Accent effects on the recognition of speech in noise: Second-language proficiency, accent similarity and adaptation.

Mélanie C.C. Pinet

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

UCL Division of Psychology and Language Sciences,
Department of Speech, Hearing and Phonetic Sciences

2012

Declaration

I, Mélanie C.C. Pinet, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Portions of this thesis have been published in peer-reviewed journals and conference proceedings.

Acknowledgements

First and foremost I would like to sincerely thank my primary supervisor, Paul Iverson for his dedicated support throughout my PhD and help with my application to obtain funding for my research. Paul has been a great source of advice and inspiration for this thesis, but also helped me to develop my research career by contributing to the publication of several articles that were outcomes of my PhD research and our collaborative side work. I am extremely grateful for his help and moral support, and hope to work with him for many more years. My secondary supervisor, Bronwen Evans has equally been an invaluable source of support, both theoretically and morally, especially in the last stages of my PhD and I am extremely grateful for her help and advice, as well as for the teaching experience she offered me.

There are many people in the Department of Speech, Hearing and Phonetic Sciences who I would like to thank. Stuart Rosen has offered a lot of technical support for this thesis and the other projects we collaborated on, and is a great source of inspiration. I am also very grateful for his moral support and kindness throughout my PhD. Valerie Hazan has continuously offered theoretical advice and great insights on my work, as well as providing materials for this thesis and I am very grateful for her advice. Mark Huckvale has collaborated with Paul and I on a published section of this thesis by providing computational support, which has shaped a major part of my research, and I am very thankful for his contribution. Andrew Faulkner and Souhila Messaoud-Galusi have provided statistical advice, and Gaston Hilkhuysen technical advice with noise generation. I am also extremely grateful to Katrin Skorrupa, Lorna Halliday, Rachel Baker, Anita Wagner, Michael Ashby, Michèle Pettinato, Jirky

Tuomainen, Páraic Scanlon and Joanna Przedlacka who always had kind words of advice. I must also thank David Cushing and Steve Nevard who have provided unconditional technical support for experiment set up and recordings; their knowledge and patience is remarkable, and I am extremely grateful to them for all their support, especially in my ‘panic moments’! The admin team is equally exceptional, and have helped me in many ways.

I would also like to thank my fellow PhD students who had to put up with my moments of madness for so many years, even though I outlived most of them due to the structure of my PhD. They all have been great support and I enjoyed providing support for them. Amongst them, I must particularly thank Kayoko Yanagisawa, Jeong Kyong, Young Shin Kim, Yasna Pereira, Chierh Cheng, Jin Shin, Kathleen McCarthy, Yasuaki Shinohara, Mark Wibrow, Wafa’a Alshangiti, Tim Schoof, Katherine Mair, Sophie Gates, Kota Hattori, Sam Evans, Nada Sari, José Joaquín Atria and Sonia Granlund.

Researchers outside of UCL have also been a great source of inspiration for my research through our exchanges and collaborations, and I must particularly thank Volker Dellwo, Patti Adank, Ann Bradlow, Tessa Bent, Cathi Best, Alexis Hervais-Adelman for their valuable inspirational advice.

I must thank the hundreds of participants who took part in my experiments, both for this PhD research and for my other research projects. I must say that without them, this thesis would have never been. They all have painfully sat through long experiments, listening to what sounded to them like weird distorted speech, and most of them did this in exchange for home baked cakes. I am extremely grateful for them giving up their time, travelling to the lab and talking to me, that weird French tattooed lady. All of my participants were very dedicated to this research and there are no words to describe how grateful I am to them.

My friends and family have been an invaluable source of support during the course of my PhD. I want to thank my family in France where I set my temporary lab to conduct my experiments on French speakers. My sister, Noémie has taken part in the majority of my experiments by helping me to recruit participants, taking part in recordings, piloting and coming to the UK during her holidays to conduct an experiment. I am very grateful for all her help. My friends have been a great source of anecdotal inspiration and moral support, and I must particularly thank Marie and Ghis Melou for their support. Most importantly, I am very grateful to my boyfriend, Sam Green, who had to put up with my general grumpiness in the last year of my PhD. He kept me physically alive, helped me through my illness, and his moral support has helped me submit this thesis on time. There are not enough cakes I could bake for him to thank him. I would also like to thank Kasia Ludka, my loving flatmate who isn't sick of me yet, and fed me nice Polish food for the past three years.

Finally, I would to thank my cats, Coucouille, Ninja and Marquise for their daily entertainment and love. They have been a big moral support and kept me sane through this process. I must warn the readers that they have made several attempts to write parts of this thesis by repeatedly jumping on the keyboard, so any typos should be blamed on them.

In loving memory of my father, Thierry Pinet, 1960-1990. I may not have become a gymnast or a professional ballerina, but instead I did that, I hope he is proud of me.

Abstract

One of the key factors that determine speech intelligibility under challenging conditions is the difference between the accents of the talker and listener. For example, normal-hearing listeners can be accurate at recognizing a wide range of accents in quiet, but in noise they are much poorer (e.g., 20 percentage points less accurate) if they try to understand native (L1) or non-native (L2) accented speech that does not closely match their own accent. The aim of this PhD research is to provide a more detailed account of this talker-listener interaction in order to establish the underlying factors involved in L1 and L2 speech communication in noise for normal-hearing populations. Study 1 examined the effects of L2 proficiency on the L1-L2 accent interaction in noise, with Study 2 investigating the contribution of acoustic similarity to accent intelligibility. Study 3 examined L1 listeners' adaptation processes to unfamiliar accents in noise. Finally, Study 4 took a cross-linguistic approach and investigated how language experience and accent similarity affect the talker-listener accent interaction in noise across languages. Overall, the results revealed that several factors contribute strongly to the L1-L2 accent interaction in noise, with the emerging findings contributing to our general understanding of speech in noise perception. For instance, acoustic similarity in the accents of the talkers and the listeners accounted for a great amount of the variance in intelligibility. Linguistic background and L2 experience were also shown to play a major role in the interaction, shaping the listeners' accent processing patterns in their L1 and L2, as well as general speech-in-noise processes, with bilingual and highly proficient L2 listeners showing facilitation effects for speech processing in both their languages. Finally, the selective tuning processes found for standard accents in English

Accent effects on the recognition of speech in noise

were not replicated for French, indicating that accent processing varies across languages.

Table of Contents

Figures

Tables

1. Chapter one: Introduction

Several factors have been identified to degrade speech communication, with noise and accent widely reported to affect intelligibility. Noise has been shown to be particularly troublesome for individuals with hearing impairments (e.g., Dubno et al., 1984; Plomp, 1978). Likewise, accent differences represent another type of degradation that has been shown to affect hearing-impaired listeners (e.g., Gordon-Salant et al., 2010) and older listeners (Adank and Janse, 2010). Studying the processing of accented speech in noise is more reflective of realistic communicative situations, compared to "lab speech" (i.e., standard-accented speech presented in quiet listening conditions) and therefore reveals speech processes that are typical of every day talker-listener interactions. Indeed, listeners use very different mechanisms to process speech in quiet compared to noisy listening conditions, and while individuals can recognize speech in quiet with high accuracy, adding noise to the signal reveals underlying cognitive mechanisms and listening strategies designed to overcome degraded listening conditions. In addition, in quiet listening conditions, accent acts as a social marker and stressor of L2 proficiency and listeners can recognize speech with little difficulty, even when the phonetic content diverges from the native form of the language spoken. In noise, however, accent variation has been shown to have a much stronger impact on speech communication for normal-hearing listeners.

Speech intelligibility in noise is determined by the differences in the accents of the talkers and the listener and, while normal-hearing listeners are able to recognize a wide range of accents in quiet, in noise, they have more difficulties recognizing L1 and L2 accents that differ from theirs in noise (e.g., Evans and Iverson, 2007; Lane, 1963; Munro, 1998; van Wijngaarden, 2001). Even highly fluent, normal-hearing L2 listeners can behave as if they have a hearing impairment in noisy listening conditions, having elevated speech-in-noise thresholds (e.g., 15 dB) compared to L1 listeners (e.g., Rogers et al., 2006). Although the effect of accent on speech-in-noise recognition has been well established, exactly why and how this occurs is largely unknown. For example, it is not clear what types of mismatches are important (e.g., differences between the speaker and listener in vowels, consonants, rhythm, or prosody) and what kinds of compensations L1 and L2 listeners make when they 'tune into' an accent. Furthermore, it has yet to be established how L2 experience affects the talker-listener accent interaction.

Listeners' perception of speech in their L2 when spoken by a native speaker of that language is affected by their experience with the language; there is a correlation between the amount of L2 experience and speech recognition, particularly in noise, (e.g., Florentine et al., 1984). Several L2 experience-related factors have been found to affect the interaction between L2 experience and L1 speech recognition. For instance, age of L2 onset (i.e., age of L2 acquisition) has been shown to affect L2 listeners' performance on L1 speech in noise, with early bilingualism (i.e., L2 acquisition during childhood) leading to better performance on L1 English in noise compared to late bilingualism (i.e., L2 acquisition post-puberty; e.g., Meador et al., 2000; Flege et al., 1995; Flege et al., 1999). The age of L2 exposure has also been shown to be a more important factor in L1 speech recognition in noise than the length of L2 exposure. For instance, Mayo et al. (1997) tested L1 Mexican-Spanish listeners with different ages of L2 English acquisition and found that extensive exposure to the L2 did not result in native-like performance if the listener did not acquire the L2 in early childhood. Other factors associated with L2 experience such as the age of arrival in the L2 speaking country (e.g., Flege et al., 1995; Flege et al., 1999), length of residency (e.g., Mayo et al., 1997), L2 proficiency (e.g., van Wijngaarden et al., 2002), amount of continued L1 use (e.g., Meador et al., 2000) and lexical structure (e.g., Bradlow and Pisoni, 1999) have also been shown to affect L2 listeners' L1 speech perception in noise.

Previous work has also suggested that the degree of a listener's L2 experience may affect the interaction between the accents of the talker and listener. For example, van Wijngaarden et al. (2002, see also Hayes-Harb et al., 2008; Imai et al., 2005) tested L1 Dutch listeners who were highly proficient in English and less proficient in German on speech-in-noise recognition for L1 and L2 accents of Dutch, English, and German. The results demonstrated that listeners were more accurate at recognizing L1-English speech in noise than Dutch-accented English, but they were more accurate at recognizing Dutch-accented German than L1- German speech. This suggests that listeners may lose their advantage for L2-accented speech as they gain more experience with an L2, given that the subjects were more accurate for L2 accented speech only in their less proficient language (German).

Familiarity with accents has accounted for the differences in intelligibility for L1 listeners, with standard accents such as SSBE or RP consistently shown to be highly intelligible to listeners with very different accents (e.g., Glaswegian), suggesting that individuals may have an advantage in noise for multiple familiar accents (i.e., their own as well as accents widespread in the media; e.g., Adank et al., 2009; Clopper and Bradlow, 2008). Indeed, British English speakers may be better at recognizing accents that most often occur in the media compared to regional varieties to which they don't get a great deal of exposure. However, the role of the media in accent intelligibility is controversial and its contribution to making standard accents such as SSBE widely intelligible is not clear (e.g., Stuart-Smith, 2007). Familiarity with L2 accents may very much depend on socio-economic factors as well as the listener's personal circumstances. For instance, listeners may have close links with a linguistic community or live in a very linguistically diverse community (e.g., London) and therefore benefit from daily exposure to L2 accented speech. Other individuals, on the other hand, may have very little interaction with L2 speakers, notably in geographical areas where there is a low rate of immigration.

Taken together, the relative contribution of accent familiarity to speech intelligibility and the factors determining what constitutes accent familiarity are unclear and no concrete measure of familiarity has been established so far. There is also growing evidence that accent intelligibility could be driven by the similarity between the accents of the talker and listener, with listeners being more accurate at recognizing the speech of talkers that matches their own (e.g., Bent and Bradlow, 2003). However, there are no reliable methods to date that can quantify accent similarity reliably and help distinguishing its contribution to speech intelligibility from familiarity. Some methods have relied on listeners' perceptual categorization of talkers' accents but this measure is strongly affected by other talker-dependent variation, such as gender or voice quality (e.g., Clopper and Bradlow, 2008; Clopper et al., 2005a).

The aim of the present research was to provide a detailed investigation of the talker-listener accent interactions in noise in order to establish the contribution of the underlying factors involved in the L1-L2 speech communication to speech intelligibility for normal-hearing populations (e.g., L2 experience, accent similarity and familiarity). Study 1 examined the effects of talker and listener accent interactions in noise as well as L2 proficiency to speech intelligibility. Study 2 was focused on the production aspect of the accent interaction, investigating how the acoustic similarity in the accents of the talker and the listener accounts for the speech recognition patterns observed in Study 1.

Study 3 examined L1 listeners' adaptation processes to unfamiliar accents in noise with and without social interaction. Finally, Study 4 took a cross-linguistic approach and investigated how language experience and accent similarity affects the talker-listener accent interaction in noise.

2. Chapter two: Talker-listener accent interactions in speech-in-noise recognition.

2.1. Introduction

Even though the talker-listener accent interaction is well established, it is unclear why it occurs and what mechanisms are involved in it. One possibility, as discussed above, is that it could be due to familiarity or experience. For example, Adank et al. (2009) found an asymmetry in accent processing; Glaswegian-accented English speakers were equally fast at comprehending Glaswegian and Southern British English accents, but Southern British English speakers were slower with Glaswegian-accented speech. This may have been due to familiarity, because Southern British English is the dominant accent in the UK whereas Glaswegian is a more regional variety (i.e., both would be familiar to Glaswegian listeners, but Glaswegian would be less familiar to southerners). Likewise, several studies have shown that L1 listeners are able to readily adapt to L2 accents (Bradlow and Bent, 2007; Clarke and Garrett, 2004). This rapid adaptation is comparable to an effect of familiarity because the listeners get exposure to the accent during the course of the experiment. For instance, in Clarke and Garrett (2004), L1 listeners demonstrated an initial processing cost associated with exposure to L2 unfamiliar accents, followed by decreased reaction times over the first few trials as the listeners adapted (within one minute of exposure).

Likewise, longer-term L2 experience can also be interpreted as a familiarity effect; L2 speakers tend to become more accurate with L1 accents as they become more experienced (i.e., more familiar with L1 accents), and inexperienced L2 speakers tend to be more accurate with L2-accented speech. For instance, in Pinet and Iverson (2010), experienced L2 French listeners showed an advantage for L1 Southern British English accented speech over L2 French accented English in quiet and moderate levels of noise when performing an English sentence recognition task. Inexperienced L2 French listeners, however, displayed a strong advantage for the French accented English sentences at all noise levels (see also van Wijngaarden et al*.,* 2002).

Another possibility is that L2 talkers are more intelligible to L2 listeners because they share an interlanguage, which Selinker (1972:214) describes as a "separate linguistic system based on the observable output which results form a learner's attempted production of a target language norm", i.e., a combined phonetic/phonological knowledge base that develops when learning more than one language (see also Bent and Bradlow, 2003). That is, L2 accents are more intelligible to L2 listeners because there are more global listening strategies available to them that mutually increase speech intelligibility. They share the phonetic and phonological knowledge of both their L1 and L2, and thus are more equipped to interpret the acoustic-phonetic features in the speech of the L2 talkers, even though they may deviate from the target language. For example, individuals who speak both French and English may be mutually intelligible because the speakers and listeners both share a French and English phonological system. This interlanguage effect may also generalize across speakers of different L1s; Bent and Bradlow (2003) found that L2 Chinese listeners were more accurate at recognizing both Chinese- and Korean-accented English than L1 English. Therefore, speakers of different L1s could have common strategies when learning an L2, such that they are able to find other L2 speakers to be more intelligible.

The aim of this study was to investigate how L2 experience contributes to the talker-listener accent interaction in noise. L1 and L2 English speakers were tested on their speech-in-noise recognition of English sentences spoken with a range of L1 and L2 accents. The L2 listeners were French native speakers with a wide range of English experience (i.e., inexperienced with spoken English, experienced L2 English speakers living in London, and English-French from-birth bilinguals). The accents of the stimuli were chosen to match and mismatch the accents of the listeners so that they would span a range of familiarity to them: L1 Southern British English (highly familiar to more experienced English speakers), French-accented English from both inexperienced and experienced talkers (familiar to these L2 listeners), Northern-Irish English (L1, but relatively unfamiliar), and Korean-accented English (L2, and relatively unfamiliar).

2.2. Method

2.2.1. Subjects

There were a total of 93 subjects: one group of 21 monolingual native Southern British English listeners (‘SE’), and three groups of L1 French speakers with varying English experience (16 English-French bilinguals, ‘FB’, 24 experienced, ‘FE’, and 32 inexperienced, ‘FI’). The SE listeners were 18 to 48 years old at test (mean = 28 years).

The FB listeners had acquired both English and French from birth or at a very young age (age of acquisition of French: 0-18 months, mean = 2.6 months; age of acquisition of English: 0-9 years, mean = 11 months) and had a native-like command of both languages; their spoken fluency in both languages was assessed by the author (a native French speaker with high level of fluency in English). They were 18-36 years old at test (mean = 21 years).

The FE listeners were residing in London at the time of testing and therefore were very familiar with Southern British English accented speech. They had lived in an English speaking country for a period of time ranging from 1 month to 8 years (mean = 15.5 months), with 3 of the listeners having resided in Anglophone countries outside of the UK for a short period of time (e.g., Canada, Australia, United States). They had started learning English at school in France from the age of 6 to 14 years old (mean = 11 years) and were 18 to 48 years old at test (mean = 25 years).

Likewise, the FI listeners had started studying English at school from the age of 7 to 13 years old (mean = 11 years). Most of these subjects had spent little time in English speaking countries (i.e., no more than 8 weeks), apart from 2 subjects who had spent 6 and 12 months in English-speaking countries in the past but were not fluent in their L2. They were 18 to 54 years old at test (mean = 25 years). The FI listeners were from the same small community in north-eastern France as the FI talkers who recorded the stimuli; some of these individuals knew each other but they were unfamiliar with each other's English accent.

The FI group was tested in north-eastern France, and all other groups were tested in London. All listeners were given a language background questionnaire that included questions on their familiarity with the accents presented in the test.

2.2.2. Stimuli and apparatus

Four talkers each (two males and two females) of Standard Southern British English ('SE'), Northern Irish English ('IE'), Korean-accented English ('KO'), experienced French-accented English ('FE') and inexperienced French-accented English ('FI') were recorded reading the complete set of the 336 Bamford-Kowal-Bench (BKB) sentences (Bench et al., 1979; appendix 1). The FE and FI talkers matched their respective subject groups in terms of L2 experience and spoken proficiency. The listeners' experience with the talkers' accents presented in the experiment was assessed by the author to measure the effects of accent familiarity on the data. The recordings were made in quiet rooms with 44 100 16-bit samples per second.

Speech-shaped noise was created for each talker based on the smoothed long-term average spectrum of their recordings. The recordings were embedded in this noise with signal-to-noise (SNR) ratios of -9, -6, -3, 0 and +3 dB, and were also presented in quiet (i.e., no added noise). All stimuli were played to the subjects using a laptop over headphones at a user-controlled comfort level.

2.2.3. Procedure

The subjects performed a sentence recognition task in which they listened to BKB sentences and verbally repeated what they had heard, with the author logging the number of correctly identified keywords. Each sentence was presented only once (i.e., not repeated within or across conditions). Sentences for the different accents and SNR levels (including quiet) were presented in a random order within each block. The subjects were given a practice session of 16 stimuli at the start of the experiment to familiarise themselves with the test. The 16 sentences used in the practice were divided between accents and noise levels (including quiet). The practice sentences were not repeated in the main experiment. The practice block was followed by two blocks of 140 stimuli (i.e., 28 sentences for each of the five accent conditions, creating a total of 324 sentences). The sentences were counterbalanced to ensure that they were played in all accents and noise levels across subjects. In between the two main experimental blocks, the subjects were also given a block of 28 sentences of Korean-accented speech (i.e., only one accent). The mixed accent design was intended to avoid accent tuning effects. The short block of Korean sentences allowed for a limited evaluation of whether the mixed presentation reduced recognition performance and also evaluated adaptation processes in these listeners for use in future investigations (see Chapter 3).

