrido	4
82	gonabeema	4
83	hamiderly	4
84	iblergaber	4
85	illamboyder	4
86	kamashillan	4
87	kenidasher	4
88	kimaleum	4
89	lablaruam	4
90	legelliom	4
91	leshormuthon	4
92	linobander	4

	Filler non-words	syllables
93	noraliu	4
94	nomemoli	4
95	nomanuo	4
96	radoliush	4
97	rimoriun	4
98	rymolinger	4
99	therodosher	4
100	unchimeeling	4
	Average:	**3.23**

Appendix 3.1: L1 English-L2 German study #2 with low acceptors (all participants)

This analysis of the 2nd L1 English-L2 German study is based on the results of all participants except *including* the "low acceptors", who accepted fewer than 70% of the ambiguous critical stimuli. The participants in the study performed well overall in the lexical decision task with a mean overall accuracy rate of 94.79% (SD 3.3%), ranging from 84.5% to 98.5%. Table 6.1 below shows that the man accuracy for all words (critical stimuli and filler words) was 95.7% (SD 4.1%); individual scores ranged from 84.0% to 100%. Table 6.1 below further shows that the mean accuracy for all non-words (filler non-words) was 93.9% (SD 6.4%); individual scores ranged from 70.0% to 100%. Participants were also faster to correctly accept words (295.3 ms, SD 148.9) than to correctly reject non-words (461.8 ms, SD 213.0).

Table 6.1: English-German study 2 (all participants): Lexical Decision (all stimuli)

	words	**non-words**
% correct* (SD)	95.7% (4.1%)	93.9% (6.4%)
RT (in ms)** (SD)	295.3 (148.9)	461.8 (213.0)

* *Mean accuracy (i.e. "yes" responses to critical items)*

** *Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Table 6.2 below illustrates that the mean overall accuracy for the unmodified and modified critical stimuli for all participants (i.e. including the three low acceptors). Overall, the response time (RT) and acceptance rate data suggest that the critical stimuli sounded acceptable to the listeners, although participants responded about 12.4 ms slower to ambiguous [?f] items than unmodified [f] stimuli (311.5 ms. compared to 299.1 ms., respectively), and accepted fewer (15.3%) of the ambiguous [?f] compared to natural [f] items (81.5% compared to 96.8%, respectively). Participants also responded slower to ambiguous [?s] items than to unmodified [s] stimuli (115.3 ms slower) and accepted slightly fewer (0.6%) of the ambiguous [?s] items compared to natural [s] items (99.1% vs. 98.5%, respectively). (Note, though, that each participant heard either only ambiguous [?f] and natural [s] words OR ambiguous [?s] words and natural [f] words, just like in the English-only study.)

Table 6.2: English-German study 2 (all participants): Lexical Decision (critical stimuli)

	Natural stimuli		Ambiguous stimuli	
	[f]	**[s]**	**[?f]**	**[?s]**
% correct* (SD)	96.8% (3.4%)	98.5% (3.4%)	81.5% (16.8%)	99.1% (2.0%)
RT (in ms)** (SD)	299.1 (180.8)	214.3 (128.5)	311.5 (125.5)	329.6 (178.3)

** Mean accuracy (i.e. "yes" responses to critical items)*
*** Mean RT for correct items (i.e. "yes" responses to critical items) measured from word offset*

Separate RM ANOVAs were carried out on the lexical decision results, one on acceptance rates of the critical stimuli and one on RT data. The analyses were performed separately using participants (F1) or the critical stimuli (F2) items as repeated measures. Additional factors in the ANOVAs were 1) training condition, i.e. whether participants heard ambiguous [?f] and unambiguous [s] or ambiguous [?s] and unambiguous [f] pronunciations, a between-participants but within-items factor, and 2) stimulus type, i.e. whether the critical stimuli contain an /f/ or an /s/ phoneme, a between-item but within-participants factor.

The acceptance rates of the critical items shows an interaction between condition and stimulus type (F1(1,19)=7.500, p=.013*; F2(1,38)=18.901, p<.001). Overall, participants accept more /s/-words ([?s] items: 99.1%, [s] items: 98.5%) than /f/-items ([?f] items: 81.5%, [f] items: 96.8%), as indicated by a main effect of stimulus type (F1(1,19)=12.844, p=.002*; F2(1,38)= 16.348, p<.001). The acceptance rate of critical stimuli also shows an effect of condition (F1(1,19)=9.190,p=.007*; F2(1,38)=16.197, p<.001).

The RT data show a significant interaction in the by-subject analysis (F1(1,19)=15.850, p=.001*) but not – or only marginally – in the by-item analysis (F2(1,38)=3.820,p=.058). The RT data also show a marginally significant effect of stimulus type in the by-subject analysis (F1(1,19)=4.321, p=.051*) but no significant effect in the by-item analysis F2(1,38)=2.548, p=.119), providing no conclusive evidence that participants responded slower to /f/-stimuli ([f] stimuli: 299.1; [?f] stimuli: 311.5) than to /s/-stimuli overall ([?s] stimuli: 329.6, [s] stimuli: 214.3). The RT data also show an effect for condition in the by-item analysis (F2(1,38)=10.374, p=.003*) but not in the by-subject analysis (F1(1,19)=.596,p=.450).

There was no significant three-way interaction between language, phoneme contrast and condition (F(1,19)= 3.257, p=.087). There was, however, a statistically significant interaction between condition and language (F(1,19)=4.996, p=.038), indicating the difference in percent labial responses between conditions was significantly larger in German than English: the effect of condition was larger in German (44.4% vs. 31.6%, i.e. a difference of 12.8%) than in English (50.7% vs. 45.2%, i.e. a difference of 5.5%). There was also a significant main effect of condition (F(1,19)=8.536; p=.009) . Keeping in mind that there was also a (dominating) interaction between condition and language, this main effect of condition indicates an overall effect of training condition: Participants who heard ambiguous [?f] words categorized overall more items on the English and German f-s and v-z continua as labial (%f and %v) (47.5%) than participants who heard ambiguous [?s] words (38.4%).

There was also a statistically significant interaction between language and phoneme contrast (F(1,19)=63.647, p<.001). This effect, however, is not important to the analysis. The effect shows that the v/z phoneme contrast had a larger difference in percent labial responses between languages (55.7% in English vs. 31.7% in German) than the f/s phoneme contrast (40.2% in English vs. 44.3% in German). Finally, there was a significant main effect of language (F(1,19)=36.415, p<.001). This effect indicates that the mean percent labial responses for both contrasts and both conditions in English (47.9%) were significantly larger than in German (38.0%). (Moreover, the interaction between phoneme contrast and condition was not significant (F(1,19)=.111, *p* =.743), and neither was the main effect of phoneme contrast (F(1,19)=.714, *p* =.409).

Finally, simple effects for the English-German study 2 including all participants – i.e. also including the three low acceptors – reveals again a statistically significant effect for the English f/s contrast (F(1,19)=5.635, p=.028*), no statistically significant effect for English v/z (F(1,19)=.480, p=.497), and a significant effect for the German v/z contrast (F(1,19)=14.366, p=.001*). Most importantly, however, this version of the analysis does not lead to a statistically significant perceptual learning effect for the German f/s phoneme contrast (F(1,19)=3.005, p=.099), although this effect is arguably marginally significant. The details for each of the simple effects are provided in the sections below.

As predicted, participants showed a perceptual learning effect on the English f/s contrast, the language and phoneme contrast of training: Participants in the ambiguous [?f] condition categorized more items of the f-s continuum as the labial consonant “f” (44.19%) than participants in the ambiguous [?s] condition (36.16%), a significant simple effect (F(1,19)=5.635, p=.028*), obtained through a planned pairwise comparison. The difference in percent “f” responses between the two conditions amounts to 8.03%, still a strong effect size but smaller than the 10.85% effect size for the English [f]-[s] continuum obtained above when the low acceptors were not included in the analysis. Figure 6.1 below illustrates successful perceptual learning on the English [f]-[s] continuum in the English-German study #2 when analyzing all participants, including the low acceptors.

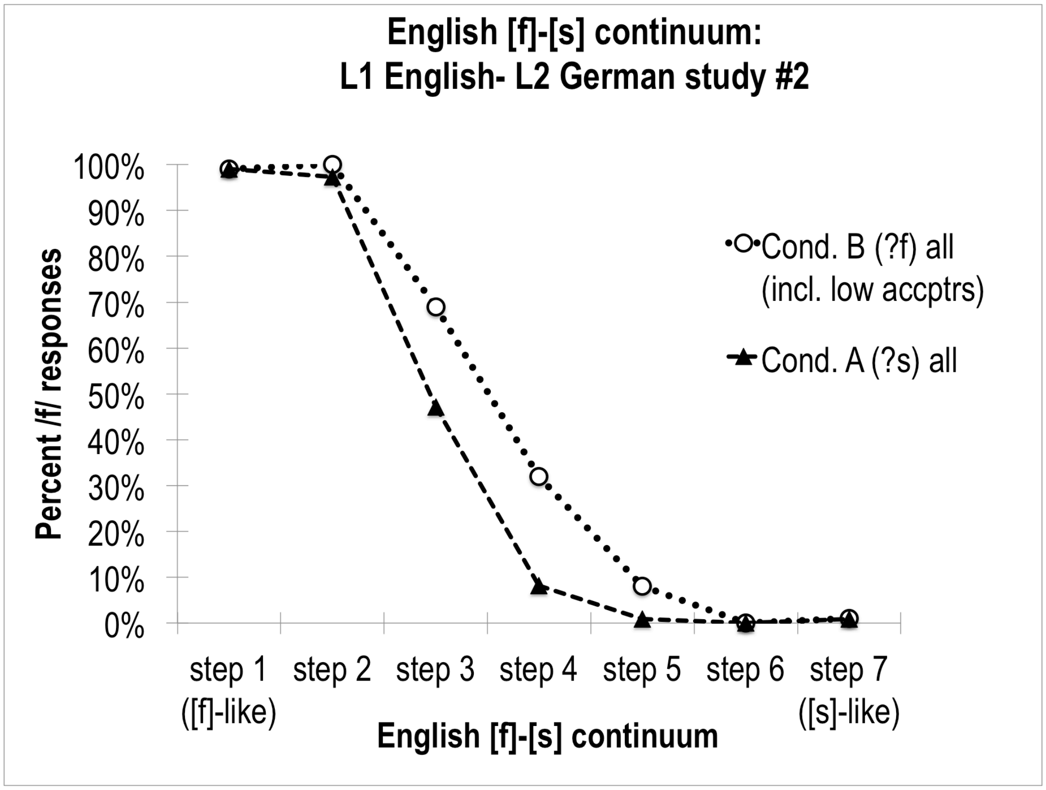

Figure 6.1: English-German study 2 (including low acceptors): Perceptual Learning on the English [f]-[s] continuum.