2.3. Results

2.3.1. Main analysis

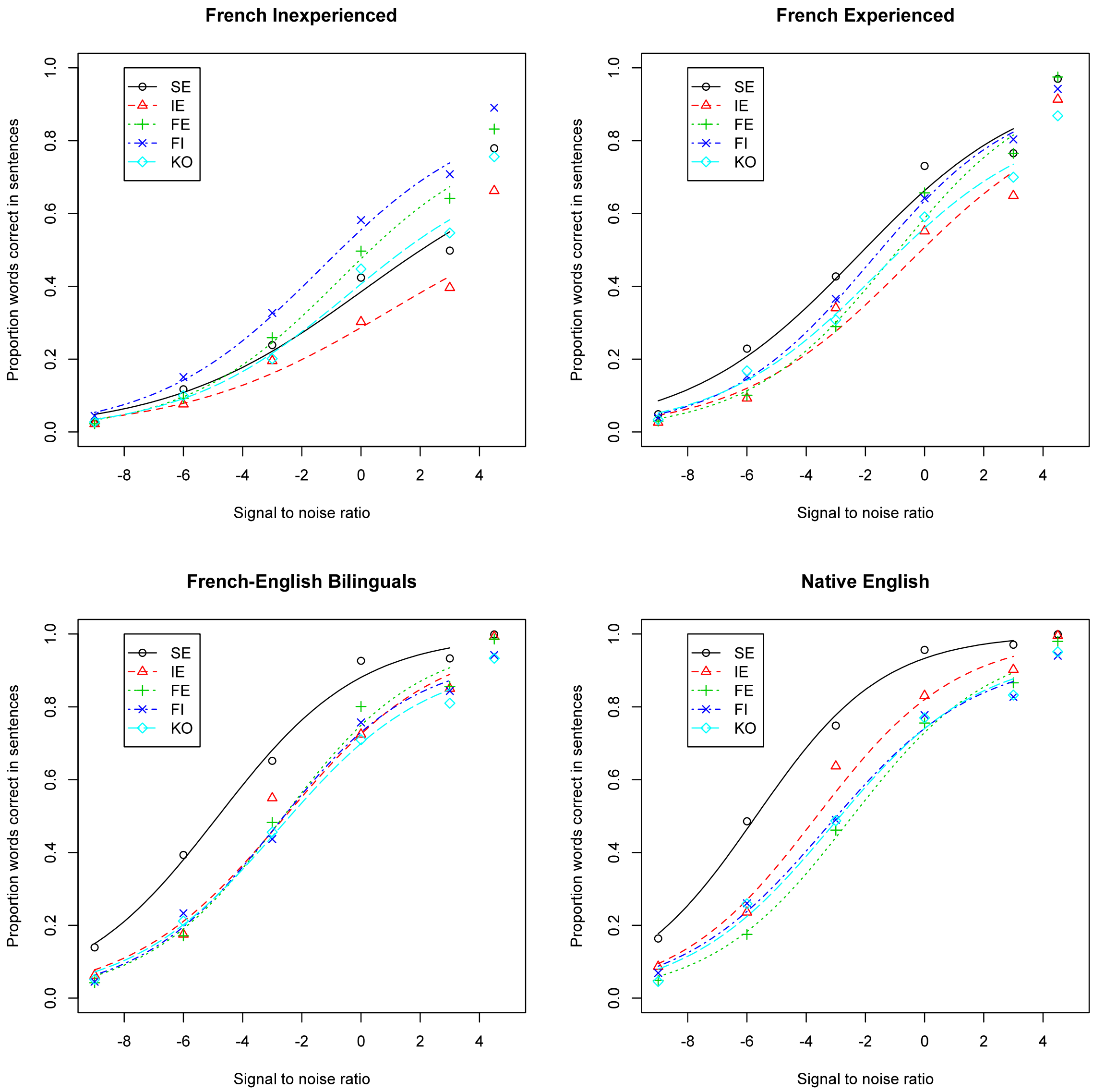


Figure 2.1. Psychometric functions of the proportion of correctly identified keywords as a function of SNR. The recognition scores in quiet were used to set the maximum of the psychometric functions. The results demonstrated a strong talker–listener accent interaction, with the least experienced L1 French speakers (FI) having the highest accuracy for French-

accented speech and the most experienced English speakers (SE and FB) having the highest accuracy for southern-English speech.

Figure 2.1 displays recognition accuracy (i.e., the proportion of words correctly identified in sentences) for the five accent conditions across all listening conditions. Experience with L1 English clearly affected how listeners recognized the various accents, with the SE listeners being most accurate overall, followed by FB, FE, and FI listeners being least accurate.

It also appeared that the intelligibility of the different accents varied with listener group (e.g., SE listeners being more accurate with SE speech, but FI listeners being most accurate with FI speech). In order to test these differences, a mixed-model analysis was conducted with accent condition and listener group as fixed factors and subject as a random factor. The percentage correct was averaged across noise levels to obtain an overall measure of how each listener performed in each condition[1], and the analyses were conducted on arcsine-transformed scores; the quiet condition was not included in this average, and was only used to set the ceiling of the psychometric functions. Figure 2.2 displays boxplots of the listeners' performance on all accents averaged across noise levels for all listener groups.

[1]The possibility of fitting psychometric functions to the individual data, then entering slopes and thresholds into the statistical models had been explored. However, the average percentage correct across noise levels proved to be a less variable and more robust measure of performance.

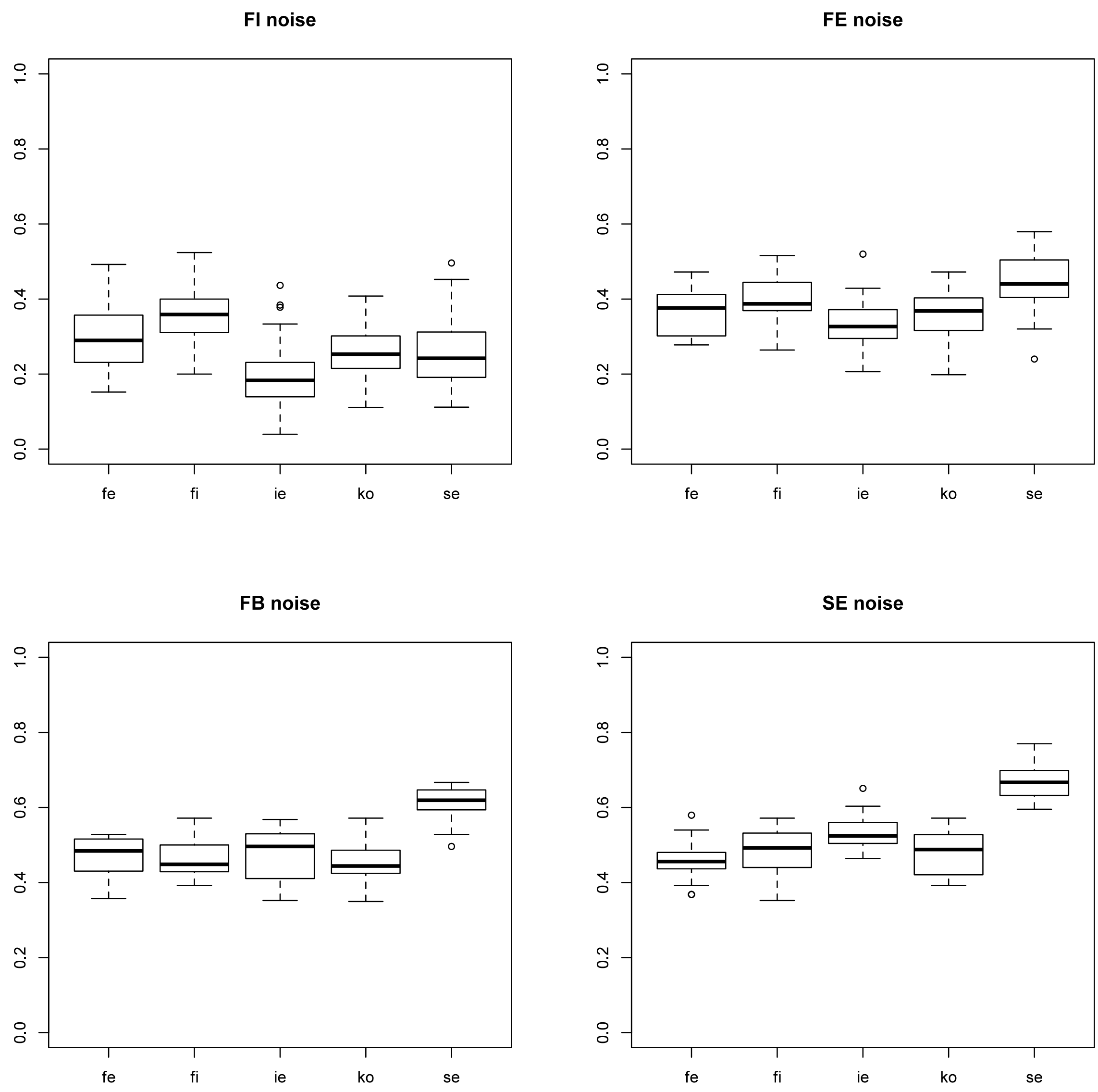


Figure 2.2: Boxplots showing the proportion of correctly identified keywords on each accent, averaged across noise levels (excluding quiet) for all listeners groups.

There were significant main effects of listener group, $F(3, 89) = 86.05$, $p < .01$, accent condition, $F(4, 356) = 89.44$, $p < .01$, and a significant interaction, $F(12, 356) = 41.53$, $p < .01$. In order to further investigate the interaction between the accents of the talkers and the listeners, mixed-model analyses were conducted separately for each group of listeners with accent as fixed factor and subject as a random factor. For FI listeners, there was a significant main effect of accent condition in noise, $F(4, 124) = 61.40$, $p < .01$. Tukey tests revealed that the intelligibility of almost all accents were significantly different, except SE and KO-accented speech ($p > .05$). As displayed in Figures 2.1 and 2.2, FI listeners were most accurate at recognizing sentences produced by FI speakers, then FE, similarly accurate for SE and KO, and least accurate for IE. These listeners were thus highly affected by accent, and had graded levels of recognition accuracy, possibly depending on the similarity of the accent to their own speech.

In contrast, the data in Figure 2.1 suggested that SE listeners were selectively tuned to their own accent, being most accurate at recognizing SE speech and having uniformly lower levels of accuracy for the other accents. The mixed-model analysis likewise demonstrated that there was a significant main effect of accent, $F(4, 80) = 89.78$, $p < .01$. Tukey tests confirmed that the listeners performed significantly better on SE speech than on the other accents, $p < .05$. However, IE was also significantly more intelligible than the L2 accents (FI, FE, KO), although the magnitude of this effect was small. The L2 accents were not significantly different from each other, $p > .05$.

The data from FE listeners in Figures 2.1 and 2.2 suggests that their accuracy was less affected by accent differences compared to the other groups. Nonetheless, the mixed-model analysis revealed a significant main effect of accent, $F(4, 92) = 26.57$, $p < .01$. Tukey tests demonstrated that the accents fell into an intelligibility ranking of SE, FI, FE, KO, IE; all accent pairs except FE-KO and KO-IE were significantly different. Therefore, even though the differences in accent intelligibility appeared reduced compared to the FI listeners, FE listeners were still affected by accent differences.

The FB listeners appeared to have a similar pattern of recognition to the SE listeners, being most accurate at recognizing sentences produced by SE-accented speakers but similar with the other accents. The mixed-model analysis demonstrated that there was a main effect of accent, $F(4, 92) = 43.58$, $p < .01$, and Tukey tests showed that only SE speech was significantly different from all the others. Therefore, the FB listeners were selectively tuned to SE speech.

2.3.2. Blocking analysis

In order to investigate whether the above results were affected by the mixed-accent blocking design, the data from KO-accented speech was compared from the mixed- and same-accent presentation blocks and averaged across noise levels. The boxplots for the two listening conditions are displayed in Figure 2.3. A mixed-model analysis was conducted with blocking (mixed or same accent presentation) and listener group as fixed factor and subject as a random factor. The condition with all Korean-accented speakers in the same block was significantly more intelligible than when these speakers were in a mixed block, $F(1, 89) = 4.69$, $p = .03$, although the magnitude of this effect was very small (i.e., 1.8 percentage points different, averaged across listener groups). The analysis also revealed a significant main effect of listener group, $F(3, 89) = 65.82$, $p < .01$, but there was no significant interaction with blocking. The effects of having a mixed-accent presentation design in the main experiment thus appeared to be fairly minimal, at least for this accent, although it is fair to acknowledge that experimental designs that specifically focus on accent learning may be able to find a stronger mixed vs. single accent difference.

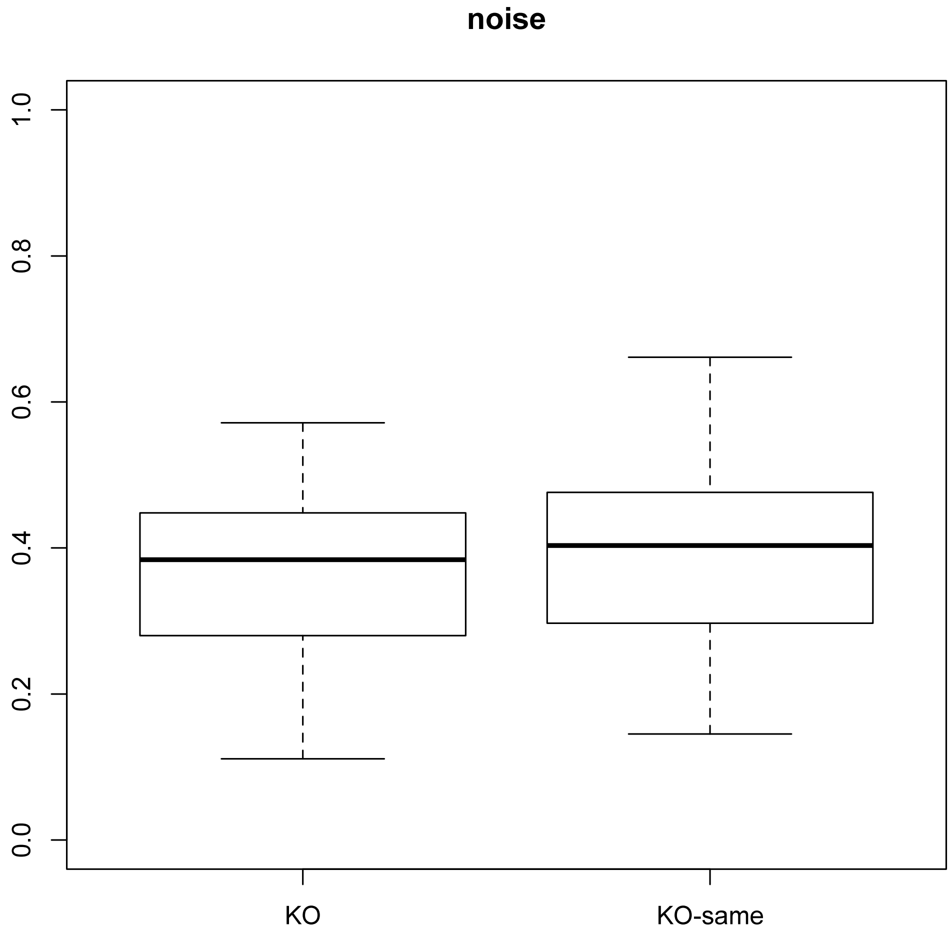


Figure 2.3. Boxplots displaying the proportion of correctly identified keywords for Korean-accented speech in mixed ('KO') vs. single accent ('KO same') listening condition, averaged across noise levels.

To examine learning effects within the mixed blocks, the overall percentage correct (i.e., across all accents) was compared for the first and second mixed blocks (i.e., the mixed-accent trials were split into two blocks, with the single-accent Korean block in the middle). A mixed-model analysis was conducted with block (first or second) and listener group as fixed factor and subject as a random factor. The difference between the first and second mixed block was significant, $F(1, 89) = 17.99$, $p < 0.01$, although the amount of the learning effect was small (i.e., 2.2 percentage points improvement for the second block, averaged across listener groups). There was also a significant main effect of listener group, $F(3, 89) = 86.58$, $p < .01$, but there was no significant interaction of block. There was thus improvement in speech recognition over the course of the entire experiment, but the magnitude of learning was not substantial compared to our effects of interest (i.e., between-accent differences).

2.4. Discussion

The results demonstrated a clear talker-listener interaction. For example the SE listeners were more accurate at recognizing speech produced by SE talkers but had no difference for L2 accents and were only marginally better on IE speech. They were thus selectively tuned to their own accent. The FI listeners, on the other hand, performed better on FI-accented speech, then FE speech, and became progressively worse on the other accents, thus showing graded sensitivity. The more experienced L2 listeners (FE and FB) were better at SE speech in noise, becoming selectively tuned to it as their experience with L1 speech increased.

It is arguable that the results are in support of the interlanguage benefit hypothesis (e.g., Bent and Bradlow, 2003) because the FI listeners displayed a clear advantage for FI speech. However, the FE and FB listeners did not show such a clear advantage for French-accented speech, despite the fact that these listeners also had learned both French and English phonological systems. It is thus unlikely that having a shared phonological background alone determines which accents are most intelligible. Likewise, there was no evidence that L2 listeners had broader advantages for L2 speech from listeners with other language backgrounds, because KO speech was difficult to understand; for instance the FI listeners showed no advantage for KO speech over SE speech but instead had higher recognition levels for French accents. This parallels Stibbard and Lee's (2006) study in which they demonstrated that speakers of Korean-accented English had no advantage for Saudi-Arabic-English accents. To be fair, Bent and Bradlow (2003) speculate that part of the interlanguage benefit effect could be driven by accent similarity. It is possible then, that while the similarity between the accents of the Chinese and Korean talkers and listeners could have enhanced the interaction, revealing an interlanguage effect, the lack of similarity in the speech of the L2 speakers in Stibbard and Lee (2006) and the present study produced the opposite effect. The contribution of accent similarity in the talker-listener interaction will be investigated further in the next chapter.

Language experience had a strong impact on the listeners' accent tuning processes in noise and likely modulated their recognition of their different accents in noise. At its most basic level, the results are in accord with previous findings that intelligibility is enhanced when talkers and listeners share the same L1, particularly in adverse listening conditions (e.g., van Wijngaarden, 2001; van Wijngaarden et al., 2002; Bent and Bradlow, 2003; Pinet and Iverson, 2010; Stibbard and Lee, 2006), but the present study additionally demonstrated that intelligibility was further enhanced for listeners when there was a match in accent with the talkers. The SE listeners displayed a strong selective tuning for their own accent in noise while the FI listeners showed an advantage for French-accented English speech in both quiet and noise. However, the accent processing pattern in the more experienced L2 listeners was more complex. Indeed, the results mirrored the findings of van Wijngaarden et al. (2002) in that only the least experienced L2 listeners benefited from the allophonic productions of the low proficiency talkers (i.e., FI talkers and listeners). This benefit was not shown for the more experienced L2 listeners. The FB listeners processed English in a similar manner to the monolingual listeners, selectively tuning their recognition processes to SE speech in noise. The FE listeners, on the other hand, had a less pronounced advantage for SE speech, with reduced, but significant, differences in accent intelligibility compared to the FI listeners. This processing pattern is likely due to their proficiency in English, which lies between that of the FI and FB listeners. Indeed, their reduced exposure to French accents and increased familiarity with SE could explain their advantage for SE speech in noise. Likewise, the FB listeners' experience with SE, which is substantially more considerable than the FE listeners', also likely modulated their accent tuning processes in English. Still, if L2 experience is comparable to accent familiarity, we may speculate on the extent to which familiarity accounts for the present findings. Indeed, while familiarity accounts well for the L2 listeners' performance on SE speech, it doesn't account for the fact that the experienced French listeners (FE and FB) couldn't use their French phonological knowledge and familiarity with French accents to perform better on French-accented speech, if not equally as well as they did on SE speech. The contribution of accent familiarity will be further examined in relation to accent similarity in the next chapter.

The listening conditions also clearly affected the listeners' accent recognition patterns. The results showed a shift in accent processing between the quiet and noisy listening conditions, revealing the listeners' more automatic speech processing techniques. Indeed, the listeners were able to recognise a wide range of accents in quiet, but when the signal was degraded, their focus shifted to acoustic cues that would likely enhance their speech recognition processes, with the effect reducing at high SNRs. For instance, in the quiet listening condition, the SE and FB listeners performed equally well on both L1 accents (SE and IE accents) and FE speech (the most proficient L2 talkers), despite them reporting having no particular familiarity with IE or FE accents. However, in the presence of noise, the listeners reverted to a selective type of accent tuning, recognising SE speech more accurately than the other accents. Interestingly, IE speech was processed on the same level as L2 accents for the FB listeners, and was only marginally more intelligible than the L2 accents for SE listeners group.

It is particularly surprising that the SE listeners showed such poor recognition levels for IE speech in noise, considering that it is a L1 accent. It is possible, then, that the noise masked the L1 cues that the listeners could easily take advantage of in quiet and that several of the acoustic features constituting IE speech that deviate from SE speech could have become more prominent in noise (such as prosodic and spectral cues), thus making speech recognition increasingly challenging in noise.

The FE listeners found SE speech and their own accent to be the most intelligible in quiet, followed by FI speech, but the intelligibility for French accents shifted in noise, with FI speech being more intelligible than FE speech. It is very likely that the L1 acoustic-phonetic cues in the allophonic realizations of the FI talkers enabled the listeners to overcome the degraded listening conditions, as they are more prominent in noise than the FE talkers' due to the talkers' degree of accentedness. This shows a flexible approach to the use of acoustic cues to recognise accented speech compared to the other listener groups, which has already been reported in Pinet and Iverson (2010). Indeed, listeners have been previously shown to shift to using prosodic cues more heavily for word segmentation under acoustically degraded conditions when segmental cues become unavailable (see also Mattys et al., 2005). For instance, in Pinet and Iverson (2010), FE listeners took advantage of the presence of their L1 and L2 segmental and prosodic cues in the signal to recognise increasingly degraded speech, modulating their reliance on the cues according to noise level. However, noise can also be thought of as a stressor to the speech recognition system, and French experienced listeners may simply revert to a more French-like way of perceiving the stimuli when the listening conditions become difficult. It thus seems conceivable that experienced L2 listeners may be able to recruit either or both of their L1 and L2 experiences to fit the demands of the listening situation, which may offer options to speech processing that monolingual listeners do not have. Likewise, the FI listeners consistently made use of the presence of their L1 cues in the speech of the French accented talkers in all listening contexts. Still, the FB listeners' inability to take advantage of the French L1 cues in noise despite having a French phonological system is puzzling. It is possible however, that they revert to a more monolingual-like way of processing accents in noise when listening to English, and the same pattern may be true when they process French. The bilinguals' speech accent processing patterns in both French and English will be further investigated in Chapter 5.

To some extent, some of the conclusions could be affected by the design used in the present experiment, in which multiple accents were presented within the same block. This design was used in order to avoid any accent tuning effects, but it is possible that the more experienced listeners in this kind of experiment used processes or strategies that favoured the more standard English accent. Previous research has shown that listeners can adapt to a novel or L2 accent (e.g., Bradlow and Bent, 2007; Clarke and Garrett, 2004), and it could be the case that listeners might have been able to adjust to the different accents better within single-accent blocks. For example, FB listeners may have been better able to recruit their French phonological system to help understand FI speech if there had not been competing accents. This possibility was tested with KO speech and there was no strong evidence for adaptation, but this was only a limited test and the chosen accent was an unfamiliar L2 accent to all listeners (i.e., the listeners couldn't recruit their L1 or L2 phonological knowledge to process this particular accent). Chapter 4 examines the role of blocking to investigate whether adaptation can modulate the talker-listener interaction, as well as observing any effects of flexibility of processing in experienced L2 listeners and bilinguals.

3. Chapter three: Acoustic similarity contribution to the talker-listener accent interaction

3.1. Introduction

In the previous chapter, it was suggested that the talker-listener accent interactions could be driven by the phonetic similarities between the accents of the talker and the listeners. For example, Bent and Bradlow (2003) speculated that L2 Chinese listeners could have found Korean-accented speech to be more intelligible because they share similarities in their L1 phonological systems, and Stibbard and Lee (2006) found no intelligibility benefit for L2 speakers with very different L1 phonological systems (Korean- and Saudi Arabic-accented English). Likewise, some of the familiarity effect due to L2 experience (e.g., Pinet and Iverson, 2010; van Wijngaarden et al., 2002) could be driven by accent similarity. That is, as L2 listeners become more proficient, their own productions become closer to L1 speech, and this could prompt a change in accent intelligibility, with L1 speech becoming more intelligible than L2 accented speech. For instance, in van Wijngaarden et al. (2002), the L1 Dutch listeners who were highly proficient with English and less proficient with German showed an advantage for L1 English speech in noise over L2 Dutch-accented English, but they were more accurate at recognizing Dutch-accented German than L1 German speech. Thus, the basic phonetic similarities of the talkers' and listeners' accents could affect intelligibility, irrespective of the cause of this similarity (e.g., familiarity or interlanguage).

Likewise, several studies have shown that familiarity with a L1 accent accounts for a benefit in intelligibility for L1 listeners, in particular when there is a match in accent (e.g., Adank et al., 2009; Floccia et al., 2009), but this advantage could also be enhanced by the presence of acoustic similarities in the accents of the talkers and the listeners. For instance, the Southern British English (SE) listeners in Adank et al. (2009) performed better on Glaswegian-accented (GE) speech in the second than in the first study. This was likely due to a reduction in accent-related variation as only two of the four speakers present in the first study were selected, making the phonetic-phonological or acoustic differences between the accents of the talkers and the listeners less salient. Hence, even though the listeners were unfamiliar with the accent of the talker, this reduction in acoustic differences contributed to enhancing the intelligibility in the interaction.

However, the very notion of accent familiarity and its contribution to intelligibility remains unclear. For instance, in Adank et al.'s study, the SE listeners were also tested on their recognition of Spanish-accented English (SpE) and performed only marginally better on GE speech compared to SpE. Interestingly, the listeners reported having no familiarity with the accents presented in the experiment, but it is very likely that they had more familiarity with and exposure to SpE than GE speech, as they were tested in Southern England (London) where the Spanish accented community is more widespread than the Glaswegian one. However, this advantage for GE speech is likely due to acoustic similarity between the SE and GE accents. Likewise, speakers of non-standard L1 accents have been shown to perform equally well on SE speech and their own accent, which is comparable to a familiarity effect as listeners get regular, almost daily exposure to SE speech, which is the most widespread L1 accent in the UK, particularly in the media (e.g., GE speech in Adank et al., 2009, and Evans and Taylor, 2010). In addition, the between- and within-talker phonological variation is much less significant in non-standard L1 accents than in L2 accents and therefore the processing cost associated with unfamiliar L1 accents is small to negligible in quiet listening conditions.

The aim of the present study was to investigate whether talker-listener accent similarity can account for L1-L2 accent intelligibility in noise and examine the relative contribution of accent familiarity to this interaction. One difficulty is that it is not clear how to assess and quantify accent similarity. Most studies using accent assessment have focused on evaluating the degree of L2 accent using perceptual (e.g., Flege et al., 1999) or computational methods (Cincarek et al., 2009; Cucchiarini et al., 2000; Franco et al., 2000; Neumeyer et al., 2000), but this is a different issue from comparing the similarity of arbitrary pairs of accents. Clopper and colleagues (e.g., Clopper et al., 2005b; Clopper and Bradlow, 2007, 2008; Clopper, 2008) have developed a free-classification task to assess accent similarity, in which listeners sort accents into groups based on perceptual similarity. However, such a perceptual task would be difficult to apply in the present study given that the aim is to compare a large number of speech samples with each other (113 talkers and listeners; speech samples taken from the participants in Chapter 2). In addition, this technique relies exclusively on listeners' perceptual judgement of similarity and can be affected by the same processes involved in speech recognition rather than being an independent measure (i.e., the ratings could be affected by accent-independent factors).

Instead, a computational method was applied to the data (ACCDIST; Huckvale, 2004, 2007a, 2007b), a more reliable and objective accent measurement method, in order to measure the acoustic similarity between the accents produced by pairs of speakers. Using ACCDIST, acoustic measurements are made automatically on phonetically-transcribed recordings (e.g., vowel spectra, duration, and pitch), which are then compared to each other to create a table of phonetic similarities for each speaker (e.g., measuring the distance between how a speaker produces every pair of vowels in the corpus). The assessment of relative phonetic similarity within each talker reduces the influence of global speaker characteristics (i.e., factors that are not specific to individual segments, such as differences due to vocal tract size, F0 range, or speech rate), leaving the phonetic differences that are more indicative of accent. These matrices of within-speaker segmental acoustic distances are then compared between pairs of talkers (correlation coefficients). Thus far, the measure has only been applied to the classification of British English accents (Huckvale, 2004). Although this approach could also be used for consonants (future investigation work), vowels are easier to compare spectrally and these vowel measures have reliably correlated with accent differences. The present study extends this measure to a wider range of accents and examines whether it can account for intelligibility data.

Following the sentence recognition task, all the listeners who took part in the experiment described in Chapter 2 were also recorded reading a subset of the test materials so that their spoken accent could be acoustically compared to the accents of the stimuli, using ACCDIST (Huckvale, 2004, 2007a, b). The aim was to examine how well this accent similarity measure could account for the relative intelligibility of the different accents for each listener, and whether there were effects of familiarity or interlanguage that could not be explained by accent similarity.

3.2. Method

3.2.1. Subjects

The subjects are described in the method section in Chapter 2. They produced speech samples after completing the speech recognition task in the first study. They were recorded reading 31 of the BKB stimuli sentences presented in the experiment (highlighted in appendix 1). The recordings were conducted in a sound proof booth for the subjects tested in London and in a quiet room for the subjects tested in France. ACCDIST (Huckvale, 2004, 2007a, b) was used to compare the accents of the subjects to the accents of the same sentences used for the stimuli.

3.2.2. Accent analysis

To identify the regions of the acoustic signal associated with phonological vowels, a process of automatic alignment of phonetic labels was performed, followed by manual checking and correction by the author. The automatic alignment was performed separately for each sentence against a single phonological transcription. Alignment was performed using two sets of hidden Markov models (HMMs). The first set of HMMs was used to establish a basic phonetic alignment, and then those alignments were used to initialise the second set of models. The models were trained with an embedded training procedure using the HTK toolkit (Hidden Markov Modelling Toolkit, 1989), which was also used to generate the forced alignment of the phonological transcription to each sentence. The first set of HMMs was trained on a standard British English database (WSJ-CAM0). The second set of HMMs was trained on all recordings from all speakers used in this experiment. One context-independent HMM was trained for each of 44 phones plus silence. Each HMM consisted of three states and used mel-frequency cepstral coefficients (MFCCs, Davis and Mermelstein, 1980) as observations. An MFCC vector was computed every 10ms and consisted of the first 12 cepstral coefficients plus one coefficient of overall energy. The alignments resulting from the second set of HMMs were then checked manually and sentences that resulted in very poor alignments were discarded. Discarded sentences were mainly those containing dysfluencies in production.

The ACCDIST distance between two speakers was performed on measurements from every vowel segment (except the unstressed vowel /ə/) in all the aligned BKB sentences that were common to both speakers. The region of the signal identified as a vowel by the aligned transcription was divided into two equal halves, and the mean MFCC vector was computed for each half. The MFCC vectors for vowels were not averaged across different instances of the same phonological vowel in the sentences, except in the case where the same word was repeated. That is, we treated the vowels in *clown*, *down*, *ground* as distinct, while we averaged the vowels found in two instances of the word *boy*.

The MFCC vectors were then used to calculate an intra-subject vowel distance table, which assessed the vowel spectral contrasts that an individual made when speaking the words in the sentences. Calculating vowel distances within speaker is effectively a normalization procedure, focusing on only the spectral distances that contrast particular vowels rather than on more global spectral differences between the recordings (e.g., associated with voice quality or vocal tract length). Each element in the distance table for a speaker contained the unweighted Euclidean distance between the MFCCs vectors for two vowel instances. Thus for speaker *S1*, and vowel list *V*, the distance table D^{S1} was computed from each vowel's concatenated MFCC vectors *f*, as in Equation 1.

$$D_{i,j}^{S1} = \left[\sum_{k=1}^{26}(f_{ik}^{S1} - f_{jk}^{S1})^2\right]^{\frac{1}{2}}, i \in V, j \in V, i \neq j$$

Equation 1.

Finally, the correlation between distance tables was calculated across speakers, ensuring that each vowel distance that was correlated corresponded to the same pair of vowel instances in the recorded sentences. For each stimulus accent (e.g., SE, FI), a listener's accent was compared to each of the four speakers who recorded stimuli for that accent, and then the average was taken as the listener's similarity to this accent.

A distance based on vowel duration was calculated in a similar fashion. The vowel durations were calculated based on the alignments described above (e.g., averaging by word). However, the durations were correlated directly between pairs of speakers without first calculating intra-speaker matrices. This was because duration is a one-dimensional measure (as opposed to the 26-dimensional MFCC vectors used above), and for a one-dimensional measure the correlation statistic already normalizes for rate etc. by eliminating differences in means and standard deviations.

3.3. Results: Comparisons to ACCDIST

The ACCDIST measures based on MFCC spectra and vowel duration are displayed in relationship to the intelligibility data in Figures 3.1 and 3.2.

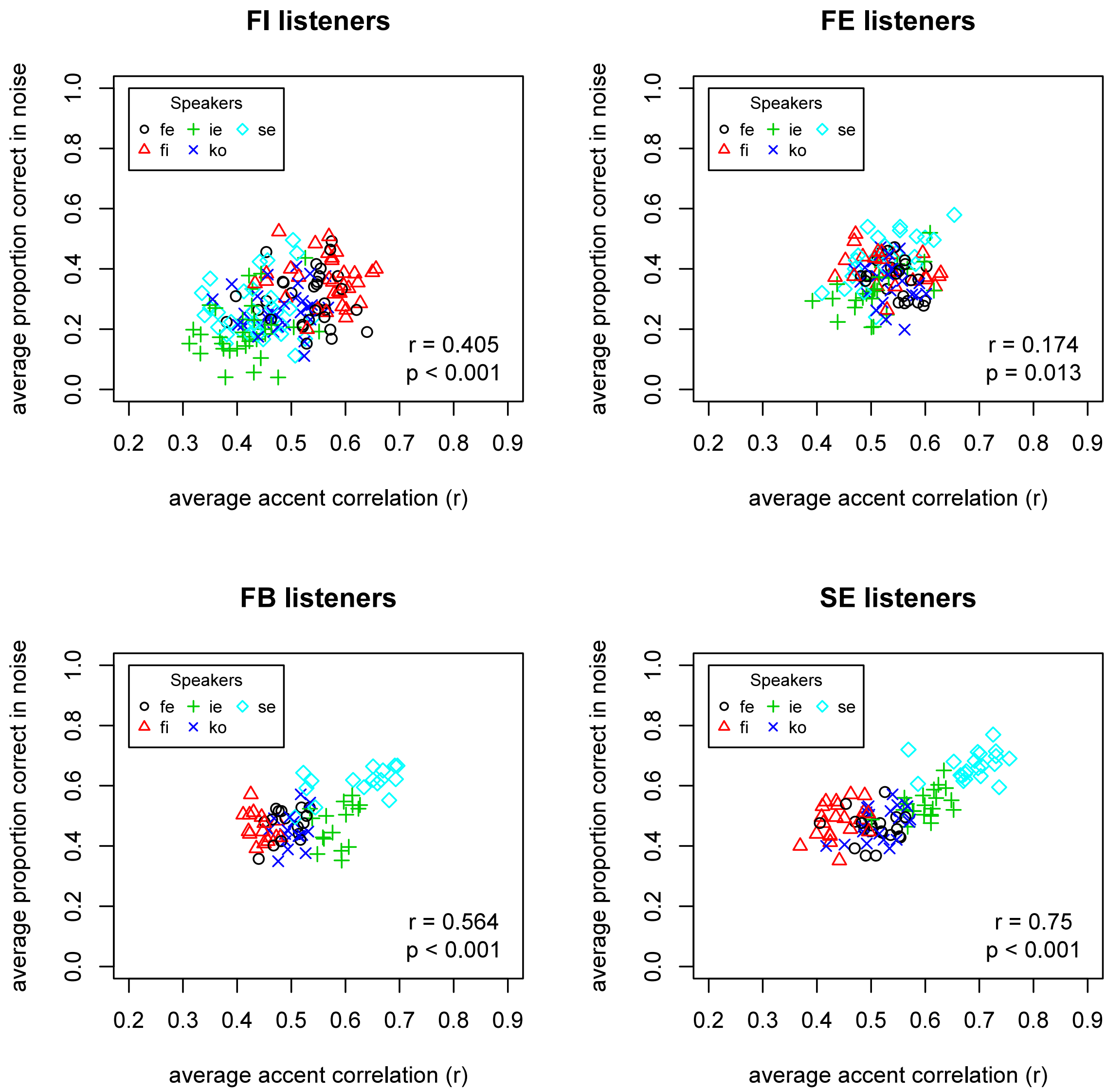


Figure 3.1. Scatterplots of accent correlations based on vowel spectra vs identification accuracy for each listener group. The *r* value represents the correlation between the two variables. The *p* value is taken from the mixed-factor analysis of the same relationship, with subject added as a random factor.