Participants showed only a marginally significant perceptual learning effect on the German f/s contrast, the phoneme contrast of training in an untrained language: Participants in the ambiguous [?f] condition categorized more items of the f-s continuum as the labial consonant "f" (48.79%) than participants in the ambiguous [?s] condition (39.74%), but the simple effect, obtained through a planned pairwise comparison, reached significance only marginally ($F(1,19)=3.005$, $p=.099$). The difference in percent "f" responses between the two conditions amounts to 9.05%, down from 15.07% in the analysis without the low acceptors. Figure 6.2 below illustrates successful perceptual learning on the German [f]-[s] continuum in the English-German study #2 when the low acceptors are included in the analysis.

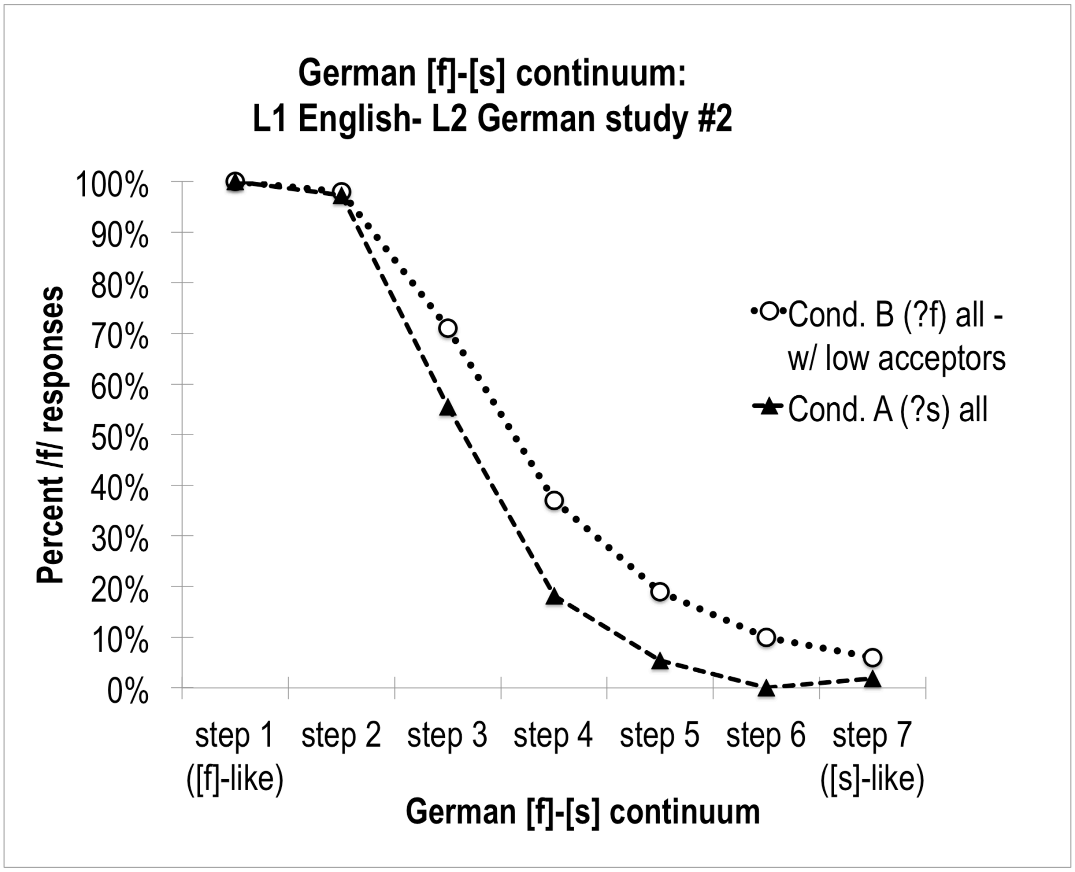

Figure 6.2: English-German study 2 (including low acceptors): Perceptual Learning on the German [f]-[s] continuum (only marginally significant).

Moreover, participants showed a perceptual learning effect on the German v/z contrast, an untrained phoneme contrast in an untrained language: Participants in the ambiguous [?f] condition categorized more items of the German [v]-[z] continuum as the labial consonant "v" (40.00%) than participants in the ambiguous [?s] condition (23.42%) a significant simple effect ($F(1,19)=14.366$, $p=.001$*) obtained through a planned pairwise comparison. The difference in percent "v" responses between the two conditions amounts to 16.59%, down from 20.46% in the analysis without the low acceptors. Figure 6.3 below illustrates successful perceptual learning on the German [v]-[z] continuum in the English-German study #2 when including the low acceptors in the analysis.

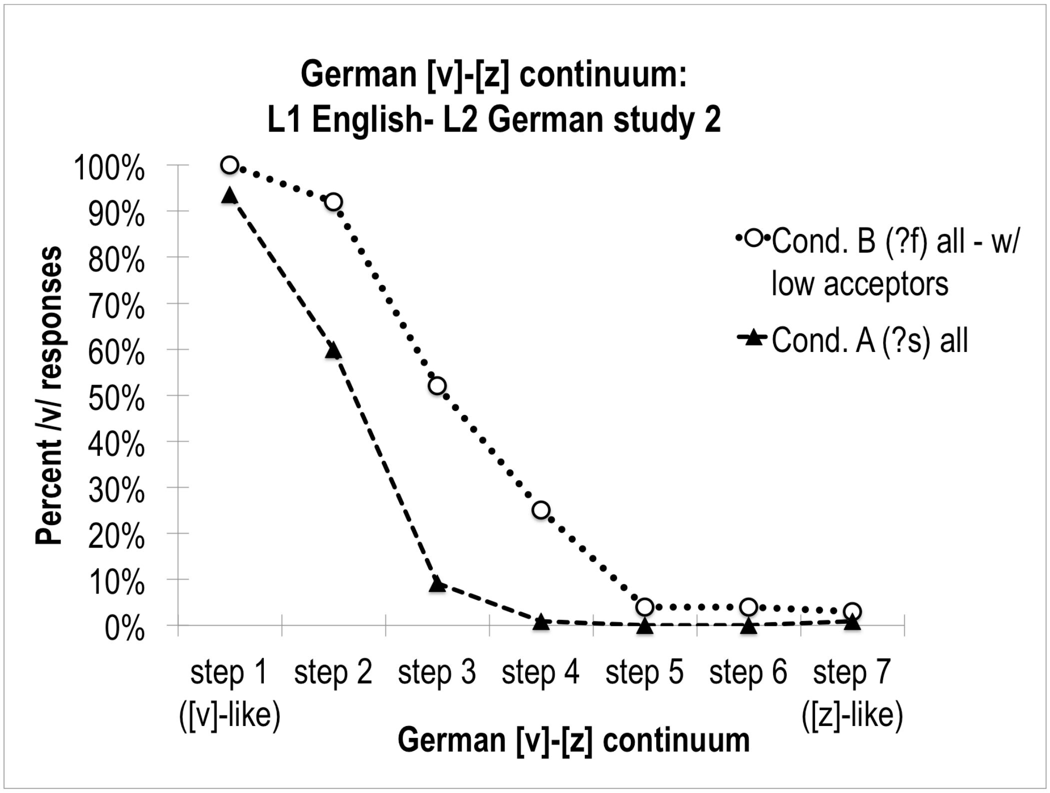

Figure 6.3: English-German study 2 (including low acceptors): Perceptual Learning on the German [v]-[z] continuum.

Rather unexpectedly in this context, participants did not show a statistically significant perceptual learning effect on the English v/z contrast, an untrained phoneme contrast in the language of training: Participants in the ambiguous [?f] condition categorized more items of the English [v]-[z] continuum as the labial consonant "v" (57.14%) than participants in the ambiguous [?s] condition (54.29%) a non-significant simple effect (($F(1,19)=.480$, $p=.497$) obtained through a planned pairwise comparison. The difference in percent "v" responses between the two conditions amounts to 2.85%, down from 4.89% in the analysis without the low acceptors. Figure 6.4 below illustrates no perceptual learning on the English [v]-[z] continuum in the English-German study #2 when analyzing the results of all participants, including the low acceptors.

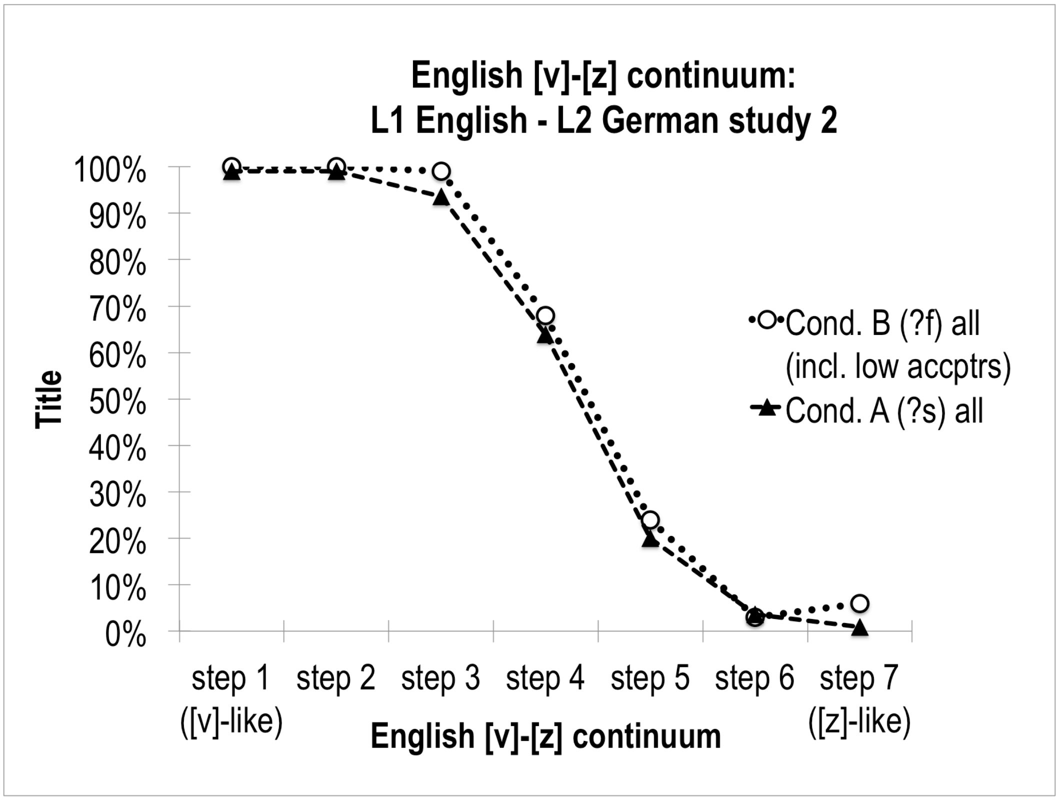

Figure 6.4: English-German study 2 (including low acceptors): Perceptual Learning on the English [v]-[z] continuum (not significant).

Appendix 3.2: Comparison of language background variables between L1 English-L2 German study #1 and L1 English-L2 German study #2.

Significant difference between studies:	No significant difference between studies:
Hours of German used per week • F(1,35)= 5.395, *p*=.026*; • (4.93 vs. 8.83 hs/wk, respectively), English-German study #1 (excluding 1 participant with +/- 3 s.d.): mean: 4.93 hrs/week, range: 0.0-21.0 hrs/wk, median: 5.00 hrs/wk; s.d. 5.81; 95% CI intervals: 2.13-7.73, N=19[20-1]); • English-German study #2: mean: 8.83 hrs/week, range: 1.5-17.5 hrs/wk; median: 9.50 hrs/wk; s.d.: 4.23, 95% CI intervals: 6.73-10.94, N=18[21-3].	German perception proficiency test scores; • F(1,36)=.204; *p*=.655 (n.s.); • English-German study #1: mean: 82.78%, range: 57.89% – 100%, median: 78.89%, s.d.: 11.42%, 95% CI intervals: 77.44%-88.13%; N=20; • English-German study #2: mean: 84.14%, range: 75.00%-95.83%, median: 85.42%, s.d.: 6.03%, 95% CI intervals: 81.15%-87.14%; N=18[21-3]. [*Notes: scores are percentage points out of a possible 100%.*]
	Proficiency in German (self-assessment score) • F(1,36)=1.265, *p*= .268 (n.s.); • English-German study #1: mean[24]: 1.73, range: 1.0-3.0, median: 2.0, s.d.: .62, 95% CI intervals: 1.44-2.01; N=20; • English-German study #2: mean: 1.97, range: 1.0-3.0, median: 2.0, s.d.: .74, 95% CI intervals: 1.61-2.34; N=18[21-3].
	Age of acquisition of L2 German • F(1,36)=.158, *p*=.694; • English-German study #1: mean (in years): 15.85, range: 0-24, median: 16.5, s.d.: 5.01, 95% CI intervals: 13.51-18.19; N=20; • English-German study #2: mean: 15.17, range: 0-20, median: 16.5, s.d.: 5.61, 95% CI intervals: 12.38-17.96; N=18[21-3].
	Time in L2 environment (L2-speaking country) • F(1,34)=.050, *p*= .825; • English-German study #1: mean (in years): 4.18, range: 0.00-30.00, median: 1.0, s.d.: 9.23, 95% CI intervals: -.27 – 8.63; N=19[20-1]; • English-German study #2: mean: 4.97, range: 0.00-45.00, median: 0.00, s.d.: 11.90, 95% CI intervals: -1.15 – 11.09.)[25]

[24] Note: data coded as follows: 1=beginner, 2=intermediate, 3= advanced, 4= native speakers; choices between major labels coded with half-point steps, e.g. 2.5 = between intermediate and advanced.

	Years of L2 German study • F(1,36)=.256; *p*=.616; • English-German study #1: mean (in years): 3.05, range: .70-7.20, median: 2.30, s.d.: 2.21, 95% CI intervals: 2.02-4.09; N=20; • English-German study #2: mean: 2.67, range: .10-8.00, median: 1.80, s.d.: 2.45, 95% CI intervals: 1.45-3.89; N=18[21-3].
	Years of L2 German experience • F(1,35)=.828; *p*=.369; • English-German study #1: mean (in years): 3.29, range: .70-9.00, median: 2.60, s.d.: 2.42, 95% CI intervals: 2.13-4.46; N=19[20-1]; • English-German study #2: mean: 4.60, range: .10-18.00, median: 1.85, s.d.: 5.75, 95% CI intervals: 1.74-7.46; N=18[21-3].[26]
	Current age • F(1,36)=.995; *p*=.325; • English-German study #1: mean (in years): 21.35, range: 18.0-30.0, median: 20.50, s.d.: 3.27, 95% CI intervals: 19.82-22.88; N=20; • English-German study #2: mean: 20.28, range: 18.0-29.0, median: 19.0, s.d.: 3.36, 95% CI intervals: 18.61-21.95; N=18[21-3].
	Status in childhood as L1 English monolingual or bilingual • chi-square Test (Pearson chi-Square χ^2 = 2.54, df = 1, p = .111) • English-German study #1: 14 monolinguals, 6 bilinguals • English-German study #2: 8 monolinguals, 10 bilinguals

[25] Excluding two outliers (+/- 3 s.d.), one from each study; There was also no significant difference between the participants in the first and the second English-German study for time spent in an L2 environment (L2-speaking country) when the two outliers [+/- 3 s.d.] are *not* excluded (F(1,36)=.097, *p*= .758).

[26] Excluding one outlier (+/- 3 s.d.) from the first study; There was also no significant difference between the participants in the first and second English-German study with regards to years of L2 experience when the one outlier [+/- 3 s.d.] is *not* excluded (F(1,36)=.047; *p*=.829).

Appendix 3.3: ANCOVA's conducted for English-German study #1

1) Monolingual vs. bilingual status as child: Condition * MonoBil interaction for:
    - all three continua: $F(1,16)=.844$, $p=.372$ (n.s.)
    - English_fs continuum: $F(1,16)=3.669$, $p=.073$ (n.s.)
    - for German_fs continuum: $F(1,16)=.004$, $p=.949$ (n.s.)
    - for German_vz continuum: $F(1,16)=.495$, $p=.492$ (n.s.)

2) Hours of German used per week: Condition * HrsGerPrWk interaction for:
    - all three continua: $F(1,16)=.839$, $p=.373$ (n.s.)
    - English_fs continuum: $F(1,16)=1.391$, $p=.255$ (n.s.)
    - German_fs continuum: $F(1,16)=.689$, $p=.419$ (n.s.)
    - German_vz continuum: $F(1,16)=3.833$, $p=.068$ (n.s.)

3) German proficiency perception score:
    a. Continuous variable: Condition * GerProfScr for:
        - all three continua: $F(1,16)=1.930$, $p=.184$ (n.s.)
        - English_fs continuum: $F(1,16)=1.697$, $p=.211$ (n.s.)
        - German_fs continuum: $F(1,16)=.082$, $p=.779$ (n.s.)
        - German_vz continuum: $F(1,16)=2.940$, $p=.106$ (n.s.)
    b. Two groups (binned categorical variable): Condition * GerProfGrp for:
        - all three continua: $F(1,16)=.924$, $p=.351$ (n.s.)
        - English_fs continuum: $F(1,16)=1.070$, $p=.316$ (n.s.)
        - German_fs continuum: $F(1,16)=.126$, $p=.728$ (n.s.)
        - German_vz continuum: $F(1,16)=.814$, $p=.380$ (n.s.)

4) Self-assessment proficiency score (in L2 German): Condition *GerSelfProf for:
    - all three continua: $F(1,16)=2.202$, $p=.157$ (n.s.)
    - English_fs continuum: $F(1,16)=2.979$, $p=.104$ (n.s.)
    - German_fs continuum: $F(1,16)=.403$, $p=.534$ (n.s.)
    - German_vz continuum: $F(1,16)=1.712$, $p=.209$ (n.s.)

5) Age of Acquisition of L2 German: Condition * AoAL2 for:
    - all three continua: $F(1,16)=1.982$, $p=.178$ (n.s.)
    - English_fs continuum: $F(1,16)=.424$, $p=.524$ (n.s.)
    - German_fs continuum: $F(1,16)=1.023$, $p=.327$ (n.s.)
    - German_vz continuum: $F(1,16)=2.117$, $p=.165$ (n.s.)

6) Time spent in an L2 German environment: Condition * TimeL2Cntr for:
    - all three continua: $F(1,16)=2.432$, $p=.138$ (n.s.)
    - English_fs continuum: $F(1,16)=.136$, $p=.718$ (n.s.)
    - German_fs continuum: $F(1,16)=1.637$, $p=.219$ (n.s.)
    - German_vz continuum: $F(1,16)=2.783$, $p=.115$ (n.s.)