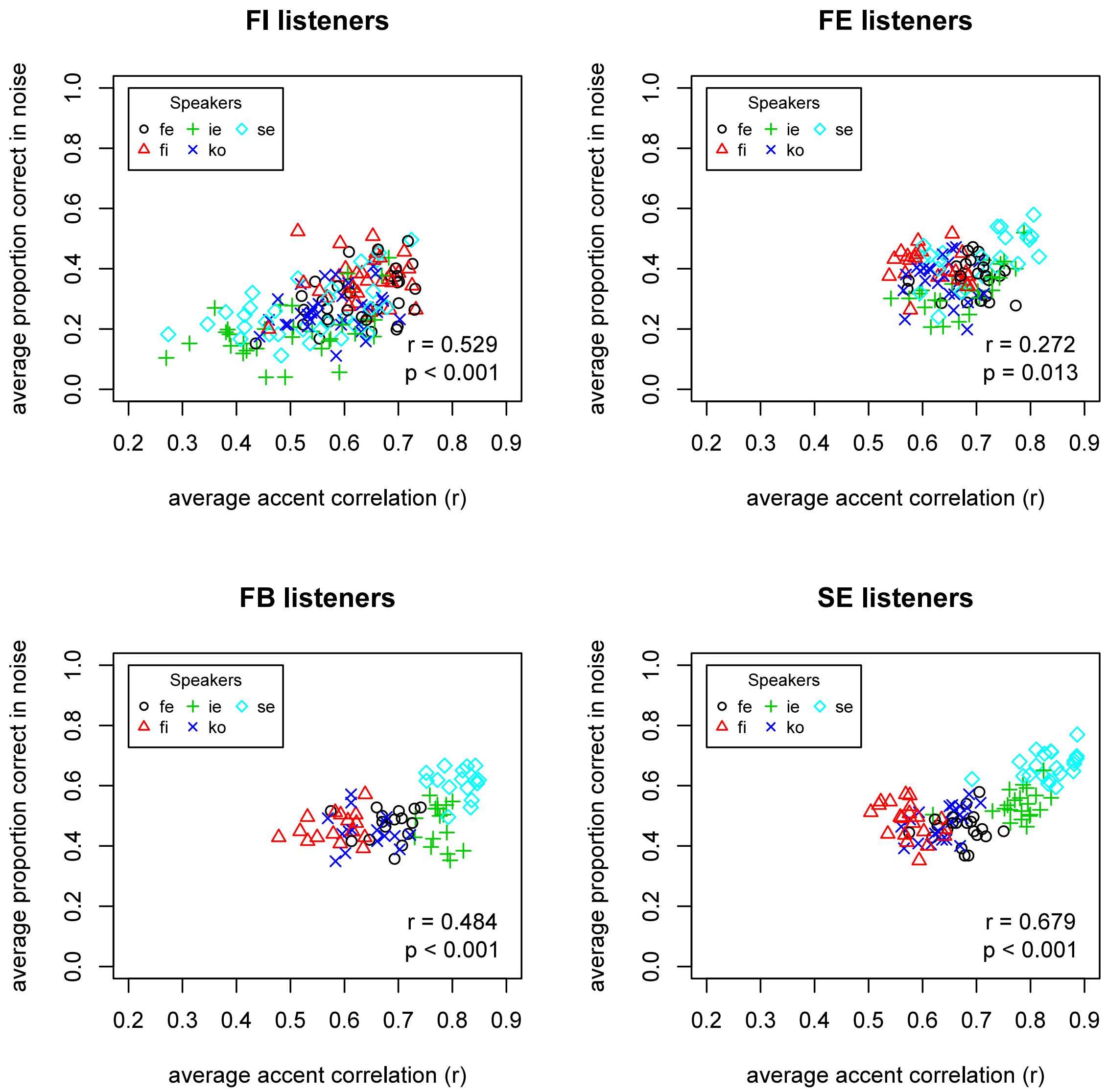


Figure 3.2. Scatterplots of accent correlations based on vowel duration vs identification accuracy for each listener group. The *r* value represents the correlation between the two variables. The *p* value is taken from the mixed-factor analysis of the same relationship, with subject added as a random factor.

For each listener, it was plotted how similar his or her own accent was to each of the 5 accents (i.e., averaged across talkers); this is represented by 5 separate points per listener on the scatterplots (i.e. 1 point for each of the 5 accents). Strong relationships were apparent between measured accent similarity and mutual intelligibility for pairs of listener types, with relatively little difference in performance between the spectral- and duration-based accent similarity metrics. A mixed-effects analysis was conducted with average percentage correct in noise as the dependent variable, ACCDIST for spectra and duration as linear independent variables, listener type as a categorical variable, and subject as a random factor. The analysis revealed significant main effects of vowel spectra, $F(1, 370) = 383.90$, $p < .01$, and duration, $F(1, 370) = 52.82$, $p < .01$. The similarity of the accents of the listeners and talkers on these measures thus were both able to account for differences in intelligibility, even when entered into the same model. For example, the scatterplots and correlations for FI listeners (Figures 3.1 and 3.2) demonstrate that their accents were closest to those of other French-accented talkers (FI and FE), equally far from SE and KO despite the fact that these are very different accents, and was the furthest away from IE accents; this mirrors the intelligibility data. The measured accents of FB and SE listeners, on the other hand, were closest to that of the SE stimuli, again mirroring the relative intelligibility of the different accents.

FE listeners had a weaker relationship between ACCDIST and intelligibility compared to the other groups of listeners, which could have been due to their smaller ranges of scores. That is, their spoken accent was not as distinctive, being relatively similar to a range of accents rather than particularly close to a single accent. They likewise had a narrower range of intelligibility scores. This could have occurred because they had an intermediate level of spoken proficiency, and thus had an accent that was neither highly distinctively French nor highly native like.

There was also a main effect of listener type, $F(3, 89) = 54.37$, $p < .01$. This suggests that, in addition to accent similarity, there were overall effects of proficiency. That is, irrespective of accent, SE listeners were more accurate at English speech recognition than were FI or FE listeners. This proficiency difference can also be observed in the variance of data in the scatterplots. For example, the best fit between the accent measures and intelligibility is found for the SE listeners (e.g., $r = 0.750$, $p < .01$, for vowel spectra), and presumably this group was fairly uniform in their abilities to understand English, whereas the fits are weaker for the FI and FE groups, but the variance could be higher because these individuals differed more in their English abilities.

To some extent, these effects of proficiency and talker-listener accent similarity can work in opposition. For example, if a listener has an accent that is closer to FI speakers this will make them more accurate with FI speech, but having such a strong French accent is also an indicator that the listener is less proficient with English, and will thus likely have more difficulty with English speech in noise. Figure 3.3 displays data across listener groups for the SE accent only, because in this condition talker-listener accent and English proficiency work in the same direction (i.e., more proficient speakers have an accent that is more similar to SE). In this circumstance, the correlation between accent distance and intelligibility becomes high for both vowel spectra measurements, $r = 0.853$, $p < .01$, and vowel duration measurements, $r = 0.868$, $p < .01$. However, this relationship becomes reversed and weaker for FI accents (for vowel spectra: $r = -0.537$, $p < .01$; vowel duration: $r = -0.301$, $p < .01$) because the proficiency effect is stronger than accent similarity on its own.

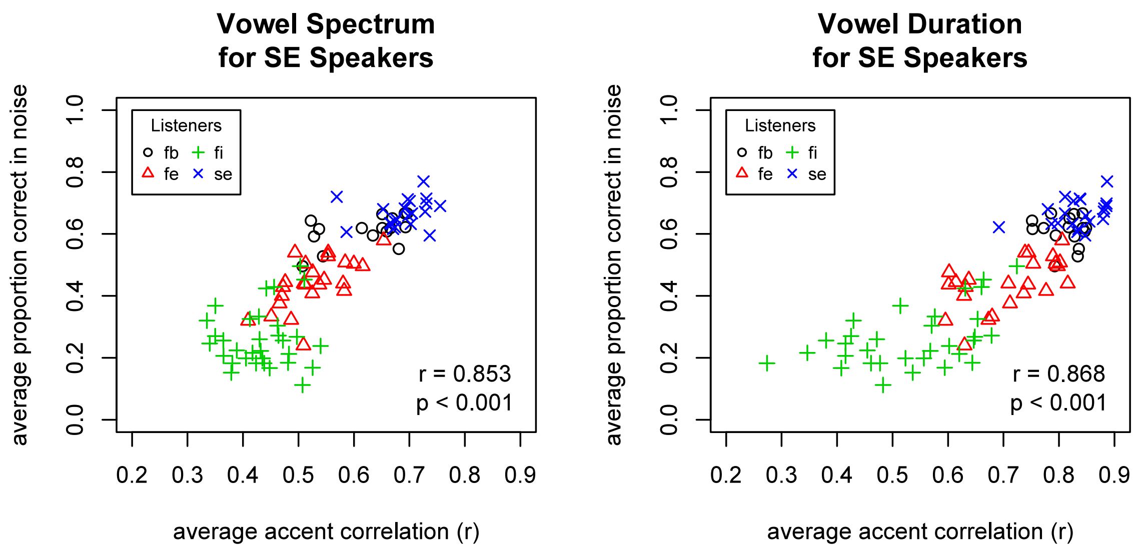


Figure 3.3.Scatterplots for southern-British English stimuli only (SE) of accent similarity (both spectra and duration) vs identification accuracy in noise, for each listener group.

There is thus strong evidence that the ACCDIST measures used here are able to effectively assess talker-listener accent similarity, and that this similarity can explain many of the differences in the ability of listeners to understand speech in noise.

To help illustrate the accent differences between the talker and listener groups, the ACCDIST measure was recalculated by averaging across vowel (e.g., averaging MFCC values for words that would normally be produced the same, such as *clown*, *down*, *ground*), and then using multidimensional scaling (Kruskal, 1964) to plot these vowels in two-dimensional spaces. As displayed in Figure 3.4, all accent groups had an English-like vowel space, with significant deviations.

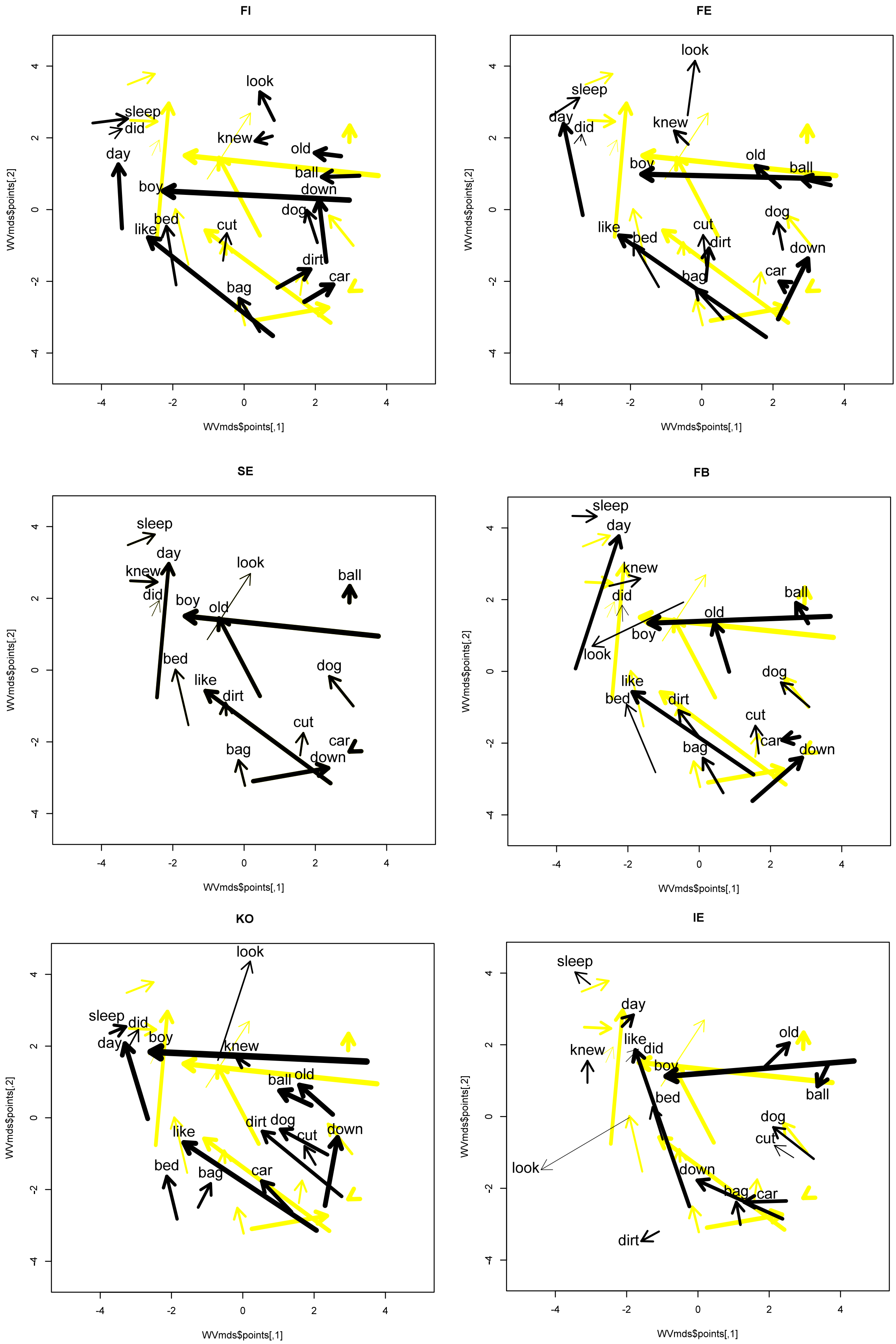

FI
look
sleep
did
knew
old
day
ball
boy
down
dog
bed
like
cut
dirt
car
bag
WVmds$points[,2]
WVmds$points[,1]
FE
look
sleep
day
did
knew
boy
old
ball
dog
like
bed
cut
dirt
down
car
bag
SE
sleep
day
knew
did
look
ball
boy
old
bed
dog
like
dirt
cut
car
bag
down
FB
sleep
day
knew
did
ball
old
boy
look
dog
like
bed
dirt
cut
car
bag
down
KO
look
sleep
did
day
boy
knew
ball
old
dirt
dog
cut
down
like
bed
bag
car
IE
sleep
day
like
did
old
knew
boy
ball
bed
dog
cut
look
down
bag
car
dirt
4
2
0
-2
-4

Figure 3.4. MDS plots of vowel spaces for all groups of talkers and listeners. Each arrow shows formant movement (i.e., starting from the MDS coordinates for the MFCC spectrum calculated over the first half of the vowel and ending at the coordinates calculated over the second half of the vowel), with the line weight of the vowel indicating duration (thicker lines for longer vowels), and an example word for each vowel.

For example, KO speakers assimilated the vowels in the word pairs *old-ball, dog-dirt, did-sleep,* and IE speakers tended to have strong fronting for vowels such as in *knew* and *look*. The FI accent may have been equidistant between KO and SE speech because it shared some of the assimilations with KO speech, but had other aspects that were more like SE (e.g., difference between *bed* and *bag*). Likewise, the fronting of IE may have contributed to this accent being highly dissimilar from that of FI speakers. As listeners gained more experience with English (FE and FB), their vowel spaces became closer to SE speech.

3.4. Discussion

The most important finding that emerged from the acoustic analysis is that much of the variance in intelligibility could be accounted for in terms of the acoustic similarity of the accents of the talker and the listener, both in terms of duration and spectral distance, either independently or in the same statistical model. That is, listeners were more accurate at recognizing the speech of talkers whose accents closely matched their own acoustically, and the accuracy decreased with increasing accent distances. If acoustic similarity can account for a great deal of the talker-listener interaction, it is questionable to what extent familiarity or interlanguage still plays a role.

To some extent, one could argue that accent familiarity can explain the present data because the listeners who were more experienced with SE speech were the most accurate at understanding this accent in noise. However, accent familiarity cannot account for how the listeners performed on the other accents. For instance, all of the French-speaking subjects were highly familiar with French-accented English, but only the least experienced listeners had an advantage for French-accented speech. In particular, the FB listeners all reported having some familiarity with French-accented speech (assessed by a questionnaire), having been raised in a mixed French and English speaking environment, having a French-accented parent or family member, and having French-accented peers in their community (i.e., FE speakers living in the UK or FI speakers living in France). However, they had no intelligibility advantage for French-accented speech, recognizing it similarly to KO and IE accents, and with about the same accuracy as did L1 English speakers. Likewise, accent familiarity cannot account for the fact that FI listeners performed equally well on SE and KO-accented speech; they didn't report having any exposure to KO speech at all, but had some exposure to SE speech through the media and short travels to the UK. In contrast, accent similarity can account both for why FB listeners had no advantage for FI speech (i.e., their own English accent was far away from that of FI speakers), and why FI listeners found KO and SE accents to be equally intelligible (i.e., both were equidistant from the FI accent).

Acoustic similarity can also account for the more complex FE listeners' reduced accent sensitivity and lack of selective tuning found in the other listener groups. Their intermediate proficiency (when compared to that of the FB and FI speakers) implies that acoustically, their productions are neither native-like in the same manner as the FB listeners, nor close to the allophonic realizations of the FI listeners. Instead, they display a more 'adaptable' phonetic space with more global acoustic-phonetic features matching some features of the talkers' accents, with, for instance, their durational patterns being closer to that of SE-accented speech and some of their spectral features matching that of FI and FE accented speech. This explains the advantage, despite its small size, for SE speech, followed by FI and FE speech, i.e., the listeners have acquired near-native durational patterns while retaining some French spectral features in their accent and were therefore able to take advantage of these acoustic features in noise.

The interlanguage benefit effect observed in Bent and Bradlow, (2003) and absence of it in Stibbard and Lee (2006) can also be accounted for in terms of accent similarity for L2 speakers, regardless of whether the talker and listener shared the same L1. For instance, none of the French listeners demonstrated a particular advantage for KO speech, and performed quite poorly on this accent. The acoustic analysis revealed that this poor recognition was due to the large acoustic distances between the accents of the French listeners and the Korean talkers, mirroring the findings in Stibbard and Lee (2006). It can be speculated that an interlanguage benefit effect could have been created if the French listeners had been exposed to a L2 accent that is acoustically closer to FE speech, such as Spanish or Italian-accented English, as shown in Bent and Bradlow (2003) with Chinese and Korean-accented English. In fact, a similar L2-L1 intelligibility effect was found for a different group of L2 listeners. The same experimental design was used in a Master student's research project in which the speech recognition experiment was conducted on three groups of listeners: L2 experienced, L2 inexperienced German and SE listeners. The SE and L2 experienced listeners showed a selective tuning pattern for SE speech, much like the SE and FB listeners in the present study, while the inexperienced listeners showed a more graded accent processing pattern, also with an advantage for SE speech. The acoustic analysis revealed that both the German listeners groups' accents were closer to SE speech, which explains their advantage for this accent, despite the inexperienced listeners' low proficiency in English and low experience with SE speech.

The interlanguage benefit effect shown between talkers and listeners sharing the same L1 also appears to be largely affected by L2 proficiency, with low proficiency listeners showing an intelligibility benefit for low proficiency L2 speakers (e.g., van Wijgaarden et al., 2002; Pinet and Iverson, 2010). In the present study, only the FI listeners showed a strong advantage for FI speech. Again, this intelligibility effect was shown to be enhanced by the acoustic similarity in the accent of the talkers and listeners, indicating that only listeners who are at a beginner stage of their L2 acquisition benefit from the allophonic realisations present in the speech of fellow inexperienced talkers. Therefore, the findings from the present study strongly suggest that any cross-language interlanguage intelligibility benefit effects depend heavily on the talker's and listener's L1 phonological system and thus acoustic similarity in their L2 accents.

The mechanism for how talker-listener accent similarity in production affects perception is not entirely clear. The conclusions could be seen to imply that there is a strong perception-production link (e.g., motor theory; Liberman and Mattingly, 1985). However, it is likely to be more broadly true that the phonetic detail of one's productions tend to become correlated, through experience and development, with the current state of the underlying phonological processes used in perception, even though many of the mechanisms underlying perception and production may be independent. For example, there may indeed be cases where an individual can understand an accent that is considerably different from their own spoken accent (e.g., GE listeners' perception of SE speech in Adank et al., 2009). But one's spoken accent is likely indicative, in most cases, of the types of phonological processes and expectations that will also be used when understanding speech. Moreover, the production measures used here were very broad-based (e.g., measurements of all vowels); measures that focus more on individual phonetic contrasts among L2 learners often show weak perception-production links (e.g., Oliver and Iverson, 2010). However, the measures used here may assess more general accent skills that apply more readily to perception and production.

A future direction is to apply the ACCDIST metric to individual talkers (Iverson and Pinet, in prep.). The present study was concerned with the overall effects of accent, and thus the data was averaged across multiple speakers of that accent to reduce idiosyncratic talker differences. Such individual differences can be due to several factors that are relatively independent from accent, such as basic acoustic characteristics (amount of energy in the 1-3 KHz range, speech rate), style differences (e.g., clear speech), or gender differences (e.g., Markham and Hazan, 2002; Hazan and Markham, 2004; Smiljanic and Bradlow, 2007; 2008). Individual differences constitute an important part of speech intelligibility, and even though they were controlled for to some extent by averaging results across talkers, it would be interesting to investigate their role on the L1-L2 talker-listener interaction in future research. In addition, so far, the metric has only been tested on vocalic measurements and, even though the results revealed strong correlations, future implementations would involve expanding the measure to consonants in order to undertake a thorough investigation of L1 accent interactions, with the aim of exploring the contribution of acoustic similarity to the high intelligibility of standard or 'prestige' accents in the UK. The next step was to apply the metric to cross-linguistic measures of accent similarity, since the contribution of acoustic similarity to accent intelligibility in noise has, so far, been investigated solely in English. In Chapter 5, the metric was tested on vocalic measurement of both French and English in order to address this issue.

4. Chapter four: British English listeners' perceptual adaptation processes to unfamiliar accents

4.1. Introduction

The work described in the two previous chapters has shown that speech recognition in noise is facilitated when the listener's accent matches the talker's and that much of this variance in intelligibility could be accounted for in terms of the acoustic similarity in the accents of the talker and the listener. For instance, the SE listeners performed equally well on both L1 accents in quiet (SE and IE accented speech), but in noise, their recognition processes became selectively tuned to their own accent. That is, they had similarly low levels of intelligibility for unfamiliar L1 and L2 accents, and only had an intelligibility advantage for the accent that matched their own spoken accent. Likewise, the FB listeners showed the same advantage for SE speech in noise, despite their familiarity with French-accented speech. This selective tuning, again, could be explained by the acoustic similarity in the English productions of the listeners that matched the accent of the SE talkers the closest.

It is plausible that accent adaptation over a prolonged single accent exposure can overturn this selective tuning process. However, the mechanisms involved in perceptual accent adaptation processes are unclear. Previous work has shown that there is an initial processing cost associated with exposure to an unfamiliar accent, followed by decreased reaction times over the first few trials as the listeners adapt (e.g., Clarke and Garrett, 2004). Intelligibility improves over a slightly longer time frame (e.g., 40 trials) than do reaction times, such that listeners typically improve in their recognition accuracy by about 5-15 percentage points (e.g., Bradlow and Bent, 2007; Evans and Taylor, 2010).

This processing cost has been described in Floccia et al. (2009) as a two-stage normalisation process, with initial disruption followed by adaptation leading to a full or partial recovery of baseline comprehension. This initial disruption may be caused by a change in accent, thus creating a ‘surprise effect’. For instance, in Clarke and Garrett (2004), the listeners were warned that there would be a change in talker’s voice but not in accent, and whether the listeners’ recovery to baseline depicts habituation or overall accent adaptation effects is unclear. Still, it is possible that this habituation effect occurs prior to the process of accent adaptation in order for the listeners to overcome the surprise effect when encountering a novel accent. This then indicates that accent adaptation could a be multi-stage, intrinsic speech processing mechanism.

However, there is mixed evidence from previous work for a sustained and robust adaptation to a novel accent occurring after the initial disruption, demonstrated by a significant improvement in recognition accuracy or talker-independent learning (i.e. transfer of learning from one speaker of an accent to another speaker of the same accent; e.g., Clarke, 2000; Gass and Varonis, 1984; Bradlow and Bent, 2007; Clarke and Garrett 2004; Weil, 2001). For instance, Floccia et al. (2009) concluded that the perturbation caused by the presentation of a novel regional or foreign accent doesn’t habituate, at least within the timeframe of accent exposure in their study (up to 15 sentences). Other studies have shown clear effects of adaptation, with listeners showing evidence of accent learning transfer to other talkers. For instance, Bradlow and Bent (2007) showed robust, talker-independent perceptual adaptation effects to Chinese-accented speech by training L1 English listeners on multiple talkers of the accent, then testing their recognition of the accent with a novel talker.

The differences in the adaptation effects' magnitude shown in the literature could also be due to differences in methodology such as length of exposure to the accent, multiple versus single talker adaptation, type of speech (natural L1 and L2 accents, synthesized, vocoded speech) listening conditions (noisy versus quiet listening conditions) and type of measures (e.g., reaction times, word or sentence recognition) (e.g., Davis et al., 2005; Floccia et al., 2009; Clarke and Garrett, 2004; Pisoni et al., 1985). Accent- and talker- dependent factors such as L2 proficiency, gender or overall intelligibility (e.g., Bradlow and Pisoni, 1999; Bradlow and Bent, 2007) have also been shown to affect the listener's ability to adapt to a novel accent. Listening conditions could also affect the magnitude of the adaptation effect. Indeed, the cost associated with processing the phonological variation in novel accents is greater in adverse listening conditions than in quiet and is directly applicable to accent adaptation processes. For instance, Pisoni et al. (1985) compared listeners' processing speed of synthetic speech versus natural speech in quiet and noise and while they found little differences between the two types of speech in quiet listening conditions, there were much longer delays in processing synthetic speech in noise. Other studies have shown comparable effects of adverse listening conditions for novel L1 and L2 accents, with quiet listening conditions showing only small differences in adaptation between L1 and L2 accents compared to noisy listening conditions (e.g., Rogers et al., 2004). Likewise, the lack of adaptation in Floccia et al. (2006) could be due to listening conditions in the experiment since the cost associated with processing variation in novel accents in quiet conditions is relatively small.

One issue that remains to be established is whether L1 listeners can fully adapt to speech that largely deviates from theirs and can achieve the same level of recognition accuracy for both type of speech. In the two previous chapters, the L1 listeners were shown to have poor recognition of L1 and L2 accents that acoustically deviated from SSBE speech. It is predictable that, with a longer exposure to a novel L1 accent, L1 listeners would have little difficulty adapting to it, since both talker and listener would share some L1 acoustic-phonetic features, and the within- and between-speaker variation is smaller in L1 than in L2 speakers. It also likely that the magnitude of the adaptation effect could be affected by listener- or talker-dependent factors, such as accent familiarity (Adank et al., 2009), strength of accent, and intelligibility (notably L2 proficiency, e.g., Bradlow and Bent, 2007). In addition, the mechanisms involved in L2 accent adaptation are likely to differ from those involved for L1 accents. For instance, Bradlow and Bent (2007) showed that the proficiency of the L2 speakers determined the rate of adaptation and suggested that the processes involved differed for L2 and L1 accent adaptation, with possibly more levels of speech processing required (e.g., phonetic, suprasegmental, lexical levels).

Different types of exposure to a novel accent have been tested in order to instigate robust and talker-independent adaptation to novel accents in L1 listeners, with two major types: short term exposure looking at the effects of quick adaption and long term exposure ‘training’ studies. For instance, in Bradlow and Bent (2007), the exposure to the novel L2 accent involved two sessions of high variability training administered over two consecutive days, with the second training session followed by a post-test. Other studies have used a different methodology by presenting one or several blocks of the novel accent to the listener in order to observe learning effects over time (e.g., Adank et al., 2009). Both measures denote different types of adaptation that are typical of real-life speech communication occurrences, the latter reflecting rapid adaptation mechanisms involved in very short interactions, the former being more indicative of the processes involved in long-term accent exposure (e.g., listeners moving to a geographical area where the spoken accent differs from theirs) which has been shown to be resistant to decay over time (e.g., Eisner & McQueen, 2006; Evans and Iverson, 2007).

The use of visual cues and social interaction in accent adaptation has been given little attention in the literature. Indeed, lab speech removes the availability of visual cues and paralinguistic features that are likely to promote adaptation to an unfamiliar accent, in particular when the interaction occurs in adverse listening conditions. Previous work has shown that both L1 listeners listening to strongly L2-accented speech and low proficiency L2 listeners listening to L1 speech rely heavily on visual cues. For instance, in Hazan et al. (2005), auditory or audio-visual training was given to L2 Japanese listeners on a variety of English phonemic contrasts. They found that sensitivity to visual cues for L2 phonemic contrasts can be enhanced via audio-visual perceptual training, with audio-visual training shown to be more effective than auditory training alone when the visual cues to the phonemic contrast are sufficiently salient. They also showed that the availability of the talker's facial gestures lead to a greater improvement in pronunciation, even for contrasts with relatively low visual salience. Likewise, L1 listeners have been shown to rely heavily on audio-visual cues when presented with L2 speech, even with little or no adverse listening conditions. In a later study, Hazan et al. (2010) found that in an audio-visual 'clear' condition of stimulus presentation (i.e., no added noise or other adversity), L1 English listeners showed greater visual weighting for L2 speakers than did L2 listeners. These findings likely have strong implications for adaptation to a novel accent, indicating that visual cues may enhance the perceptual adaptation process. The present study goes one step ahead by using social interaction instead of visual cues to enhance further accent adaptation.

The aim of the present study was to investigate whether L1 listeners' accent selectivity observed in Chapter 2 can be reversed when given the opportunity to tune into unfamiliar accents in single accent blocks or with the presence of social interaction to enhance adaptation. In Experiment 1, L1 listeners were presented with an L1 accent that matched their own ('SE'), a relatively unfamiliar L1 accent (Northern Irish English, 'IE'), and two L2 accents (French- and Korean-accented English, 'FE' and 'KO'). The accents were presented in single blocks as well as in a mixed accent block, in order to evaluate whether individuals had broadly improved on the task within the block or specifically adapted to single accents. The aim of Experiment 2 was to investigate whether accent adaptation could be promoted by social interaction where visual cues and paralinguistic features are freely available. L1 listeners took part in a 15 minute 'spot the difference' Diapix task (Baker and Hazan, 2011) with the same French-accented talker used in Experiment 1. To evaluate whether the listeners benefited from the social interaction with the L2-accented speakers, they performed a speech-in-noise recognition task on sentences recorded by the L2 talker before and after the interaction.

4.2. Experiment 1

4.2.1. Method

4.2.1.1. Subjects

The subjects were 18 monolingual Standard Southern British English listeners ('SE'), aged 22 to 35 (mean = 28 years). They were residing in London at the time of testing and reported having no strong familiarity with the accents presented aside their own accent (assessed by questionnaire; described in the procedure below). None of the subjects reported any speech, hearing or learning difficulties.

The full set of the Bamford-Kowal-Bench (BKB) sentence recordings of one of the female talkers of Standard Southern British English ('SE'), Northern Irish English ('IE'), French-accented English ('FE') and Korean-accented English ('KO') generated in Chapter 1 were used in the present study (Bench et al., 1979; appendix 1). The French talker had low English proficiency ('FI' in Chapter 1); she had learned English at school and was residing in France at the time of recording. The Korean talker was a low proficiency speaker who was residing in the UK. The digitized recordings were embedded in speech-shaped noise with a signal-to-noise ratio from -1 to -5dB; the exact values were selected for individual talkers based on previous data (obtained in Chapter 1), in an attempt to equate intelligibility levels between accents (a target of 70% correct words in sentences). The speech-shaped noise was generated for each individual talker such that it matched the smoothed long-term average spectrum of their speech.

4.2.1.2. Procedure

The subjects performed a sentence recognition task where they listened to the stimuli and repeated what they had heard. Responses were given verbally (i.e., the author marked how many keywords were spoken correctly). Each block contained 56 sentences and each sentence was presented only once (i.e., they were not repeated within or across conditions). The stimuli were presented in a random order within each single and mixed accent block, and presentation order was counterbalanced between subjects. In addition, the mixed accents block was presented either before or after the 4 single accent blocks. In order to evaluate the listeners' knowledge and familiarity of the accents presented in the experiment, after completing each single accent block, they were asked to identify the accent of the talker first, and then their familiarity with the accent was assessed.

4.2.2. Results

Figure 4.1 displays recognition accuracy (i.e., the proportion of words correctly identified in the sentences) for the 4 accents and the mixed-accents block across time (divided into 4 time periods).

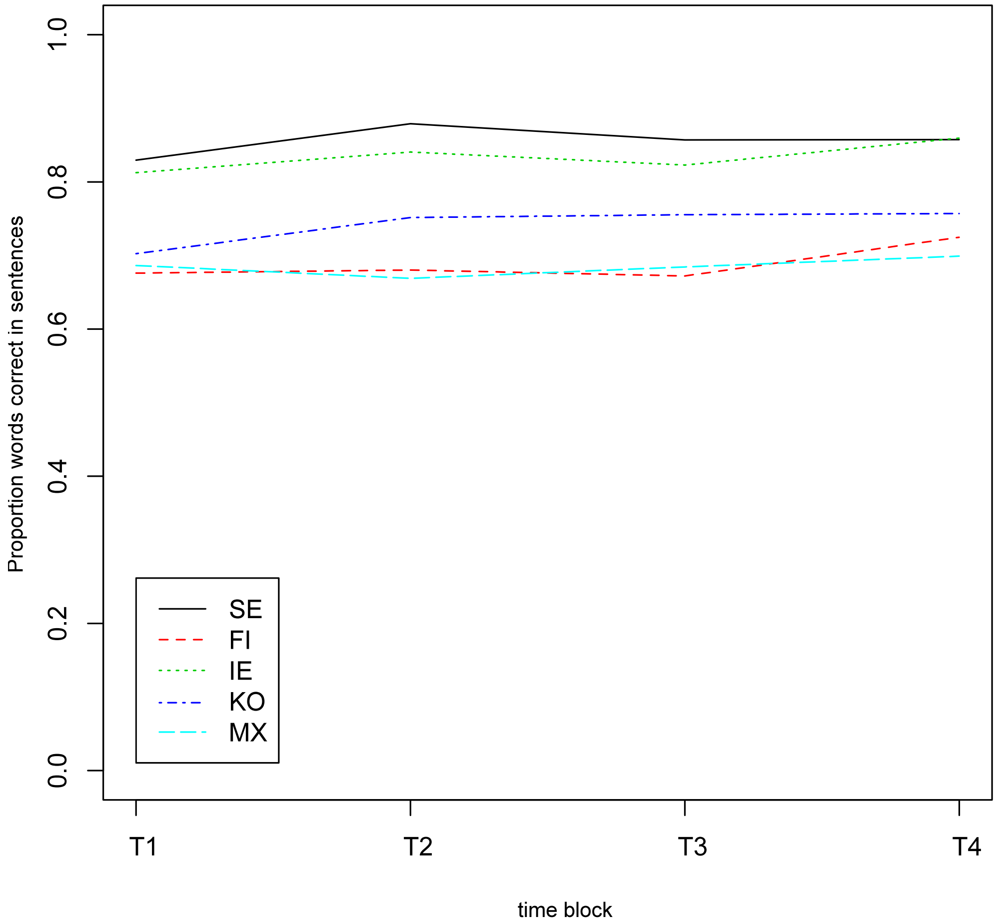


Figure 4.1: Proportion of correctly identified words for each accent condition over time. MX indicates the mixed accent block.

A mixed-effects ANOVA was conducted with time and accent condition as within-subject factors. All analyses were conducted on arcsine-transformed scores. The results revealed significant main effects of time, $F(1, 685) = 6.28$, $p = .01$ and accent, $F(4, 685) = 51.13$, $p < .01$, but no interaction between the two, indicating that the listeners performed differently on the accents presented (i.e., the levels of noise used did not fully equate intelligibility differences). The adaptation effect (i.e., change over time) was small (see Figure 4.1), but significant. However, the lack of interaction shows that the listeners adapted to all accents in a similar manner. Therefore, the overall moderate effect of learning indicates no additional learning of the unfamiliar accents over their own.

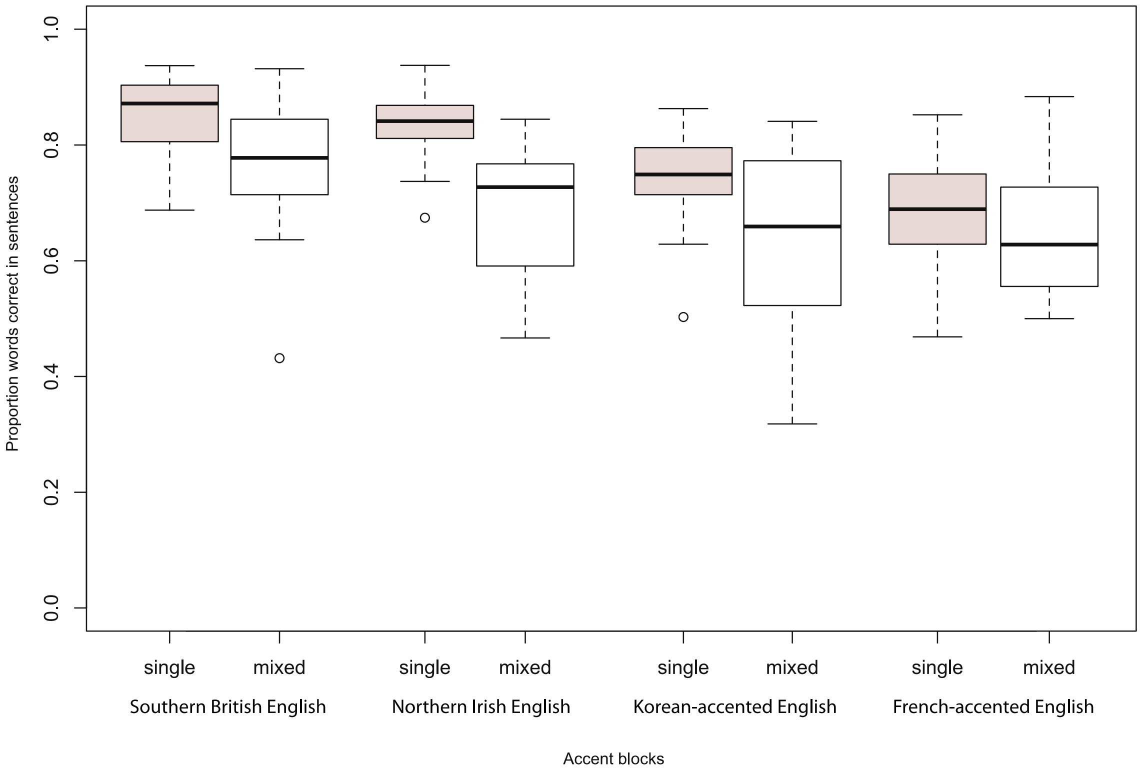


Figure 4.2: Boxplots comparing the listeners' performance on the four accents in single (grey boxplot) versus mixed (white boxplot) accent presentation conditions. Boxplots display the quartile ranges of scores.

Figure 4.2 displays boxplots of the listeners' performance on the 4 accents in single versus mixed accent blocks. The performance in the single block refers to the condition in which the accent was presented on its own, and the mixed accent block refers to the listeners' performance on that same accent when it was presented with the other accents in randomised order. The listeners performed worse on the mixed-accent blocks over the same single-accent blocks, showing that recognition was facilitated by a continuous exposure to a single accent. A mixed-effects ANOVA was conducted with blocking and accent as within-subject factors. The results revealed significant main effects of blocking, $F(1, 115) = 35.97$, $p < .01$, and accent, F $(3, 115) = 17.76$, $p < .01$, and a significant interaction between the two, $F(3, 115) = 2.83$, $p < .05$. Overall, the results thus demonstrate that there were significant advantages for individuals listening to only a single talker and accent within each block, which suggests that there is some accent adaptation. The blocking therefore likely enabled the listeners to better tune into the accent of the talkers, compared to when the accents were presented randomly. The significant interaction likely occurred due to there being slightly less of a difference for the FE accent between the two types of blocking.

4.3. Experiment 2

4.3.1. Method

4.3.1.1. Subjects

The subjects were similar to those in the first experiment. They were 19 monolingual Standard Southern British English listeners (‘SE’), aged 19 to 51 (mean = 26 years). They were residing in London at the time of testing and reported no speech or hearing difficulties. None of the subjects spoke French fluently and only one of them had some familiarity with French-accented speech (high proficiency L2 French speaker).