7) Years of L2 German study: Condition * YrsL2Study for:
   - all three continua: $F(1,16)=.695$, $p=.417$ (n.s.)
   - English_fs continuum: $F(1,16)=.320$, $p=.580$ (n.s.)
   - German_fs continuum: $F(1,16)=.018$, $p=.895$ (n.s.)
   - German_vz continuum: $F(1,16)=2.082$, $p=.168$ (n.s.)

8) Years of L2 German experience: Condition * YrsL2Exp for:
   - all three continua: $F(1,16)=1.422$, $p=.251$ (n.s.)
   - English_fs continuum: $F(1,16)=.452$, $p=.511$ (n.s.)
   - German_fs continuum: $F(1,16)=.009$, $p=.927$ (n.s.)
   - German_vz continuum: $F(1,16)=3.518$, $p=.079$ (n.s.)

9) (Current age:) Condition * AgeCurr for:
   - all three continua: $F(1,16)=1.492$, $p=.240$ (n.s.)
   - English_fs continuum: $F(1,16)=.011$, $p=.919$ (n.s.)
   - German_fs continuum: $F(1,16)=.846$, $p=.371$ (n.s.)
   - German_vz continuum: $F(1,16)=2.996$, $p=.103$ (n.s.)

Appendix 3.4: ANCOVA's conducted for English-German study #2

1) Monolingual vs. bilingual status as child: Condition * MonoBil interaction for:
   - all four continua: $F(1,14)=3.611$, $p=.078$ (n.s.)
   - English_fs continuum: $F(1,14)=1.767$, $p=.205$ (n.s.)
   - for German_fs continuum: $F(1,14)=2.069$, $p=.172$ (n.s.)
   - for German_vz continuum: $F(1,14)=.012$, $p=.916$ (n.s.)
   - English_vz continuum: $F(1,14)=3.258$, $p=.093$ (n.s.)

2) Hours of German used per week: Condition * HrsGerPrWk interaction for:
   - all four continua: $F(1,14)=1.335$, $p=.267$ (n.s.)
   - English_fs continuum: $F(1,14)=1.335$, $p=.267$ (n.s.)
   - German_fs continuum: $F(1,14)=.041$, $p=.843$ (n.s.)
   - German_vz continuum: $F(1,14)=.938$, $p=.349$ (n.s.)
   - English_vz continuum: $F(1,14)=.486$, $p=.497$ (n.s.)

3) German proficiency perception score:

   a. Continuous variable: Condition * GerProfScr for:
   - all four continua: $F(1,14)=1.402$, $p=.256$ (n.s.)
   - English_fs continuum: $F(1,14)=1.011$, $p=.332$ (n.s.)
   - English_vz continuum: $F(1,14)=1.569$, $p=.231$ (n.s.)
   - German_fs continuum: $F(1,14)=.130$, $p=.723$ (n.s.)
   - German_vz continuum: $F(1,14)=6.544$, $p=.023$ (*)

   → *The perceptual learning effect for the German v/z continuum was larger for participants with a higher score on the L2 German perception proficiency test, as illustrated in the scatterplot below.*

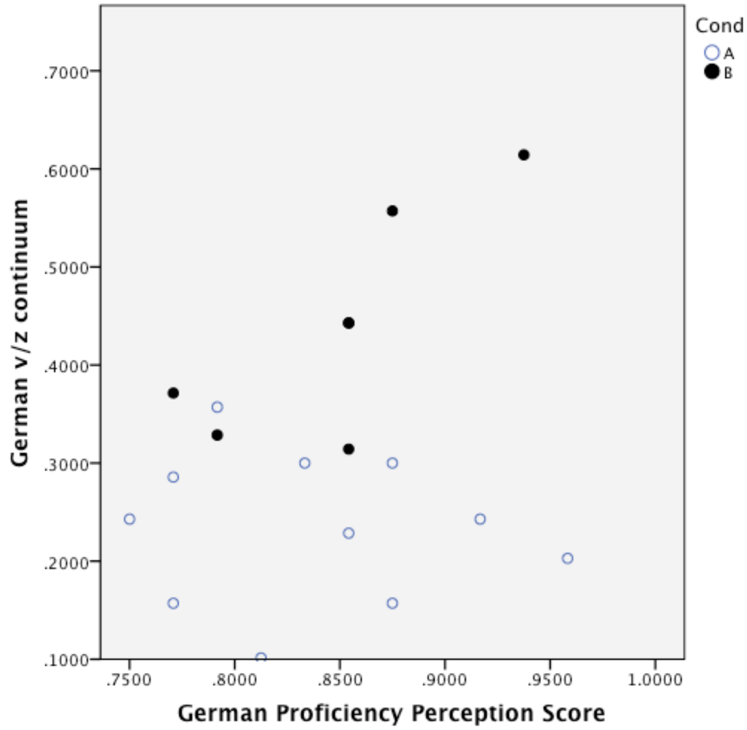

Figure 6.5: Scatterplot for the German v/z continuum (English-German study #2) with the IV condition (A/B), and the continuous covariate "L2 German Proficiency Perception Score".

b. Two groups (binned categorical variable): Condition * GerProfGrp for:
   - all four continua: $F(1,14)=.303$, $p=.591$ (n.s.)
   - English_fs continuum: $F(1,14)=2.633$, $p=.127$ (n.s.)
   - English_vz continuum: $F(1,14)=.252$, $p=.623$ (n.s.)
   - German_fs continuum: $F(1,14)=.007$, $p=.934$ (n.s.)
   - German_vz continuum: $F(1,14)=8.384$, $p=.012$ (*)

   → *The perceptual learning effect for the German v/z continuum was larger for participants with a higher score on the L2 German perception proficiency test, as illustrated in the figure below.*

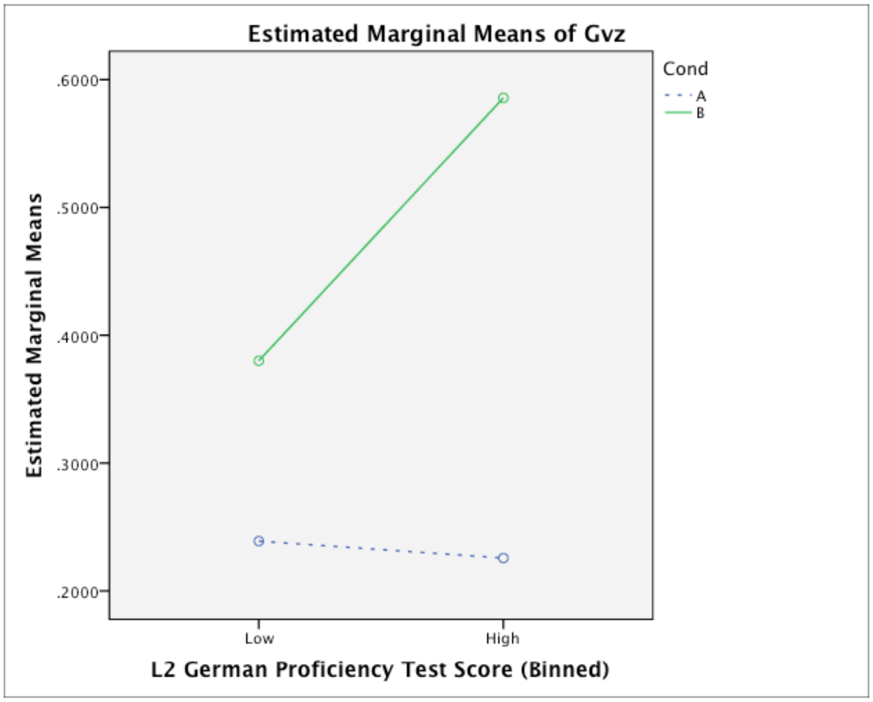

Figure 6.6: Perceptual learning effect (IV "Condition") for participants with low vs. high scores on the L2 German Proficiency Test (binned variable).

4) Self-assessment proficiency score (in L2 German): Condition *GerSelfProf for:
   - all four continua: F(1,14)=12.899, *p*=.003 (*)
   - English_fs continuum: F(1,14)=19.805, *p* =.001 (*)
   - English_vz continuum: F(1,14)=4.939, *p* =.043 (*)
   - German_fs continuum: F(1,14)=5.846, *p* =.030 (*)
   - German_vz continuum: F(1,14)=.090, *p* =.769 (n.s.)

→*As the scatterplots below illustrate, the pooled mean perceptual learning effect, and the perceptual learning effects for the English f/s continuum, the English v/z continuum and the German f/s continuum are bigger when the participants' self-assessed proficiency score in their L2 German is lower.*

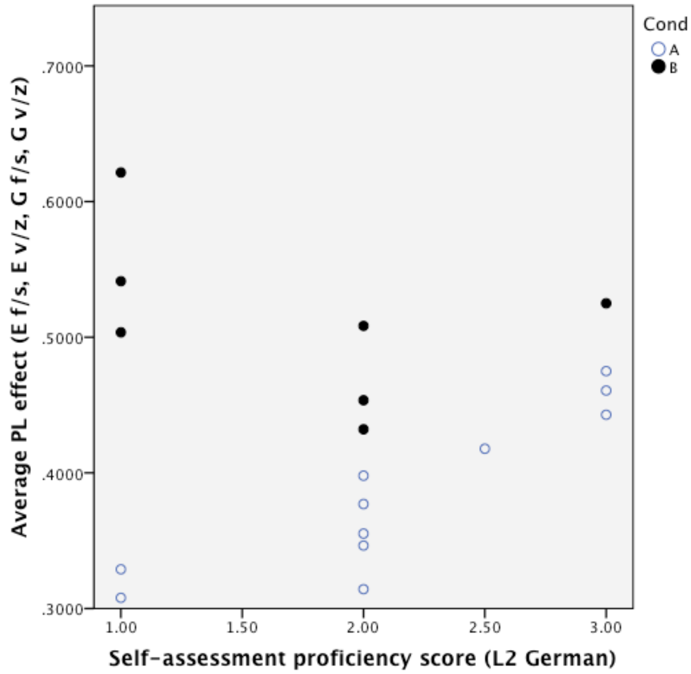

Figure 6.7: Scatterplot for the mean PL effect (pooled over all four continua of the English-German study #2) with the IV condition (A/B), and the continuous covariate "L2 German Self-assessment Score".