The low proficiency French-accented female talker from Experiment 1 took part in the Diapix task with the SE speakers (described below) and her BKB sentences recordings were used again for this experiment. The recordings were embedded in speech-shaped noise with signal-to-noise (SNR) ratios of -9, -6, -3, 0 and +3 dB and were also presented in quiet.

4.3.1.2. Procedure

The listeners were first presented with a block of 30 BKB sentences with 5 sentences for each noise level (pre-test), and performed a sentence recognition task as described in Experiment 1. The stimuli and noise levels were presented randomly and counterbalanced between subjects. Next, the listeners performed a 15 minute Diapix task (Baker and Hazan, 2011, Hazan and Baker, 2011; Figure 4.3; appendix 3) with the French-accented talker where both participants were given scenery pictures with discrepancies and had to find the differences without seeing each other's picture. The participants were sitting facing one another in order to recreate a natural conversation and optimize the availability of visual cues during the interaction. They were encouraged to have a balanced conversational exchange despite the talker's low proficiency in English so that the SE listeners could get enough exposure to the French-accented talker's speech.


PAUL & KIM'S Cafe
Shack TO RENT
PUSH ONCE
BEACH VOLLEYBALL LESSONS
SHORE FORECAST SUN


TOM & PAUL'S Cafe
PEACH JUICE
PUSH HERE
SHORE FORECAST RAIN

Figure 4.3. Example of the Diapix pictures given to the participants to elicit dialogue by conducting a 'spot the different' task.

After the Diapix task, the listeners performed another speech-in-noise recognition task on a second block of 30 BKB sentences (post-test) in order to evaluate the benefit of the face-to-face interaction with the talker on their accent adaptation processes. None of the stimuli presented in the pre-test were presented in the post-test.

4.3.2. Results

Figure 4.4 displays recognition accuracy (i.e., the proportion of words correctly identified in the sentences) for FE-accented speech across all listening conditions pre- and post-test.

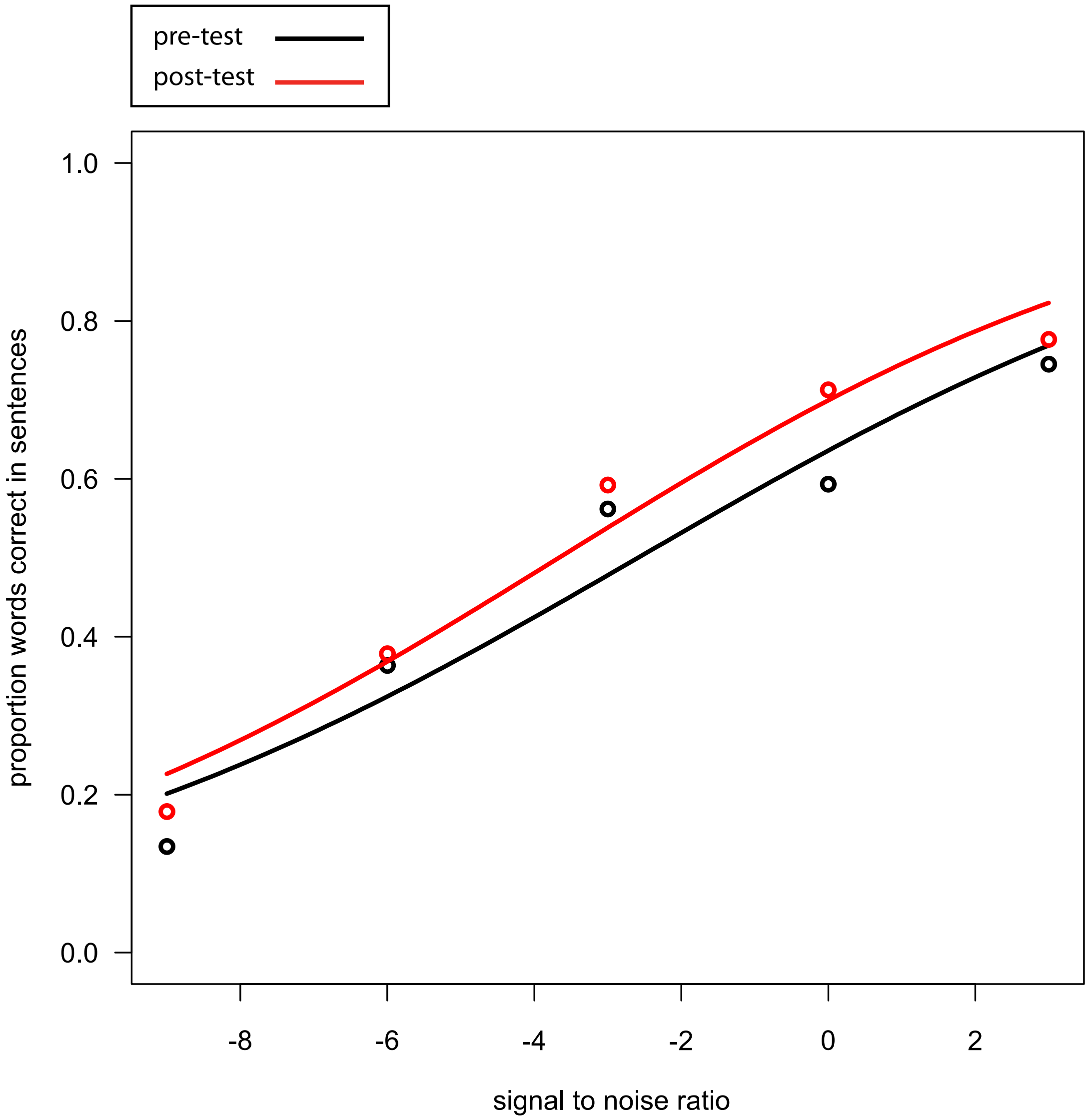


Figure 4.4: Proportion of correctly identified keywords pre- and post-test as a function of SNR.

The listeners performed better in the post-test and therefore seemed to have benefitted from the interaction with the talker. In order to test the difference in performance pre- and post-test, a mixed-effects ANOVA was conducted with test (pre- and post-test) and noise as within-subject factors. All analyses were conducted on arcsine-transformed scores. The results revealed significant main effects of test, $F(1, 9.90) = 7.22$, $p < .01$ and noise, $F(5, 9.90) = 148.73$, $p < .01$, but no interaction between the two. The listeners thus improved significantly on their recognition of the FE speaker's accent in the post-test, but the adaptation effect was relatively small. The significant effect of noise indicates that the listeners performed better on some noise levels, with the largest amount of learning reaching 10% and an average of 5% (0.559 to 0.604) improvement across noise levels. Therefore, the results showed a significant but, again, moderate overall accent adaptation effect as shown in Experiment 1.

4.4. Discussion

The results for Experiment 1 demonstrated a significant but moderate overall learning effect for all four accents. Interestingly, listeners showed similar amounts of learning for all accents, including their own accent, demonstrating that adaptation is a pervasive, if small in magnitude, effect (c.f., Evans and Taylor, 2010; Floccia et al., 2009). However, the uniformity in learning also indicates that the selectivity for an accent near one's own continues to occur even when listeners are given the opportunity to adapt to individual accents and talkers.

One possible explanation for this homogenous learning effect is that the perceptual adaptation is talker-specific rather than accent-specific. Bradlow and Bent (2007) found that listeners, at least under some conditions, are able to generalize their accent adaptation to new talkers with the same accent, as long as they are exposed to multiple talkers during adaptation. It is thus possible that accent adaptation can occur in addition to more talker-specific adaptation. However, in the first experiment, no interaction between the talkers' and listeners' accents was found. Therefore, the homogenous adaptation found for all accents (including SE), seems to indicate a talker- over accent-specific adaptation, whether the listener is exposed to single or multiple talkers. Bradlow and Bent (2007) point out that it is unclear to what extent the relationship between daily accent input and flexibility of speech perception affects the listeners' accent adaptation abilities. Indeed, listeners are exposed to accented speech through interactions with speakers of these accents but also through their environment (e.g., from media exposure), which represents a rich input. This exposure to accented speech, in turn, may result in listeners acquiring a general flexibility of speech perception, thereby enabling them to adapt to a variety of novel accents with ease alongside accent-specific learning. This general and accent-specific flexibility of speech processing could account for the listeners' ability to adapt to the novel accents presented in Experiment 1, but the fact that they also improved on their own accent indicates against it. Instead, it may be the case that the listeners showed different types of adaptation for the different accents, with a talker-specific adaptation for their own accent and accent-specific adaptation for the other, novel accents. However, this is unlikely given the uniformity in learning they displayed.

First exposure to an unfamiliar talker and accent involves a processing cost for the listener (e.g., Adank et al., 2009). Floccia et al. (2009) suggest that the speech perception system is perturbed by the presentation of a novel accent, and that this perturbation does not habituate, in particular when the accent exposure is short (e.g., a few minutes of exposure). As mentioned earlier, this perturbation or processing cost is assimilated to a surprise effect, and in their study, both regional and foreign accents triggered a delay in word identification processes and the listeners didn't show any habituation effects. It is thus possible that in Experiment 1, the listeners had to overcome this surprise effect associated with a change of talker and accent between blocks. This would represent an added element of disruption which would further slow down the talker- and accent- specific adaptation processes the listener faces when presented with a novel accent. However, in Experiment 1, unlike in Floccia et al. (2009) and Clarke and Garrett (2004), all efforts were made to minimize this surprise effect. The listeners were given a short break between each accent block and reminded about the change of talker and accent, while in Experiment 2, the listeners were told they would be listening to French-accented speech at the start of the pre-test and knew they would be listening to the same talker for the post-test. Still, the overall learning patterns mirrored that of the above-mentioned studies.

Moreover, the results showed that the amount of adaptation to the accents was quite minimal. One possible explanation for this small learning effect is the timeframe of exposure to the accent. The listeners were exposed to 56 BKB sentences per block in Experiment 1, representing 6 to 7 minutes of exposure to the accented talker, and it is possible that this length of exposure is not enough to promote robust learning effects. However, Clarke and Garrett (2004) have shown very rapid adaptation to a novel accent (within 1 minute of exposure), but it could be that adaptation continues beyond this initial rapid learning effect. Also, Clarke and Garrett (2004) measured reaction times rather than changes in intelligibility. The accent-specific improvements in intelligibility found by Bradlow and Bent (2007) included training that was split across two days, although the number of sentences was similar to that used here. It is thus possible that learning effects are better consolidated when listeners have more time to process the accent exposure. Experiment 2 addressed this issue by exposing the listeners to the talker's accent for a longer period of time (a total of 23-25 minutes). Still, the results showed only a moderate learning effect. It is also plausible, however, that rapid adaptation effects are generally quite small, and that it may take much longer-term exposure (e.g., living in a community that speaks that accent) in order to perceive a novel accent as well as ones own (Evans and Iverson, 2004, 2007).

Blocking was also shown to affect the listeners' adaptation processes. Indeed, the listeners performed significantly better on the accents when they were presented in a single-accent block compared to when they were mixed with other accents. However, there is no clear indication that the blocking design changed the SE listeners' accent selectivity, because the difference between mixed and single accent blocks did not reliably vary depending on whether the accent was familiar or unfamiliar. In addition, there was less of a difference for the FE accent, but this talker was also less intelligible overall under the selected noise levels, and the intelligibility level may affect the degree of learning. The relationship between talker intelligibility and the impact on the listeners' ability to adapt to the talker's accent has been examined in Bradlow and Bent (2007). In their study, L1 English listeners were presented with blocks of sentences spoken by single L2 Chinese and Slovakian talkers varying in L2 proficiency (low to highly intelligible). Even though the listeners showed significant and equal perceptual learning for all talkers, the adaption to the accented speech of the low intelligibility talkers was slower compared to that of the more proficient talker. They concluded that the amount of exposure required to achieve a significant improvement in intelligibility increased as baseline intelligibility decreased. That is, the extent to which listeners can adapt to a novel L2 accent relies on the quality of the talker's speech rather than the quantity of the exposure. Therefore, the SE listeners' poor learning of FE accented speech could have been caused by the quality of the talker's productions in English.

It is also plausible that the SE listeners adapted the least to the least intelligible talker because noise may have further impeded learning and affected the talker's speech, and adaptation might be promoted by a different type of exposure. Experiment 2 addressed this issue by presenting the SE listeners with the accented speech in both quiet and noise, but also with the added availability of social interaction to enhance intelligibility and learning, making feedback, visual cues and other paralinguistic features such as gestures available to the listener. Social interaction has been shown to enhance phonetic learning, notably L2 speech. For instance, Khul et al. (2003) exposed 9 and 10 month old L1 American English speaking infants to L1 Mandarin speech in audio only, audio-visual conditions in which a live speaker was shown on a television. The results showed that the infants benefited from the interpersonal interaction offered in the audio-visual condition, indicating that the language learning process doesn't simply require long-term listening but instead is enhanced by social interaction represented by the presence of a live person. Other studies (Naigles et al., 2001) have shown that older children also benefit from this type of interaction (i.e., live person on a TV screen) to learn new vocabulary items in a foreign language, but the learning didn't extend to more complex linguistic aspects such as grammatical structures. It is thus possible that exposure without human interaction may not be sufficient to elicit robust phonetic learning.

Nevertheless, the social interaction in quiet listening conditions in Experiment 2 didn't promote much additional adaptation to the accent compared to an audio-only accent exposure in noise (Experiment 1). It is clear from the literature that training listeners on novel accents with audio-visual cues helps learning compared to audio only conditions (e.g., Thompson and Hazan, 2010). There are several reasons why the social interaction didn't generate a larger learning effect. First, the benefits of exposure to a live person for phonetic learning in Khul et al. (2003) were only shown in infants, and it is possible that adult speakers don't benefit from this type of exposure in the same way. For instance, adult speakers may not use their neural plasticity for phonetic learning in the same manner as infants (i.e., exposure to the L1 reduces sensitivity to foreign language phonetic details, with the decline happening in infancy, between 6 and 12 months of age, Best et al., 1995). It is also plausible, again, that the full length of the exposure to the talker's accent was too short for the listeners to apply enough of the accent's acoustic-phonetic features they have learned to the post-test. In addition, the lack of visual cues and added noise to the signal in the post-test could have affected the listeners' performance in the post-test and therefore may explain the moderate learning effect. A condition with added audio-visual cues and social interaction (e.g., showing the talker speaking the sentences on a screen) could have enhanced the listeners' performance. Further investigation is needed to fully understand the contribution of social interaction to novel accent adaptation.

The persistent selective tuning for the listeners' own accent that held across adaptation conditions shown in Experiment 1 and the moderate L2 accent learning of Experiment 2, combined with the findings of previous research, prompt the question of whether accent adaptation actually happens, and if so, how it happens. Indeed, on the one hand, the short term or rapid adaptation effects are very small, with listeners showing only a moderate learning effect on unfamiliar accents. On the other hand, longer-term adaptation may occur but it likely has the same phonetic interaction difficulties that we find in L2 learning and happens over a prolonged length of time and continuous accent exposure. In fact, the selective tuning found for SE listeners may not be reversible after only a relatively short exposure to the novel accent. Indeed, accent adaptation is cognitively demanding; the listener not only has to learn talker-specific and accent-specific acoustic characteristics of the accented speech, but also cope with individual differences and within-accent variation (e.g., degree of accentedness, L2 proficiency). Further investigation is needed to examine the mechanisms involved in long-term accent adaptation.

5. Chapter five: Cross-linguistic accent processing in English and French speakers: effects of L2 experience and acoustic similarity in the talker-listener accent interaction.

5.1. Introduction

Most of the published research on accent processing has focused on English (i.e., how L1 and L2 listeners of English process accented English), but the overall scientific goal is to understand general principles of speech communication, not just details of a single language. The investigation of accent variation and speech in noise in other languages would provide us with a wider view of speech recognition so as to understand its architecture, irrespective of the particular language examined. A minority of studies have reported accent processing in other languages (e.g., Floccia et al., 2006), while cross-linguistic studies of accent processing in noise have mainly focused on comparisons of L1 and L2 talker-listener interactions within the same language (e.g., van Wijgaarden et al., 2002). The L1-L2 accent processing study reported in chapters 2 and 3 revealed that L2 experience and acoustic similarity strongly contribute to the talker-listener accent interaction, showing that speech recognition in noise is facilitated by a match in the talker's and listener's accent. One outstanding question emerging from this research is whether these speech processing patterns are specific to English or are language independent, and if the latter, how do the findings extend to other languages. For instance, it is unknown whether the SE listeners' selective tuning processes in noise is also characteristic of all monolingual speakers processing accented speech in their L1 (e.g., French monolingual listeners selectively tuning to Parisian

French), or specific to English. Likewise, the graded sensitivity approach to accent processing displayed by the inexperienced L2 French listeners could be language-independent and a process typical of low proficiency L2 listeners processing speech in their second language. The present study aims to address this issue by undertaking a thorough cross-linguistic investigation of accent processing.

Language experience strongly affects how multilingual individuals process speech, with L1-L2 interactions occurring at several levels of speech processing. These types of interactions have been extensively documented by Flege's Speech Learning Model, 'SLM' (Flege et al., 1995; 1999; 2002, see also Flege, 1995, 2003). According to the SLM, the elements constituting the L1 and L2 phonetic subsystems of a bilingual or L2 speaker exist within a shared phonological space, and will inevitably influence one another. The nature of the L1-L2 interactions varies as a function of the state of development of the L1 phonetic system when L2 learning begins. L1 and L2 speech sounds may interact through category assimilation, where a L2 speech sound is assimilated to the nearest L1 category (e.g., English /ɪ/ assimilated to French /i/ for L2 French speakers), or phonetic category dissimilation where a new category has been established for an L2 speech sound. For instance, in a study of age effects on L2 speech acquisition and language interaction, Flege et al. (2003) tested L1 speakers of Italian who learned English when they emigrated from Italy to Canada and varied in age of L2 acquisition on their production of English vowels. The results showed that some of the Italian–English bilinguals produced the English /eᶦ/ vowel with little tongue movement whereas others produced it with too much movement. The findings supported the hypothesis that the L1 and L2 phonetic subsystems of bilinguals interact through two distinct mechanisms, phonetic category assimilation and phonetic category dissimilation. The present study aims to further investigate these L1-L2 interactions by testing speakers with a wide range of language experience and providing a cross-linguistic comparison of their phonetic system interactions.

The pattern of linguistic interference in Flege's SLM has been widely documented for L2 learners of English, in which the term 'bilingual' is used to describe any type of L2 speaker (e.g., early, late bilingual), with minor attention given to balanced, from-birth bilinguals (simply referred to as 'bilinguals' in this thesis). The extent to which the listeners are able to dissociate their phonological systems to minimize inter-lingual interferences, as well as the role language dominance plays in this interaction remain to be clarified. For instance, in Flege et al. (2002), Italian-dominant bilinguals were found to have significantly stronger foreign accents than balanced bilinguals, who had stronger foreign accents than the English-dominant bilinguals (assessed by accent ratings). This suggests that bilinguals who become dominant in their L2 may be able to suppress the influence of their L1 system when pronouncing L2 sentences. On the other hand, Cutler et al. (1989, 1992) found that French-English bilinguals performing at native-like levels in both languages only used one rhythm-based segmentation procedure from their dominant language, showing some inter-lingual interference. In the previous chapters (2 and 3), the French-English bilinguals displayed accent processing patterns in English that paralleled the monolingual L1 English listeners' patterns. It was concluded that these listeners revert to a monolingual way of processing accents when listening to English speech, and this was accounted for by acoustic similarity. That they showed no advantage for French accents in noise thus indicated that they could suppress interferences from one language when processing another. The present study will investigate whether bilingual from birth listeners display the same type of monolingual accent processing in French in order to provide some insight into how bilinguals organize and use their two languages. An added measure of L2 proficiency and language dominance for bilingual listeners with perception and production (acoustic similarity) measures will provide a better understanding of how their phonetic systems interact, as opposed to subjective measures of accent ratings provided by L1 listeners.

In addition, Flege's notion of L1-L2 phonetic interference might be interpreted as a type of flexibility of processing that varies as a function of language experience. For instance, in Pinet and Iverson (2010), highly experienced French L2 listeners varied the L1 and L2 acoustic cues they used to process accented speech in English according to noise, showing some flexibility of processing. Thus, it is plausible that highly proficient L2 listeners may display distinctive speech processes. Even though multilingual individuals may not reach the same speech processing abilities as monolinguals in each of their languages, their multilingual experience may offer them more flexible general speech processing systems. The present study aims to test this hypothesis by including speakers with a broad range of L2 proficiency in the sample.

It is also plausible that listeners of a language other than English may process accents in a manner that differs from the one observed in the previous chapters, as phonological systems vary between languages. For instance, French has fewer vowels than English, and vowel differences are the major cause of accent differences within English. Thus, it may be that accent variation affects French listeners differently because they process vowel variation in a different manner, or that accent variation in French does not involve differences in vowels to the same extent as in English. Likewise, French employs a syllable-based segmentation procedure compared to the rhythm-based procedure used in English, a difference which is likely to affect accent processing. In particular, listeners have been shown to rely more heavily on their L1 prosodic cues when segmental cues become unavailable in high levels of noise (e.g., Pinet and Iverson, 2010).

The aim of the present study was to investigate whether the accent processing patterns found in L1 and L2 listeners of English in Chapters 2 and 3 are language-specific or paralleled in other languages. Effects of L2 experience and acoustic similarity in the talker-listener accent interaction were also examined cross-linguistically. The aim is to reveal whether the accent interactions found in English can be replicated in French and whether acoustic similarity can account for the interaction as previously shown. This will indicate if selective tuning processes extend to standard French for L1 and bilingual listeners, if graded sensitivity of accent processing patterns extend to other L2 speakers, and if bilingual from birth listeners display dual monolingual accent processing patterns in English and French. L1 and L2 listeners of both English and French (with varying proficiency in the languages), and English-French bilingual from birth listeners were tested on their speech-in-noise recognition of English and French sentences. In French, the listeners were presented with sentences spoken with Standard French, Quebecois French, and English high and low proficiency accents. In English, they were presented with Standard Southern British English, Northern Irish English, and French high and low proficiency accents. Subsequent acoustic analysis using the ACCDIST metric (Chapter 3) was conducted on the listeners' speech recordings in order to observe any accent interaction effects that were due to acoustic similarity in the accents of the talkers and the listeners.

5.2. Method

5.2.1. Subjects

There were a total of 94 subjects split across three groups: one group of 31 native Southern British English listeners ('E'), one group of 27 English-French bilinguals ('B') and one group of 36 native French listeners ('F').

The English listeners were 18 to 48 years old at test (mean = 22 years), they were native speakers of Southern British English, all had learned French as a L2 and had varying experience and spoken proficiency with the language. They all learned French at school (range age of acquisition: 3 to 13 years old; mean = 9 years) with an average of 9 years of L2 study (ranging from 4 to 18 years of continuous studies, GCSE to BA study level). Fourteen of the listeners spent some time in France and other French-speaking countries (e.g., Mauritius) prior the time of testing for a period of time ranging from 3 months to 10 years (mean = 21.5 months).

The French listeners were aged 18 to 48 years old at test (mean = 25 years). All but 5 of the subjects were residing in London at the time of testing and therefore were very familiar with Southern British English accented speech. They had lived in an English speaking country for a period of time ranging from 1 month to 6 years (mean = 22 months), with some of the listeners having resided in Anglophone countries outside of the UK for a short period of time (e.g., Australia, United States). The other 5 listeners were residing in France when they were tested. They had minimal experience with spoken English and only had taken short trips to London, except one subject who had spent a year in London in the past. The listeners had learned English at school in France (age of acquisition ranging from 6 to 14 years old; mean = 11 years). The range of the English and French listeners' L2 experience was varied on purpose in order to conduct a thorough investigation of individual differences, with L2 experience ranging from the minimal amount of L2 study to comprehend simple sentences in noise to years of residency in the L2 speaking country.

The bilingual listeners had acquired both English and French from birth or at a very young age (age of acquisition of French: 0 to 8 years old, mean = 9 months; English acquisition from birth). The subject who learned French aged 8 had lived in France from the ages of 8 to 13 years old and has had continuous use of the language with a bilingual command of both languages. All the subjects described themselves as balanced bilinguals and had a native-like command of both languages, having been raised in a bilingual environment (e.g., raised by one English and one French speaking parent; raised by English speaking parents in a francophone country). Their spoken fluency in both languages was assessed by the author (a native French speaker with high level of fluency in English). They were 18-32 years old at test (mean = 21 years) and were tested in London.

None of the subjects tested reported any significant speech, hearing or learning disabilities. Their linguistic background (age of L2 acquisition, length of L2 study, time in Francophone and Anglophone countries, languages spoken at home etc.) was assessed by a detailed questionnaire.

5.2.2. Stimuli and apparatus

Recordings were conducted in both French and English. The full set of the 336 Bamford-Kowal-Bench (BKB) sentences (Bench et al., 1979; appendix 1) was translated into French by the author (appendix 2) and was recorded by four talkers each (two males and two females) of Standard French ('SF'), Quebecois French ('QF'), experienced English-accented French ('EE') and inexperienced English-accented

French (‘EI’). In English, the BKB recordings (2 males, 2 females) of Standard Southern British English (‘SE’), Northern Irish English (‘IE’), experienced French-accented English (‘FE’), and inexperienced French-accented English (‘FI’) generated in Chapter 2 were used for the present study again. The FE and EE, and FI and EI talkers matched one another in terms of L2 experience and spoken proficiency and covered the range of proficiencies within the French and English listener groups. The recordings were made in a sound proof booth and a quiet room for the FI recordings with 44,100 16-bit samples per second. Speech-shaped noise was created for each talker based on the smoothed long-term average spectrum of their recordings. The recordings were embedded in this noise with signal-to-noise (SNR) ratios of -6, -3, 0 and +3 dB, and were also presented in quiet (i.e., no added noise). All stimuli were played to the subjects using a laptop over headphones at a user-controlled comfort level.

5.2.3. Procedure

The procedure was very much the same as the one described in Chapter 2. The subjects performed a sentence recognition task in which they listened to the BKB sentences and verbally repeated what they had heard, with the author logging the number of correctly identified keywords. The subjects were given a practice session of 8 English and 8 French stimuli at the start of the experiment to enable the subjects to familiarise themselves with the test. The 16 sentences used in the practice were evenly divided between accents and noise levels (including quiet). The practice block was followed by two blocks of 160 stimuli: one block of French and one block of English language stimuli (i.e., 40 sentences for each of the eight accent conditions, creating a total of 320 sentences for the experimental blocks). The two languages were always presented in separate blocks and the order of language presentation was counterbalanced between subjects. The sentences were also counterbalanced to ensure that they were played in each language, accent and noise level across subjects. Each sentence was presented only once (i.e., not repeated within or across conditions and no versions were presented to the same subjects in both French and English) and the practice sentences were not repeated in the main experiment. Sentences for the different accents and SNR levels (including quiet) were presented in a random order within each block. The mixed accent design was intended to avoid accent adaptation effects, already observed in Chapter 4. The listeners were given a short break half way through the blocks and between blocks when they were reminded of the language change in order to avoid surprise effects. Before the start of the experiment, the listeners were given a language background questionnaire which included an assessment of their experience with the talkers' accents presented in the experiment. This allowed an evaluation of the effects of accent familiarity on the data.

5.2.4. Accent analysis with ACCDIST

After the sentence recognition task, the subjects were recorded reading 31 English and 31 French BKB stimuli sentences that were presented in the experiment, (highlighted in appendices 1 and 2). The recordings were conducted in a sound proof booth for the subjects tested in London and in a quiet room for the 5 subjects tested in France. ACCDIST (Huckvale, 2004, 2007a, b) was used to compare the accents of the subjects to the accents of the same sentences used for the stimuli. In order to calculate the forced alignments for the French recordings, the same HMM models used for the English recordings were employed, but a special purpose French dictionary was generated in which the words of the corpus were given phonemic translations that had English phonemes for which HMM models were available. That is, each French phoneme was matched to the closest English phoneme (e.g., / y / was matched to the English vowel /ʊ/) or assigned to another phoneme if a match was not possible (e.g., the nasal vowel /ɛ̃/ was matched to the English diphthong /eɪ/). ACCDIST was then calculated in the same way for the French and English stimuli (see Chapter 3 for a full description).

5.3. Results

5.3.1. Main accent analysis

Figure 5.1 displays recognition accuracy (i.e., the proportion of words correctly identified in sentences) for the 4 accent conditions across all listening conditions, in English and French.

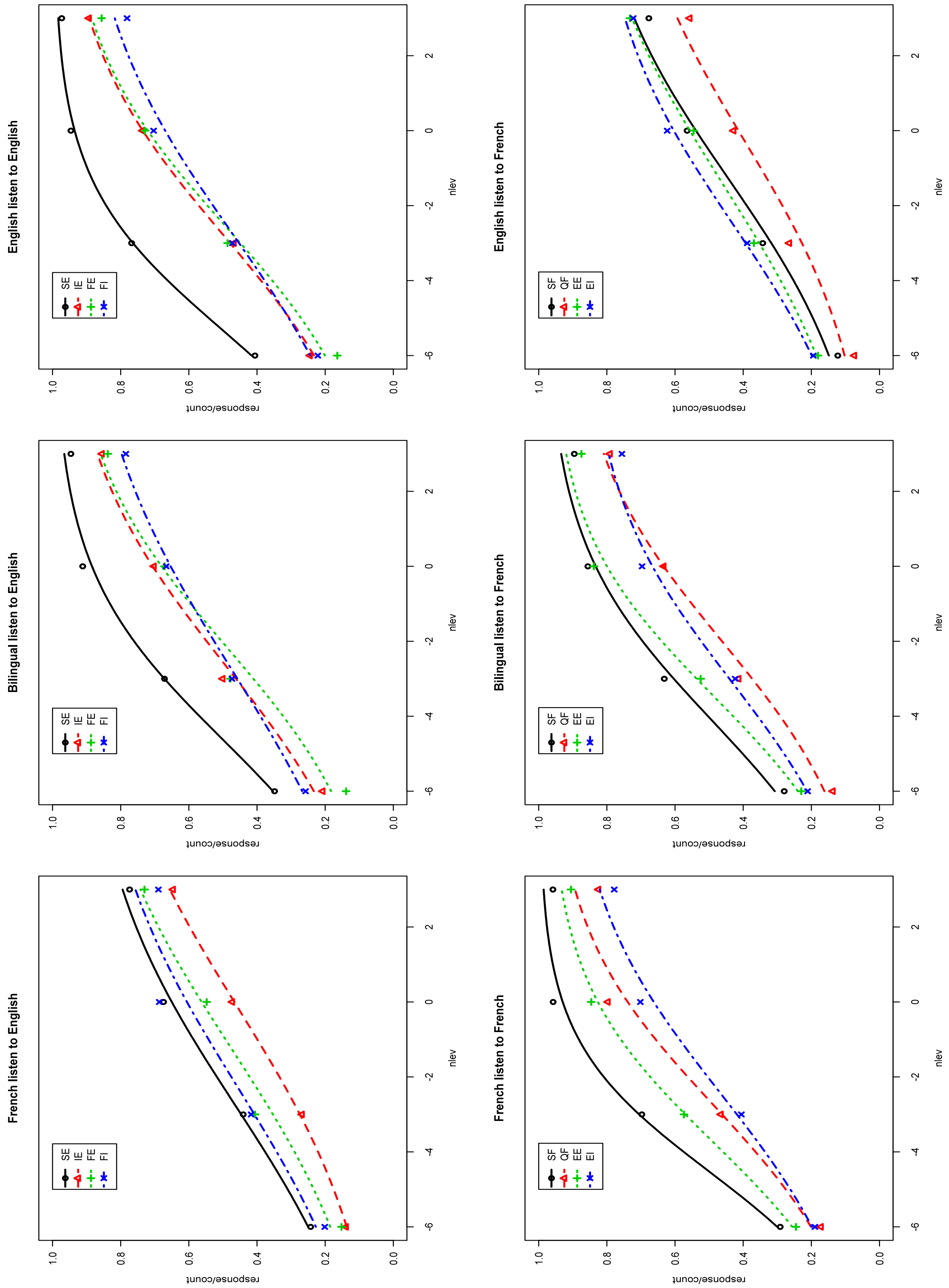


Figure 5.1: Psychometric functions of the proportion of correctly identified words as a function of SNR. The recognition scores in quiet were used to set the maximum of the psychometric functions. The results demonstrated a strong talker–listener accent interaction in both languages, with L2 listeners showing the same type of graded recognition patterns in both

languages but processing of standard accents in English and French differing for L1 and bilingual listeners.

In English, the listeners' accent recognition patterns reflect those of Chapter 2. Experience with L1 English clearly affected how listeners recognized the various accents, with the E and B listeners being the most accurate overall and F listeners being least accurate. It also appeared that the intelligibility of the different accents varied with listener group, with the more experienced listeners showing some selectivity for SE speech, while the less experienced listeners appeared to show more sensitivity to accent differences. The accent recognition pattern was somewhat different in French. The F listeners didn't show strong selective tuning processes for SF speech, but more sensitivity to the different accents. Experience with French also clearly affected the listeners' accent processing patterns. The B listeners had quite similar recognition patterns to the F listeners but had lower overall recognition levels, and the E listeners were the least accurate with the French accents, showing a similar recognition pattern to that of the FI listeners in Chapter 2.

5.3.1.1. English accent perception analysis

In order to test these differences, a mixed-model analysis was conducted with accent condition and listener group as fixed factors and subject as a random factor. The percentage correct was averaged across noise levels to obtain an overall measure of how each listener performed in each condition, and the analyses were conducted on arcsine-transformed scores; the quiet condition was not included in this average, and was only used to set the ceiling of the psychometric functions in Figure 5.1.

Figure 5.2 displays boxplots of the listeners' performance on the 4 accents in English, averaged across noise levels.

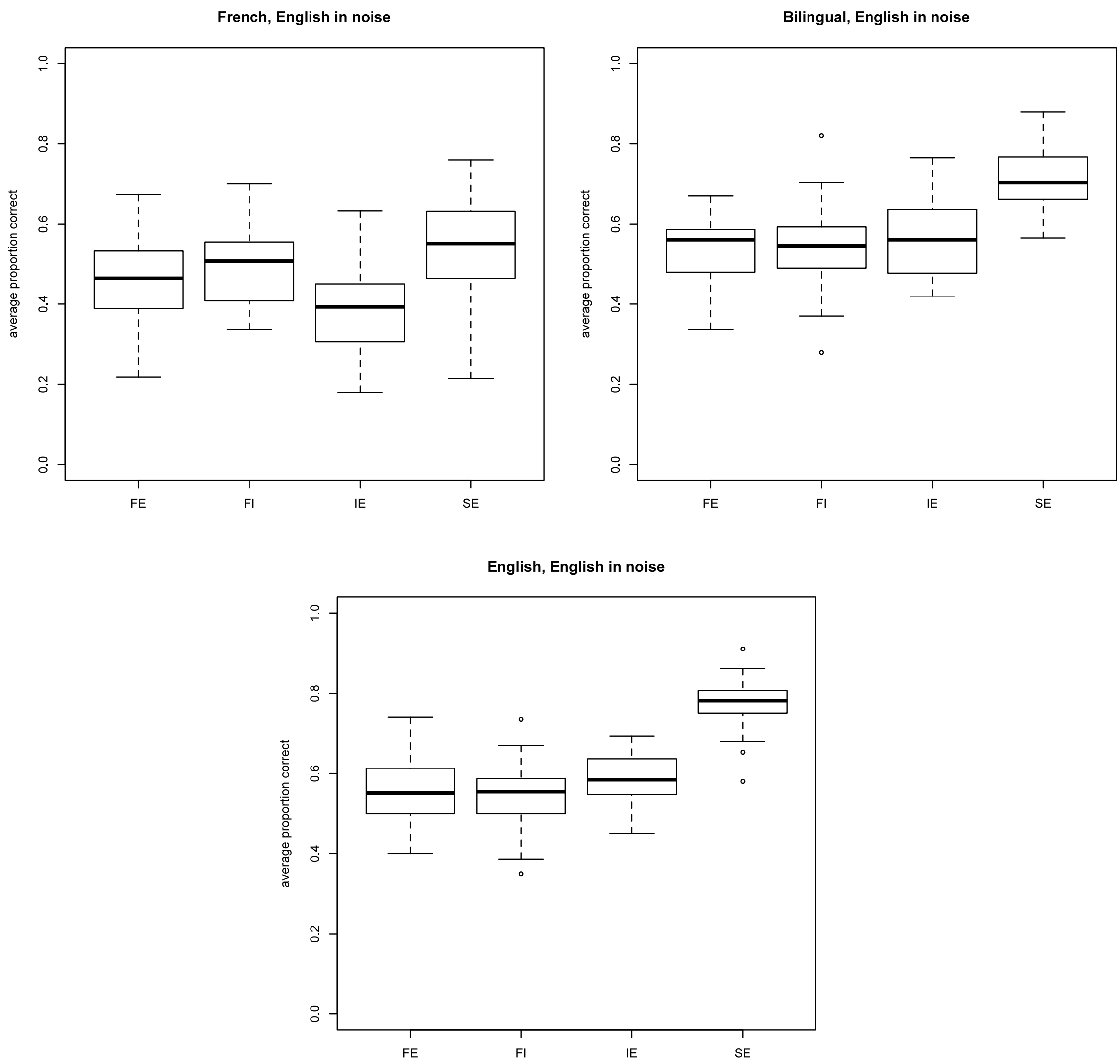


Figure 5.2: Boxplots showing the listeners' performance on each of the four accents in English, averaged across noise levels.

The E and B listeners displayed a strong advantage for SE speech with little differences for the other accents, while the F listeners showed more accent differences, with a small advantage for SE speech, and performed the worst on IE speech. The analysis revealed significant main effects of listener group, $F(2, 91) = 31.04$, $p < .01$, accent condition, $F(3, 273) = 174.20$, $p < .01$, and a significant interaction, $F(6, 273) = 24.56$, $p < .01$.

Accent effects on the recognition of speech in noise

In order to further investigate the interaction between the accents of the talkers and the listeners, mixed-model analyses were conducted separately for each group of listeners with accent as a fixed factor and subject as a random factor. For the F listeners, there was a significant main effect of accent condition in noise, $F(3, 105) = 31.28$, $p < .01$. Tukey tests on every pair of accents revealed that the intelligibility of almost all accents was significantly different, except the SE-FI and FI-FE accent pairs ($p > .05$). Therefore, the listeners showed no significant advantage for SE speech, being similarly accurate for SE and FI accents. These listeners' processing patterns thus resemble that of the FE listeners in Chapter 2 but with added variance due to the broader range of L2 experience in this group of listeners.

In contrast, the data in Figures 5.1 and 5.2 showed that the B listeners were most accurate at recognizing sentences produced by SE-accented speakers but were similar with the other accents. The mixed-model analysis demonstrated that there was a main effect of accent, $F(3, 78) = 73.63$, $p < .01$, and Tukey tests showed that only SE speech was significantly different from all the others, $p < .05$. Therefore, the B listeners were selectively tuned to SE speech, as shown in Chapter 2.