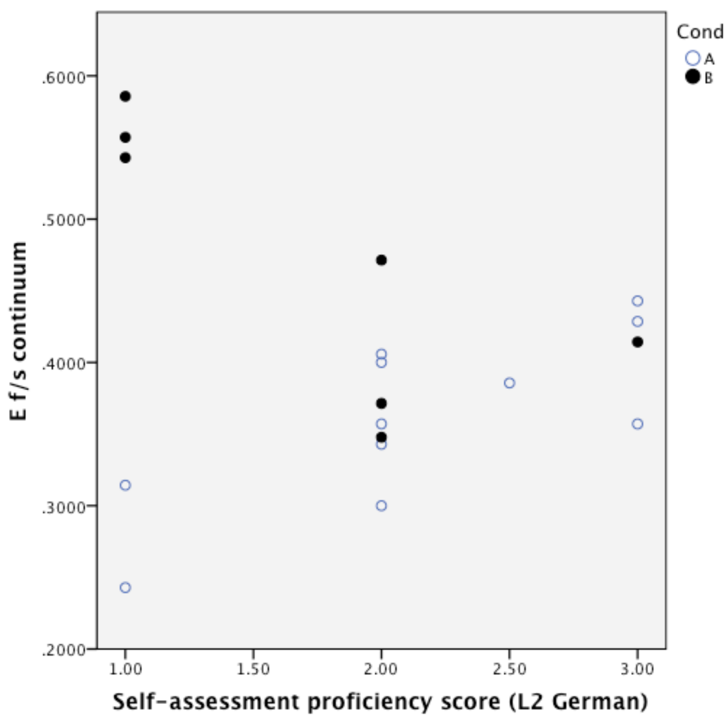

Figure 6.8: Scatterplot for the English f/s continuum (English-German study #2) with the IV condition (A/B), and the continuous covariate "L2 German Self-assessment Score".

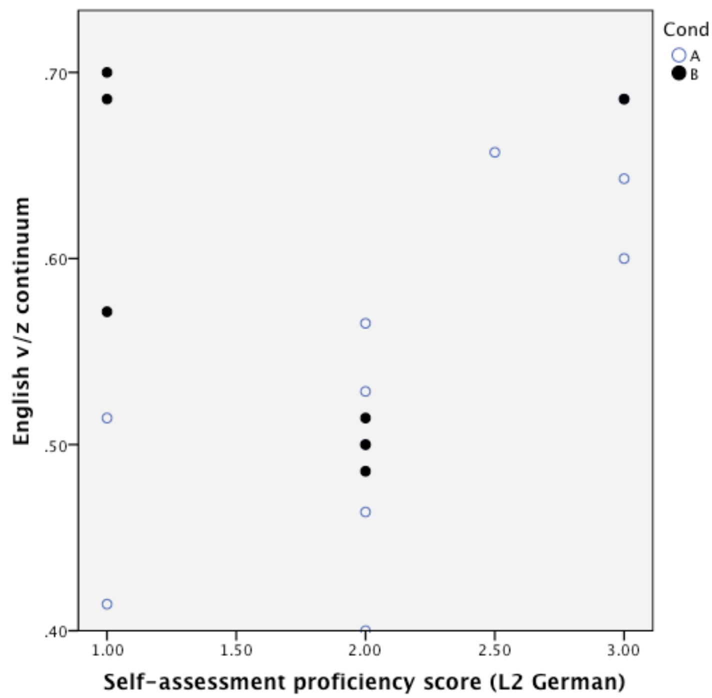

Figure 6.9: Scatterplot for the English v/z continuum (English-German study #2) with the IV condition (A/B), and the continuous covariate "L2 German Self-assessment Score".

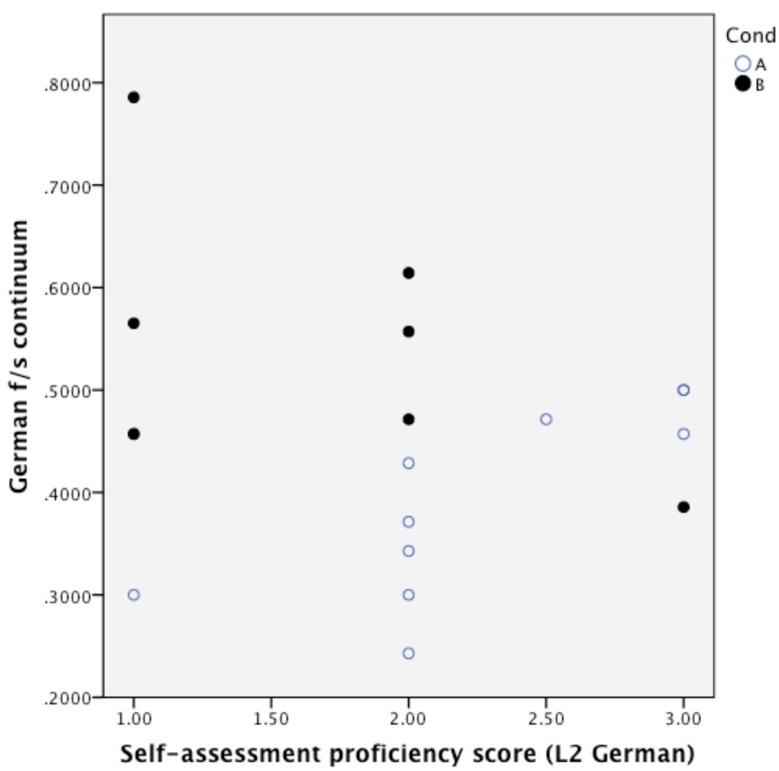

Figure 6.10: Scatterplot for the German f/s continuum (English-German study #2) with the IV condition (A/B), and the continuous covariate "L2 German Self-assessment Score".

5) Age of Acquisition of L2 German: Condition * AoAL2 for:
   - all four continua: F(1,14)=0.062 *p*=.806 (n.s.)
   - English_fs continuum: F(1,14)=.125, *p* =.729 (n.s.)
   - English_vz continuum: F(1,14)=.185, *p* =.674 (n.s.)
   - German_fs continuum: F(1,14)=.948, *p* =.347 (n.s.)
   - German_vz continuum: F(1,14)=.035, *p* =.854 (n.s.)

6) Time spent in an L2 German environment: Condition * TimeL2Cntr for:
   - all four continua: F(1,14)=.020, *p*=.889 (n.s.)
   - English_fs continuum: F(1,14)=.202, *p* =.660 (n.s.)
   - English_vz continuum: F(1,14)=1.301, *p* =.273 (n.s.)
   - German_fs continuum: F(1,14)=2.184, *p* =.162 (n.s.)
   - German_vz continuum: F(1,14)=2.053, *p* =.174 (n.s.)

7) Years of L2 German study: Condition * YrsL2Study for:
   - all four continua: F(1,14)=5.329, *p*=.037 (*)
   - English_fs continuum: F(1,14)=10.479, *p* =.006 (*)
   - English_vz continuum: F(1,14)=2.071, *p* =.172 (n.s.)
   - German_fs continuum: F(1,14)=1.132, *p* =.305 (n.s.)
   - German_vz continuum: F(1,14)=.098, *p* =.759 (n.s.)

   → *As the scatterplots in Figure 6.11 and Figure 6.12 below illustrate the perceptual learning training effect (effect of the IV ‘condition’) for all four continua pooled (mean of all four continua), and for the English f/s continua are each larger when participants have studied German for less time (continuous covariate).*

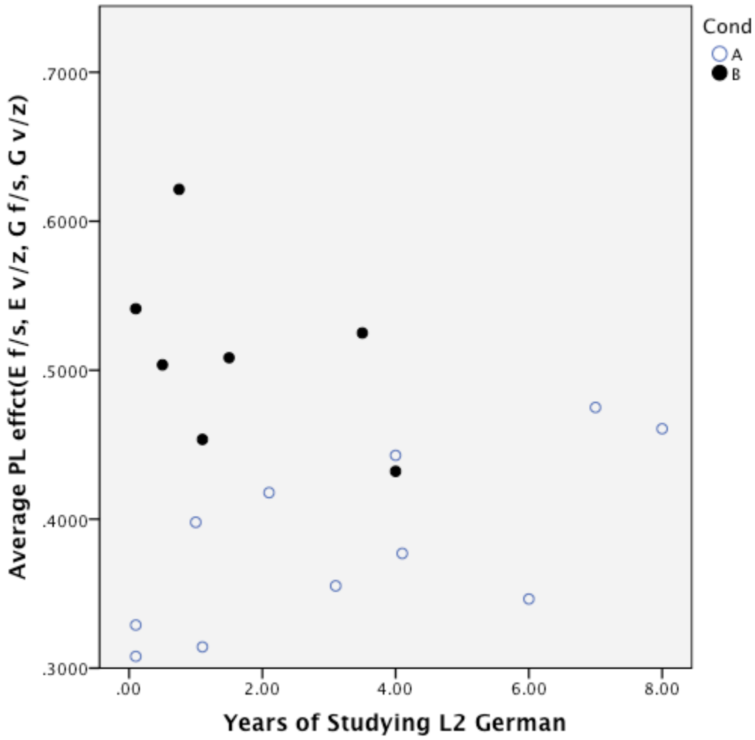

Figure 6.11: Scatterplot for the mean PL effect (pooled over all four continua of the English-German study #2) with the IV condition (A/B), and the continuous covariate "Years of Studying L2 German".

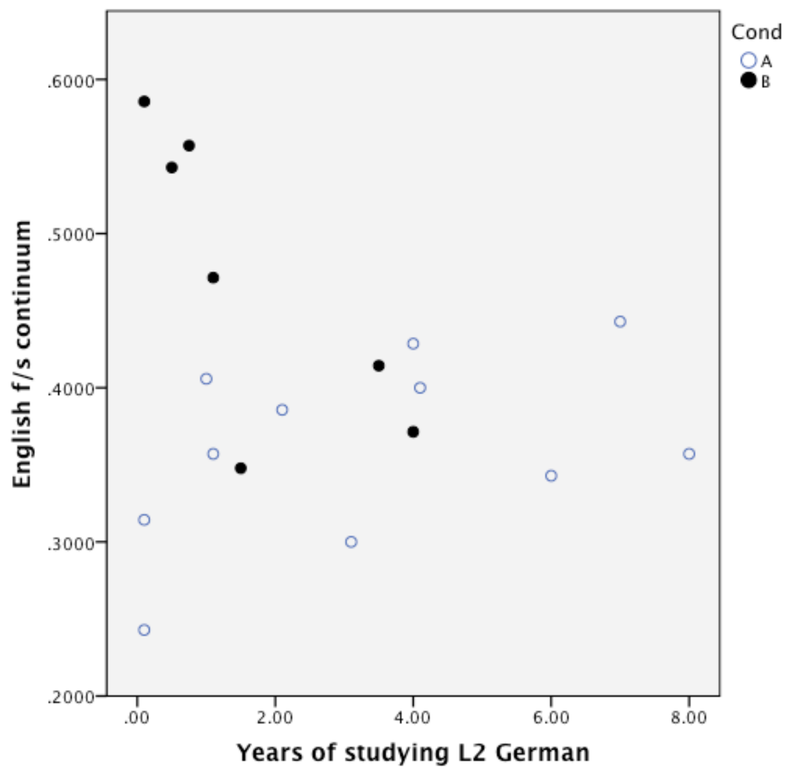

Figure 6.12: Scatterplot for the English f/s continuum (English-German study #2) with the IV condition (A/B), and the continuous covariate "Years of Studying L2 German".

8) Years of L2 German experience: Condition * YrsL2Exp for:
  - all four continua: F(1,14)=.755, *p*=.400 (n.s.)
  - English_fs continuum: F(1,14)=1.212, *p* =.290 (n.s.)
  - English_vz continuum: F(1,14)=.043, *p* =.839 (n.s.)
  - German_fs continuum: F(1,14)=2.321, *p* =.150 (n.s.)
  - German_vz continuum: F(1,14)=.013, *p* =.910 (n.s.)

9) (Current age:) Condition * AgeCurr for:
  - all four continua: F(1,14)=1.268, *p*=.279 (n.s.)
  - English_fs continuum: F(1,14)=.059, *p* =.812 (n.s.)
  - English_vz continuum: F(1,14)=.014, *p* =.907 (n.s.)
  - German_fs continuum: F(1,14)=3.389, *p* =.087 (n.s.)
  - German_vz continuum: F(1,14)=.015, *p*=.906 (n.s.)

Appendix 3.5: Language background questionnaire

1a. Date:__________________________ 1b. Location: ________________________

2a. Last name: ____

2b. First name: ____

3a. Gender: _______ 3b. Age: _________

4a. Native language: _______________________________________

4b. Did your parents speak more than one language around you in your first few years of life? __
What language(s) did your parents speak around you/to you when you were an infant/child?
____

5a. In which country/-ies did you grow up? ____

5b. In which town/state/*Bundesland* did you grow up? ___________________________________

5c. What is your dialect/regional accent? _______________________________________

6a. Mother's native language: ____

(& other languages spoken by mother: ____)

6b. Father's native language: ____
(& other languages spoken by father: ____)

7. Other languages known/studied:

Languages known/studied	**Age of acquisition**	**Number of years studied/used**	**Hrs/week *currently* exposed to each language, *on average***
a)			
b)			
c)			
d)			
e)			
f)			

8. Have you been abroad in an English-speaking country?

(or a situation where you mostly used English) Y / N, *if N skip to #9*

8a. Where: ____

8b. What year: ____

8c. Duration (in months): ____

8d. What language(s) did you speak there: ______________________________________

8e. Nature of trip: ____

(e.g. vacation/travel/au-pair/language program/student X-change/study abroad/work/business etc.)

8f. Total length of time in English-speaking countries: ______________________

9a. Currently studying in:

O undergraduate program O graduate program (Master/Doctoral)

O other: __________________

9b. Years of formal higher education since HS: _____________

10. Have you ever had any of the following:

O hearing impairment O language disability O learning disability

11. Do you know English? O Yes O No *If yes, please continue:*

12. How would you self-categorize your proficiency level in English?

a) Your proficiency in speaking English?

O Beginner O Intermediate O Advanced O Native speaker

b) Your proficiency in understanding spoken English?

O Beginner O Intermediate O Advanced O Native speaker

c) Your proficiency in reading English?

O Beginner O Intermediate O Advanced O Native speaker

d) Your proficiency in writing in English?

O Beginner O Intermediate O Advanced O Native speaker

e) Your overall proficiency level in English?

O Beginner O Intermediate O Advanced O Native speaker

[Beg =only basic functions can be performed (travel/directions/buying goods/convers. about element. topics), rudiment. grammar;
Intermed. = basic casual conversations about current topics, limited working proficiency, strong accent & many grammatical issues;
Advanced = can use language fluently, in all social and professional situations; large vocab & few gramm. errors, (slight) accent;
Native Sp =fluency in all areas (voc. & idiomatic expressions, colloquialism, cultural refer.s), could be mistaken for educated nat. sp.]

[Adapted from ILR (Interagency Language Roundtable (ILR) scale]

13a. How did you learn English? (*circle all that apply*)

O Courses in elementary, middle or high school *(circle all that apply*)

O Courses in college/graduate school

- O Private instructor/tutor
- O TV/movies/radio; technology/computer software
- O Self-taught: language books, tapes etc.
- O No formal instruction: I picked it up from family members/friends around me
- O Time abroad
- O Other: ____

14a. Total years/semesters of formal/informal *study* of English: ___________________

14b. Did you use English continuously once you started learning it? Y / N

14c. Total years/semesters of *using* English: ___________________________

14d. Do you *currently use* English? Y / N *If YES, continue:*

14e. How many hrs/week do you currently use English, on average? _________

14f. How/Where/When do you currently use English? _____________________

____

15. How often are you currently *exposed* to English in the following contexts:

Context	**Exposure to English:** *Hrs/week (or: daily, X times a week, several times a month, rarely, never)*		**Context**	**Exposure to English:** *Hrs/week (or: daily, X times a week, several times a month, rarely, never)*
Among family members			In the classroom (*besides lg. courses)*	
Among friends			TV, Radio, Music	
In a language course			Reading	
Other:				

Appendix 4.1: Vocabulary questionnaire for the twenty critical /s/-stimuli

For each word below, please indicate:

<u>How sure are you about using each of the words below in a written sentence</u>?

Indicate your answer on a scale from 0 (completely unsure) to 5 (definitely sure).

	Word	How sure are you about using this word in a sentence? *0 (completely unsure)* ←→ *5 (completely sure)*
1	accuracy	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
2	aerosol	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
3	Arkansas	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
4	chromosome	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
5	coliseum	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
6	condensation	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
7	condescend	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
8	connoisseur	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
9	cul-de-sac	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
10	delicacy	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
11	democracy	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
12	dinosaur	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
13	embassy	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
14	eraser	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
15	indecision	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
16	Johnson	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
17	legacy	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
18	medicine	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
19	reconcile	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
20	rehearsal	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*

Appendix 4.2: Vocabulary questionnaire for the twenty critical /f/-stimuli

For each word below, please indicate:

How sure are you about using each of the words below in a written sentence?

Indicate your answer on a scale from 0 (completely unsure) to 5 (definitely sure).

	Word	How sure are you about using this word in a sentence? *0 (completely unsure)* ←→ *5 (completely sure)*
1	amphibian	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
2	beneficial	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
3	cacophony	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
4	calligraphy	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
5	chlorophyll	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
6	clarification	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
7	daffodil	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
8	endorphin	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
9	gleeful	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
10	glorification	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
11	manufacture	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
12	meaningful	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
13	microphone	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
14	modification	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
15	Newfoundland	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
16	orthography	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
17	perform	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
18	qualification	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
19	qualify	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*
20	unofficial	*(completely unsure)* 0 1 2 3 4 5 *(completely sure)*

References

Akahane-Yamada, Reiko, Ann R Bradlow, David B Pisoni & Yoh'ichi Tohkura. 1997. Effects of audio‐visual training on the identification of English/r/and/l/by Japanese speakers. *The Journal of the Acoustical Society of America* 102. 3137-3137.

Baese-Berk, M. M., A. R. Bradlow & B. A. Wright. 2013. Accent-independent adaptation to foreign accented speech. *J Acoust Soc Am* 133. EL174-80.

Beebe, Leslie M. 1981. Social and situational factors affecting the strategy of dialect codeswitching. *International Journal of the sociology of language* 32. 139-149.

Behrens, Susan J. & Sheila E. Blumstein. 1988. Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics* 16. 295-298.

Berwick, Robert C. 1985. *The acquisition of syntactic knowledge*. Cambridge, Mass.: MIT Press.

Best, Catherine T. & Michael D. Tyler. 2007. Nonnative and second language speech perception. Commonalities and complementarities. In M. J. Munro & Ocke-Schwen Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production*, 13-34. Amsterdam: John Benjamins.

Bialystok, Ellen. 2001. *Bilingualism in development: Language, literacy, and cognition*. New York, NY: Cambridge University Press.

Boersma, Paul & David Weenink. 2011. Praat: Doing Phonetics by Computer. [Computer program]. Retrieved from http://www.praat.org/

Bradlow, Ann R. & Tessa Bent. 2008. Perceptual adaptation to non-native speech. *Cognition* 106. 707-29.

Broselow, Ellen, Su-I Chen & Chilin Wang. 1998. The Emergence of the Unmarked in Second Language Phonology. *Studies in Second Language Acquisition* 20. 251-280.

Brown, Amanda & Marianne Gullberg. 2010. Bidirectional cross-linguistic influence in event conceptualization? Expressions of Path among Japanese learners of English. *Bilingualism: Language and Cognition* 14. 79-94.

Caramazza, Alfonso, G.H. Yeni-Komshian, E. B. Zurif & E. Carbone. 1973. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America* 54. 421-428.

Chang, Charles B. 2010. *First Language Phonetic Drift During Second Language Acquisition*. Doctoral dissertation, UC Berkeley.

Chang, Charles B. 2011. Systemic drift of L1 vowels in novice L2 learners. *ICPhS XVII* 428-431.

Chang, Charles B. 2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics* 40. 249-268.

Chang, Charles B. 2013. A novelty effect in phonetic drift of the native language. *Journal of Phonetics* 41. 520-533.

Cook, Vivian. 2003. *Effects of the Second Language on the First*, vol. 3. Tonawanda, NY: Multilingual Matters.

Cook, Vivian J. 1991. The poverty-of-the-stimulus argument and multicompetence. *Second Language Research* 7. 103-117.

Cook, Vivian J. 1992. Evidence for Multicompetence. *Language Learning* 42. 557-591.

Cook, Vivian J. 1995. Multi-competence and the learning of many languages. *Language, Culture and Curriculum* 8. 93-98.

Costa, Albert. 2009. Lexical access in bilingual production. *Handbook of Bilingualism: Psycholinguistic Approaches: Psycholinguistic Approaches*, 308.

Costa, Albert, Martin J. Pickering & Antonella Sorace. 2008. Alignment in second language dialogue. *Language and Cognitive Processes* 23. 528-556.

Cristia, Alejandrina, Jeff Mielke, Robert Daland & Sharon Peperkamp. 2013. Similarity in the generalization of implicitly learned sound patterns. *Lab Phonol* 4.

Cutler, Anne. 2012. *Native Listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.

de Bot, K. & M. Clyne. 1994. A 16-year Longitudinal Study of Language Attrition in Dutch Immigrants in Australia. . *Journal of Multilingual and Multicultural Development* 15. 17-28.

De Bot, Kees. 2004. The multilingual lexicon: modeling selection and control. *International Journal of Multilingualism* 1. 17-32.

De Bot, Kees, Wander Lowie & Marjolijn Verspoor. 2005. *Second Language Acquisition. An advanced resource book*. London and New York: Routledge.

Delattre, P.C., A.M. Liberman & F.S. Cooper. 1964. Formant transitions and loci as acoustic correlates of place of articulation in American fricatives. *Studia Linguistica* 16. 104-121.

Dijksta, Ton. 2005. Bilingual visual word recognition and lexical access. *Handbook of Bilingualism: Ppsycholinguistic Approaches* 54. 179-201.

Eisner, Frank & James M. McQueen. 2005. The specificity of perceptual learning in speech processing. *Percept Psychophys* 67. 224-38.

Escudero, Paola. 2005. *Linguistic Perception and Second Language Acquisition.Explaining the attainment of optimal phonological categorization*. Utrecht: LOT Dissertation Series.

Escudero, Paola. 2007a. Multilingual sound perception and word recognition. *Stem, Spraak en Taalpathologie* 93-103.

Escudero, Paola. 2007b. Second-language phonology: The role of perception. *Phonology in context* 109-134.

Escudero, Paola. 2009. Linguistic Perception of "similar" L2 sounds. In Paul & Silke Hamann Boersma (ed.), *Phonology in Perception*. Berlin: Mouton de Gruyter.

Escudero, Paola & Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26. 551-585.

Fant, C. & M. Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.

Flege, James Emil. 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47-65.

Flege, James Emil. 1992. Speech learning in a second language. In C. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, and Application*, 565-604. Timonium, MD: York Press.

Flege, James Emil. 1995a. The Phonetic Study of Bilingualism. In G. Bloothooft, V. Hazan, D. Huber & J. Llisterri (eds.), *European Studies in Phonetics and Speech Communication*, 98-103.

Flege, James Emil. 1995b. Second language speech learning. Theory, findings, and problems. In Winfred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233-277. Timonium, MD: York Press.

Flege, James Emil. 2002. Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske & A. Rohde (eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*, 217-244. Trier: Wissenschaftlicher Verlag.

Flege, James Emil. 2003. Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (eds.), *Phonetics and Phonology in Language*

Comprehension and Production, Differences and Similarities, 319-355. Berlin: Mouton de Gruyter.

Flege, James Emil. 2007. Language contact in bilingualism: Phonetic system interactions. In J. Cole & J.I. Hualde (ed.), *Lab Phonol*, 353-382. Berlin: Mouton de Gruyter.

Flege, James Emil & Wieke Eefting. 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication* 6. 185-202.

Flege, James Emil & James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America* 76. 708-721.

Flege, James Emil, Carlo Schirru & Ian R. A. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40. 467-491.

Fowler, Carol A., Valery Sramko, David J. Ostry, Sarah A. Rowland & Pierre Hallé. 2008. Cross language phonetic influences on the speech of French–English bilinguals [10//]. *Journal of Phonetics* 36. 649-663.

Ganong, William F. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6. 110-125.

Goldrick, Matthew. 2004. Phonological features and phonotactic constraints in speech production. *J Mem Lang* 51. 586-603.

Goldstone, Robert L. 1998. Perceptual learning. *Annual review of psychology* 49. 585-612.

Green, David W. 1986. Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language* 27. 210-223.

Grosjean, François. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36. 3-15.

Grosjean, François. 1997. Processing mixed language: issues, findings, and models. In A. de Groot & J. Kroll (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives*. Mahwah, NJ: LEA.

Grosjean, François. 2008. The Léwy and Grosjean BIMOLA Model. *Studying Bilinguals*, 201-210. New York: Oxford University Press.

Grosjean, François & Carlos Soares. 1986. Processing mixed language: Some preliminary findings. In Annette M.B. de Groot & Judith Kroll (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives*, 145-179. Hillsdale, NJ: Lawrence Erlbaum.

Guion, Susan G. 2003. The vowel systems of Quichua-Spanish bilinguals. *Phonetica* 60. 98-128.

Gullberg, M. 2009. Reconstructing verb meaning in a second language: How English speakers of L2 Dutch talk and gesture about placement. *Annual Review of Cognitive Linguistics* 7. 221-245.

Gullberg, Marianne. 2011. 8 Language-specific encoding of placement events in gestures. In J. Bohnemeyer & E. Pederson (eds.), *Event representation in language and cognition*, vol. 11, 166-188. Cambridge: Cambridge University Press.

Gullberg, Marianne. 2013. The Handbook of Bilingualism and Multilingualism, second edition (Blackwell Handbooks in Linguistics). In K. Bhatia Tej & C. Ritchie William (eds.), 417-437. Chichester, U.K.: Wiley–Blackwell Publishers.

Haigh, C. A. & D. Jared. 2007. The activation of phonological representations by bilinguals while reading silently: evidence from interlingual homophones. *J Exp Psychol Learn Mem Cogn* 33. 623-44.

Halle, M., G.W. Hughes & J-P. A. Radley. 1957. Acoustic properties of stop consonants. *Journal of the Acoustical Society of America* 29. 107-116.

Harris, Katherine Safford. 1958. Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech* 1. 1-7.

Heinz, John M. & Kenneth N. Stevens. 1961. On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America* 589-596.

Heredia, R. R. & J. M. Brown. 2012. Bilingual Memory. In T. K. Bhatia & W. C. Ritchie (eds.), *The Handbook of Bilingualism and Multilingualism, Second Edition*, 225-248. Chichester, UK: John Wiley & Sons.

Hughes, G. W. & M. Halle. 1956. Spectral properties of fricative consonants. *Journal of the Acoustical Society of America* 28. 303-310.

Jared, Debra & Judith F. Kroll. 2001. Do bilinguals activate phonological representations in one or both of their languages when naming words? *J Mem Lang* 44. 2-31.

Jarvis, Scott. 2003. Probing the Effects of the 12 on the LI: A Case Study. In Vivian Cook (ed.), *Effects of the Second Language on the First*, vol. 3, 81-102. Tonawanda, NY: Multilingual Matters.

Jesus, Luis M. T. & Christine H. Shadle. 2002. A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics* 30. 437–464.

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108. 1252-1263.

Ju, Min & Paul A. Luce. 2004. Falling on Sensitive Ears Constraints on Bilingual Lexical Activation. *Psychological Science* 15. 314-318.

Juslin, Peter, Henrik Olsson & Anna-Carin Olsson. 2003. Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General* 132. 133-156.

Kim, M., W. S. Horton & A. R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance [May 3]. *Lab Phonol* 2. 125-156.

Köpke, Barbara, Monika S Schmid, Merel Keijzer & Susan Dostert. 2007. *Language attrition: Theoretical perspectives*, vol. 33. John Benjamins Publishing.

Kraljic, T., A. G. Samuel & S. E. Brennan. 2008. First impressions and last resorts: how listeners adjust to speaker variability. *Psychol Sci* 19. 332-8.

Kraljic, Tanya & Arthur G. Samuel. 2005. Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology* 51. 141-78.

Kraljic, Tanya & Arthur G. Samuel. 2006. Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review* 13. 262-268.

Kraljic, Tanya & Arthur G. Samuel. 2007. Perceptual adjustments to multiple speakers. *J Mem Lang* 56. 1-15.

Kroll, Judith F., Paola E. Dussias, Cari A. Bogulski & Jorge R. Valdes Kroff. 2012. Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. *Psychology of Learning and Motivation-Advances in Research and Theory* 56. 229-262.

Kuhl, Patricia K. 2000. A new view of language acquisition. *Proceedings of the National Academy of Sciences* 97. 11850-11857.

Kuhl, Patricia K. 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics* 50. 93-107.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages.* Oxford: Blackwell.

Léwy, N. & François Grosjean. 1997. A computational model of bilingual lexical access. Ms.

Li, Ping & Igor Farkaš. 2002. A self-organizing connectionist model of bilingual processing. In R. Heredia & J. Altarriba (eds.), *Bilingual Sentence Processing*. North Holland: Elsevier Science Publisher.

Liberman, Alvin M., Franklin S. Cooper, Donald P. Shankweiler & Michael Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74. 431-461.

Liberman, Alvin M., Katherine S. Harris, Howard S. Hoffman & Belver C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54. 358.

Maddieson, Ian. 1984. *Patterns of sounds*. New York: Cambridge University Press.

Marian, Viorica, Michael Spivey & Joy Hirsch. 2003. Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language* 86. 70-82.

Martínez-Celdrán, Eugenio, Ana Ma Fernández-Planas & Josefina Carrera-Sabaté. 2003. Castilian Spanish. *Journal of the International Phonetic Association* 33. 255-259.

Maye, J. & L. A. Gerken. 2001. Learning phonemes: How far can the input take us? Paper presented at Proceedings of the 21st Boston University Conference on Language Development., Boston, MA.

Maye, J., R. N. Aslin & M. K. Tanenhaus. 2008. The weckud wetch of the wast: lexical adaptation to a novel accent. *Cogn Sci* 32. 543-62.

Maye, J., D. J. Weiss & R. N. Aslin. 2008. Statistical phonetic learning in infants: facilitation and feature generalization. *Dev Sci* 11. 122-34.

Maye, Jessica & Daniel Weiss. 2003. Statistical Cues Facilitate Infants' Discrimination of Difficult Phonetic Contrasts. In BarbaraBeachley et al. (ed.), *BUCLD 27 Proceedings*, 508-518. Somerville, MA: Cascadilla Press.

McClaskey, Cynthia L., David B. Pisoni & Thomas D. Carrell. 1983. Transfer of training of a new linguistic contrast in voicing. *Perception & Psychophysics* 34. 323-330.

McReynolds, Leija V & Shannon Bennett. 1972. Distinctive feature generalization in articulation training. *Journal of Speech and Hearing Disorders* 37. 462-470.

Mitterer, Holger. 2011 (accessed). Interpolation Script. [Computer program]. Retrieved from http://www.holgermitterer.eu/HM/interpolate between voiced sounds.praat

Montrul, Silvina A. 2008. *Incomplete acquisition in bilingualism: Re-examining the age factor*, vol. 39. Philadelphia, PA: John Benjamins Publishing.

Montrul, Sylvia. 2013. incomplete L1 acquisition. In Julia Rogers Herschensohn & Martha Young-Scholten (eds.), *The Cambridge handbook of second language acquisition*, 353-371. Cambridge: Cambridge University Press.

Murphy, Victoria A. & Karen J. Pine. 2003. L2 influences on L1 linguistic representations. In Vivian Cook (ed.), *Effects of the Second Language on the First*, 142-167. Clevedon, UK: Multilingual Matters.

Nas, Gerard. 1983. Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior* 22. 526-534.

Nielsen, Kuniko. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39. 132-142.

Nishi, Kanae & Diane Kewley-Port. 2007. Training Japanese Listeners to Perceive American English Vowels: Influence of Training Sets. *Journal of Speech, Language, and Hearing Research* 50. 1496-1509.

Nishi, Kanae & Diane Kewley-Port. 2008. Nonnative Speech Perception Training Using Vowel Subsets: Effects of Vowels in Sets and Order of Training. *Journal of Speech, Language, and Hearing Research* 51. 1480-1493.

Norris, Dennis, James M. McQueen & Anne Cutler. 2003. Perceptual learning in speech. *Cognitive Psychology* 47. 204-238.

Obler, Lorraine. 1982. The parsimonious bilingual. In L. Obler & L. Menn (eds.), *Exceptional language and linguistics*, 339-346. New York: Academic Press.

Pallier, Christophe, Laura Bosch & Núria Sebastián-Gallés. 1997. A limit on behavioral plasticity in speech perception. *Cognition* 64. B9-B17.

Pallier, Christophe, Angels Colomé & Núria Sebastián-Gallés. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12. 445-449.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119. 2382.

Pavlenko, Aneta 1999. New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition* 2. 209-230.

Peperkamp, Sharon & Emmanuel Dupoux. 2007. Learning the mapping from surface to underlying representations in an artificial language. In J. Cole & J. Hualde (ed.), *Laboratory Phonology 9*, 315-338. Berlin: Mouton de Gruyter.

Pickering, Martin J & Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences* 27. 169-190.

Pisoni, David B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics* 13. 253-260.

Reinisch, Eva & Lori L. Holt. 2014. Lexically guided phonetic retuning of foreign-accented speech and its generalization. *J Exp Psychol Hum Percept Perform* 40. 539-55.

Reinisch, Eva, Andrea Weber & Holger Mitterer. 2013. Listeners retune phoneme categories across languages. *J Exp Psychol Hum Percept Perform* 39. 75-86.

Samuel, Arthur G. 1981a. Phonemic restoration: insights from a new methodology. *Journal of Experimental Psychology: General* 110. 474.

Samuel, Arthur G. 1981b. The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance* 7. 1124.

Sancier, Michele L & Carol A Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25. 421-436.

Schmid, Monika S. 2004. *First language attrition: Interdisciplinary perspectives on methodological issues*, vol. 28. John Benjamins Publishing.

Schuhmann, Katharina. 2012. Perceptual learning in Hindi-English bilinguals. *Studies in the Linguistic Sciences: Illinois Working Papers* 81-99.

Schwartz, Ana I. & Judith F. Kroll. 2006. Bilingual lexical activation in sentence context. *J Mem Lang* 55. 197-212.

Sebastián-Gallés, Núria & Judith F. Kroll. 2003. Phonology in Bilingual Language Processing: Acquisition, Perception, and Production. In A. Meyer & N. Schiller (eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, 279-318. Berlin: Mouton de Gruyter.

Sebastian-Gallés, Núria, Antoni Rodríguez-Fornells, Ruth de Diego-Balaguer & Begoña Díaz. 2006. First- and Second-language Phonological Representations in the Mental Lexicon. *J Cogn Neurosci* 18. 1277-1291.

Shadle, Christine H. 2012. On the Acoustics and Aerodynamics of Fricatives. In Abigail C. Cohn, Cécile Fougeron & Marie K. Huffman (eds.), *The Handbook of Laboratory Phonology*, 511-526. Oxford University Press.

Sidaras, S. K., J. E. Alexander & L. C. Nygaard. 2009. Perceptual learning of systematic variation in Spanish-accented speech [May]. *J Acoust Soc Am* 125. 3306-16.

Sjerps, M. J. & J. M. McQueen. 2010. The bounds on flexibility in speech perception. *J Exp Psychol Hum Percept Perform* 36. 195-211.

Sperbeck, Mieko, Winifred Strange & Kikuyo Ito. 2005. Training Japanese L2 learners to perceive difficult American vowel contrasts. *The Journal of the Acoustical Society of America* 117. 2400-2400.

Spivey, Michael J. & Viorica Marian. 1999. Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 10. 281-284.

Sridhar, Shikaripur N. & Kamal K. Sridhar. 1980. The syntax and psycholinguistics of bilingual code mixing. *Canadian Journal of Psychology/Revue canadienne de psychologie* 34. 407.

Stevens, Kenneth. 1971. Airflow and turbulence noise for fricative and stop consonants: static considerations. *Journal of the Acoustical Society of America* 50. 1180-11192.

Stevens, Kenneth N. & Sheila E. Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64. 1358-1368.

Stevens, Peter. 1960. Spectra of fricative noise in human speech. *Language and Speech* 3. 32-49.

Sundara, Megha & Linda Polka. 2008. Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition* 106. 234-258.

Sussman, Harvey M., Helen A. McCaffrey & Sandra A. Matthews. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America* 90. 1309-1325.

Tabachnick, Barbara G. & Linda S. Fidell. 2007. *Experimental designs using ANOVA*. Belmont, CA: Duxbury.

Thomson, Ron I. 2012. Improving L2 Listeners' Perception of English Vowels: A Computer-Mediated Approach. *Language Learning* 62. 1231-1258.

Tomiak, Gail R. 1990. *An Acoustic and Perceptual Analysis of the Spectral Moments Invariant with Voiceless Fricative Obstruents*. Doctoral dissertation, SUNY Buffalo.

Van Hell, Janet G & Ton Dijkstra. 2002. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review* 9. 780-789.

Warren, Richard M. 1970. Perceptual restoration of missing speech sounds. *Science* 167. 392-393.

Watson, I. 1995. The effect of bilingualism on the acquisition of perceptual categories underlying the voicing contrast. Paper presented at Proceedings of the 13th International Congress of Phonetic Sciences

Werker, J. F., K. Byers-Heinlein & C. T. Fennell. 2009. Bilingual beginnings to learning words. *Philos Trans R Soc Lond B Biol Sci* 364. 3649-63.

Wexler, Kenneth & Rita Manzini. 1987. Parameters and learnability in binding theory. In Thomas Roeper & Edwin Williams (eds.), *Parameter Setting*. Dordrecht: Reidel.

White, J. 2014. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130. 96-115.

Williams, Lee. 1977. The perception of consonant voicing by Spanish English bilinguals. *Perception & Psychophysics* 21. 289-297.

Williams, Lee. 1979. The modification of speech perception and production in second-language learning. *Perception & Psychophysics* 26. 95-104.

Young-Scholten, M. 1994. On positive evidence and ultimate attainment in L2 phonology. *Second Language Research* 10. 193-214.

Zając, Magdalena. 2013. [ˈberɚ] or [ˈbetə]? Do Polish Learners of English Accommodate their Pronunciation? A Pilot Study. 229-239.

Zampini, Mary L. & Kerry P. Green. 2001. The voicing contrast in English and Spanish: The relationship between perception and production. In J. Nicol (ed.), *One mind, two languages: Bilingual language processing*, 23-48. Cambridge, MA: Blackwell.

Zeno, Susan M., Stephen H. Ivens, Robert T. Millard & Raj Duvvury. 1995. *The Educator's Word Frequency Guide*. Brewster, NY: Touchstone Applied Science Associates.