Likewise, Figures 5.1 and 5.2 suggest that the SE listeners were selectively tuned to their own accent, being most accurate at recognizing SE speech and having uniformly lower levels of accuracy for the other accents. The mixed-model analysis demonstrated that there was a significant main effect of accent, $F(3, 90) = 182.43$, $p < .01$. Tukey tests confirmed that the listeners performed significantly better on SE speech than on the other accents, $p < .01$. IE was only significantly more intelligible than the strong French accent (FI), although the magnitude of this effect was small, $p < .01$, and not significantly different from FE speech ($p > .05$). The French accents were not significantly different from each other, $p > .05$. The SE listeners' selective tuning for their own accent thus replicates the findings in Chapter 1.

5.3.1.2. French accent perception analysis

The speech-in-noise recognition data for the French accents differed somewhat from the English data. Figure 5.3 displays boxplots of the listeners' performance on the 4 accents averaged across noise levels in French.

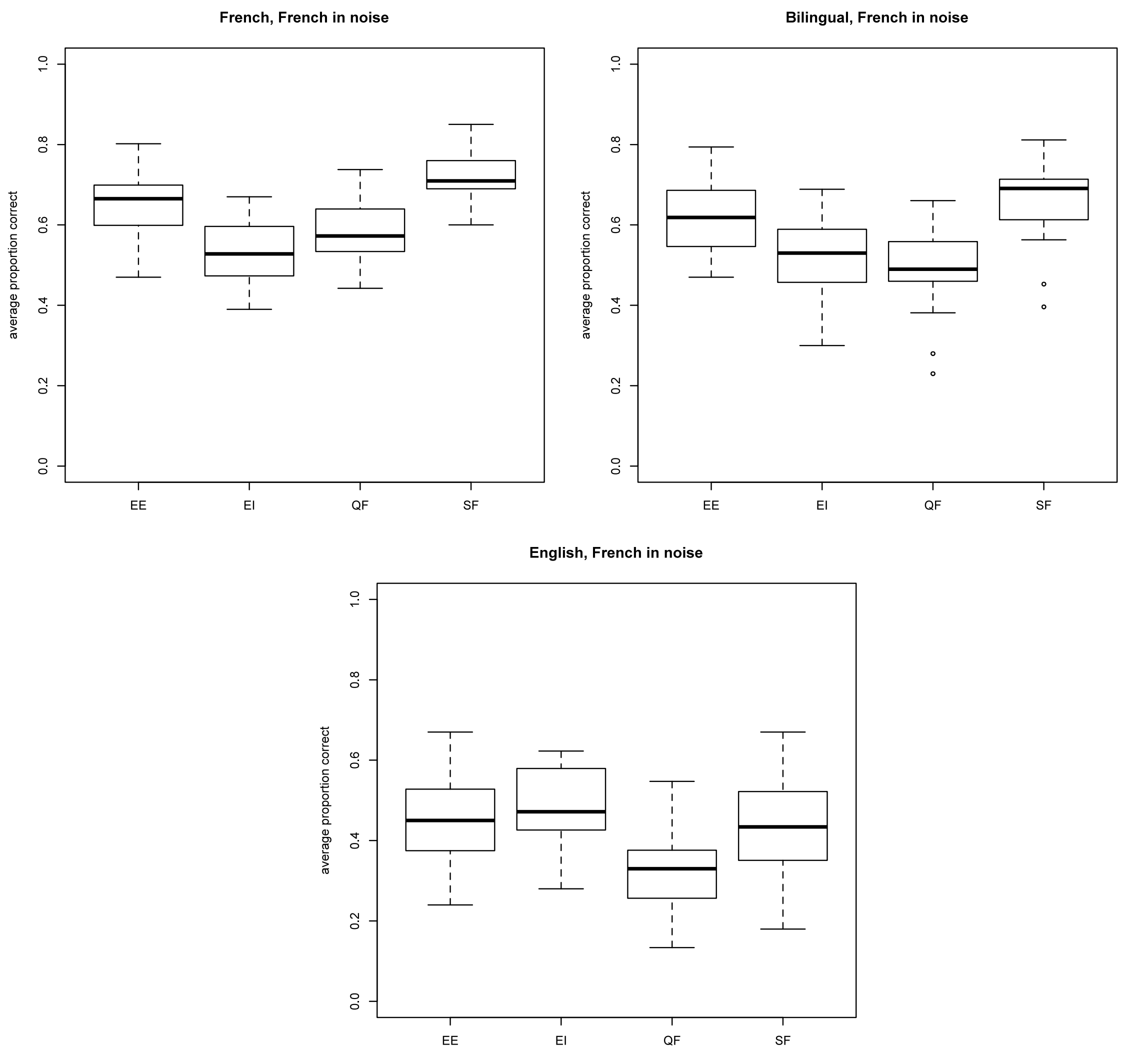


Figure 5.3: Boxplots showing the listeners' performance on each of the four accents in French, averaged across noise levels.

Accent effects on the recognition of speech in noise

The F and B listeners showed a small advantage for SF speech, closely followed by EE-accented speech, while the E listeners appeared to show little intelligibility advantage for any of the accents and performed the worst on QF speech. In order to test these differences, a mixed-model analysis was conducted on the data in the same manner as for the English stimuli. The analysis revealed significant main effects of listener group, $F(2, 91) = 52.53$, $p < .01$, accent condition, $F(3, 273) = 107.66$, $p < .01$, and a significant interaction, $F(6, 273) = 31.27$, $p < .01$.

Mixed-model analyses were conducted separately for each group of listeners to further investigate the interaction between the accents of the talkers and the listeners. For the F listeners, there was a significant main effect of accent condition in noise, $F(3, 105) = 94.76$, $p < .01$. Tukey tests revealed that all pairs of accents were significantly different from each other, $p < .01$, with SF speech being the most intelligible accent, followed by EE, QF and EI-accented speech, thus indicating some graded intelligibility in accent differences, which mirrors the FI listeners' performance in Chapter 2. This indicates that, besides showing an advantage for their own accent, the listeners were much more sensitive to the accent differences. This compares to the E listeners who were very selectively tuned to their own accent.

The boxplots in Figure 5.3 suggest a similar pattern of accent processing for the B listeners with less of a difference in intelligibility for the EI and QF speech. In addition, they had overall lower levels of recognition accuracy in French compared to the F listeners. The analysis revealed a significant main effect of accent condition in noise, $F(3, 78) = 53.66$, $p < .01$. Tukey tests revealed that all pairs of accents were significantly different from one another, $p < .01$ ($p < .05$ for the SF-EE pair), except QF and EI speech that were not significantly different ($p > .05$). Therefore, the B listeners had similar accent processing patterns to the F listeners, with an advantage for SF speech even though it was relatively small in magnitude compared to the F listeners.

The difference in accent intelligibility appeared reduced for the E listeners in Figure 5.3, with a small advantage for their accent and a marked disadvantage for QF speech. There was a significant main effect of accent condition in noise, $F(3, 90) = 32.94$, $p < .01$. Tukey tests showed that QF speech was significantly different from all accents, $p < .01$, indicating that the listeners performed the worst on this accent. The difference in intelligibility between the three other accents was small, with only a significant, but small in magnitude, difference between SF and EI speech ($p < .01$). Thus, the listeners' proficiency level is such that they benefitted from the presence of the English accents, particularly EI speech, mirroring the FI listeners' accent recognition patterns in English in Chapter 2 where they took advantage of the presence of similarly accented FI talkers.

5.3.2. Main production analysis

5.3.2.1. English production analysis

The ACCDIST measures based on MFCC spectra and vowel duration for the English data are displayed in relation to the intelligibility data in Figures 5.4 and 5.5.

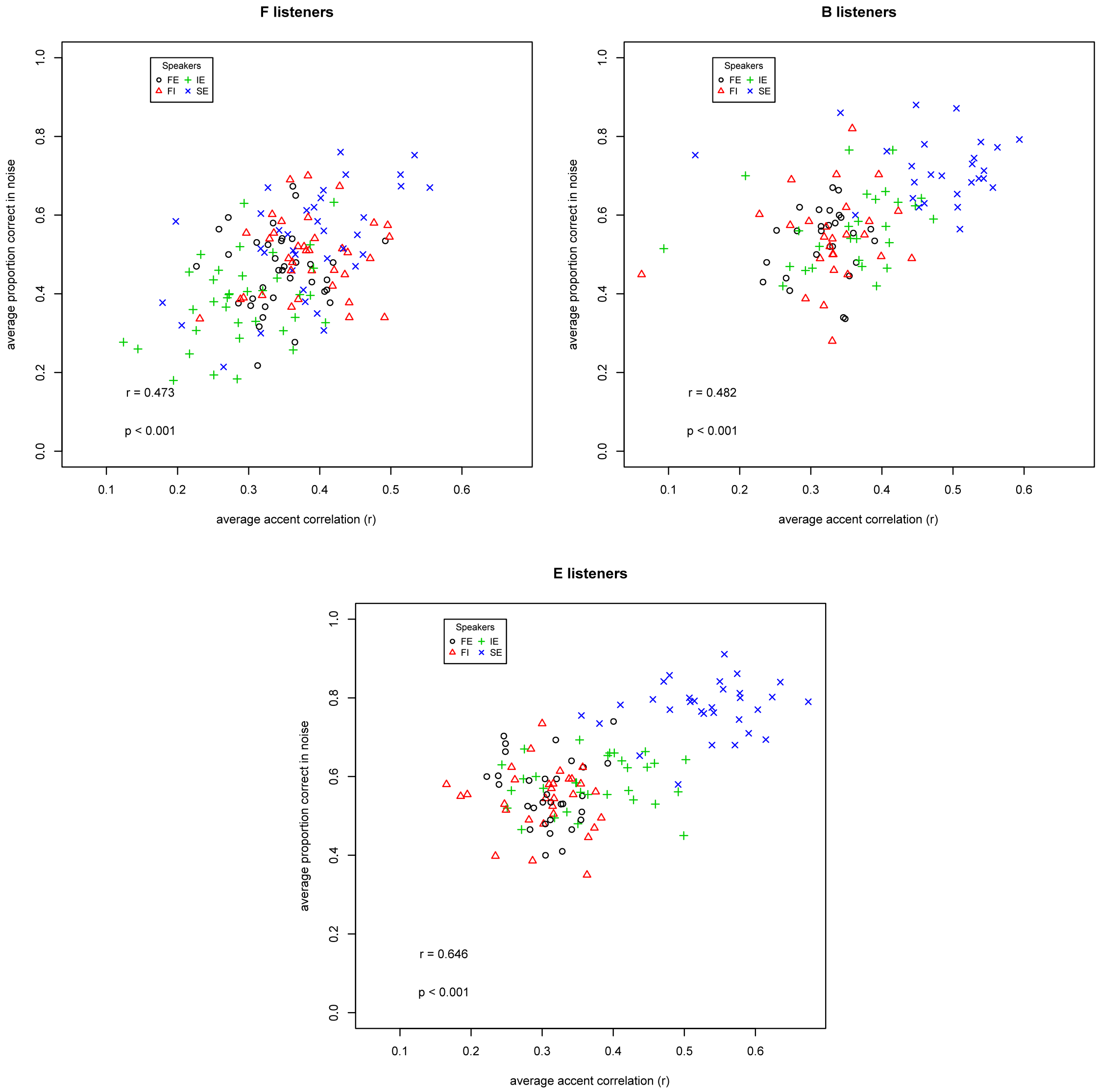


Figure 5.4: Scatterplots of accent correlations in English based on vowel spectra vs identification accuracy for each listener group. The *r* value represents the correlation between the two variables. The *p* value is taken from the mixed-factor analysis of the same relationship, with subject added as a random factor.

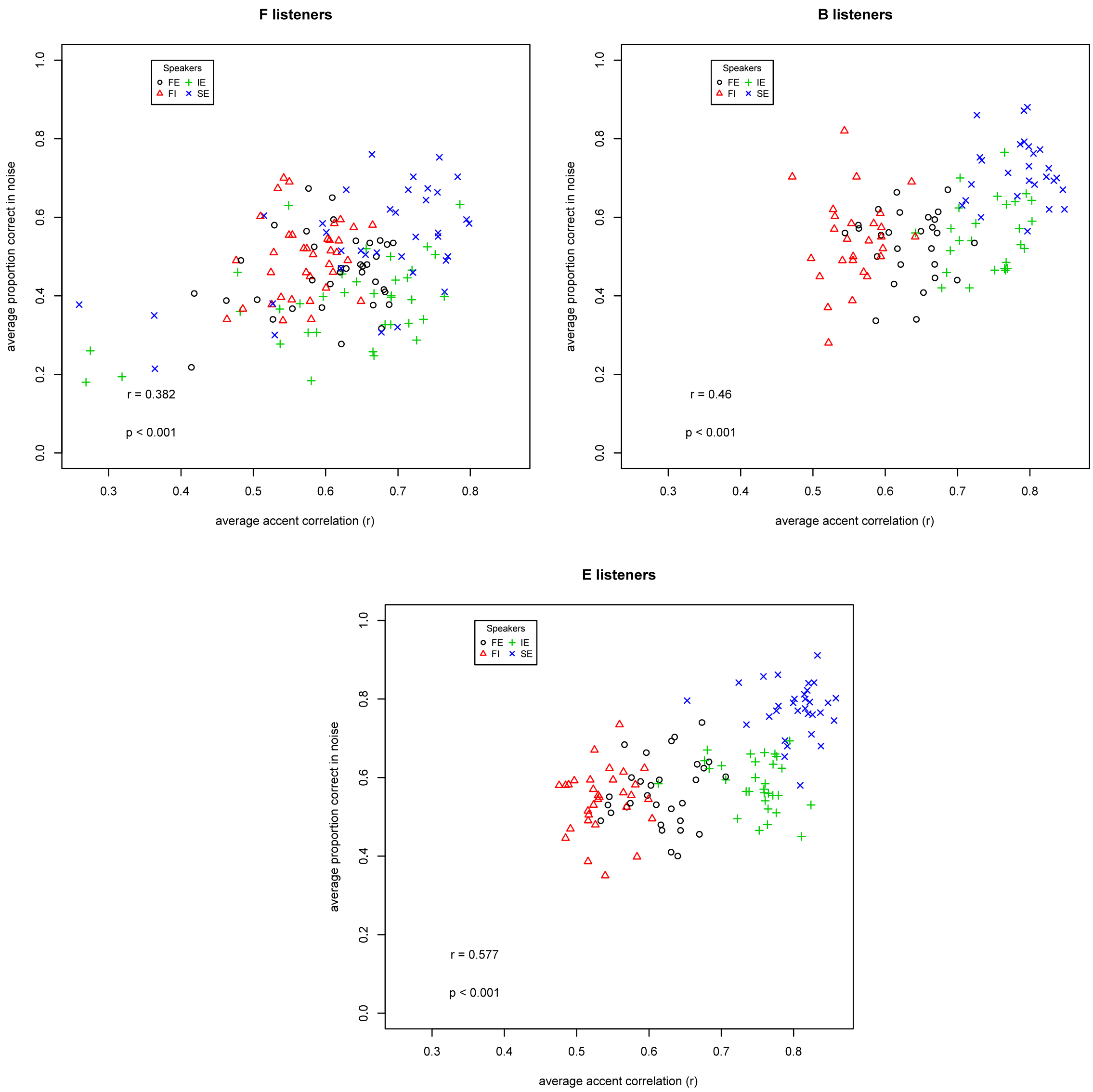


Figure 5.5: Scatterplots of accent correlations in English based on vowel duration vs identification accuracy for each listener group. The *r* value represents the correlation between the two variables. The *p* value is taken from the mixed-factor analysis of the same relationship, with subject added as a random factor.

For each listener, it was plotted how similar his or her own accent was to each of the 4 accents (i.e., averaged across talkers); this is represented by 4 separate points per listener on the scatterplots in each language (i.e. 1 point for each of the 4 accents). Overall, the data displays strong relationships between measured accent similarity and mutual intelligibility for pairs of listener types, and there was little difference in performance between the spectral- and duration-based accent similarity metrics.

The measures appear to be good predictors of how the listeners understand accents in noise. For the vowel spectra measures, the accents of the E and B listeners were closest to that of the SE stimuli, followed by IE and furthest away from the French accents (SE: $r = 0.646$, $p < .01$; FB: $r = 0.482$, $p < .01$), and the pattern of accent similarity was clearer for the durational measurement (SE: $r = 0.577$, $p < .01$; FB: $r = 0.46$, $p < .01$). This mirrored the listeners' relative intelligibility for the different accents and selective tuning processes for SE speech. The correlations were also significant for the F listeners (spectra: $r = 0.473$, $p < .01$; duration: $r = 0.382$, $p < .01$), with Figures 5.4 and 5.5 showing more overlap than the other groups of listeners, and with some of the listeners' speech closer to that of the SE talkers and some further away. This reflects well the listeners' more graded patterns of speech in noise recognition patterns and their individual differences in terms of L2 proficiency.

Figure 5.6 displays data across listener groups for the SE accent only, because in this condition talker-listener accent and English proficiency work in the same direction (i.e., more proficient speakers have an accent that is more similar to SE).

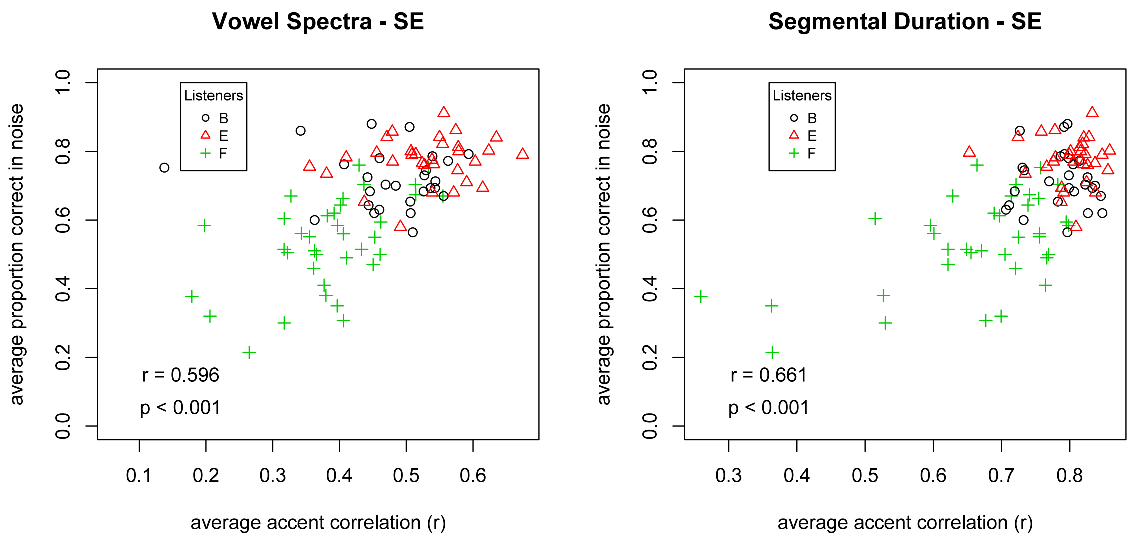


Figure 5.6: Scatterplots for southern-British English stimuli only (SE) of accent similarity (both spectra and duration) vs identification accuracy in noise, for each listener group.

The E listeners were closest to the speech of SE speakers, showing that they were the most proficient talkers, closely followed by the B listeners and the F listeners were the furthest away from SE speech, with their wide range of L2 proficiency clearly displayed. In this circumstance, the correlation between accent distance and intelligibility becomes high for both vowel spectra measurements, ($r = 0.596$, $p < .01$) and vowel duration measurements, ($r = 0.661$, $p < .01$). To illustrate the strength of this measure of proficiency, the same analysis was conducted for the FI accent only, as the relationship works the opposite way to the SE accent only data. The data is displayed in Figure 5.7.

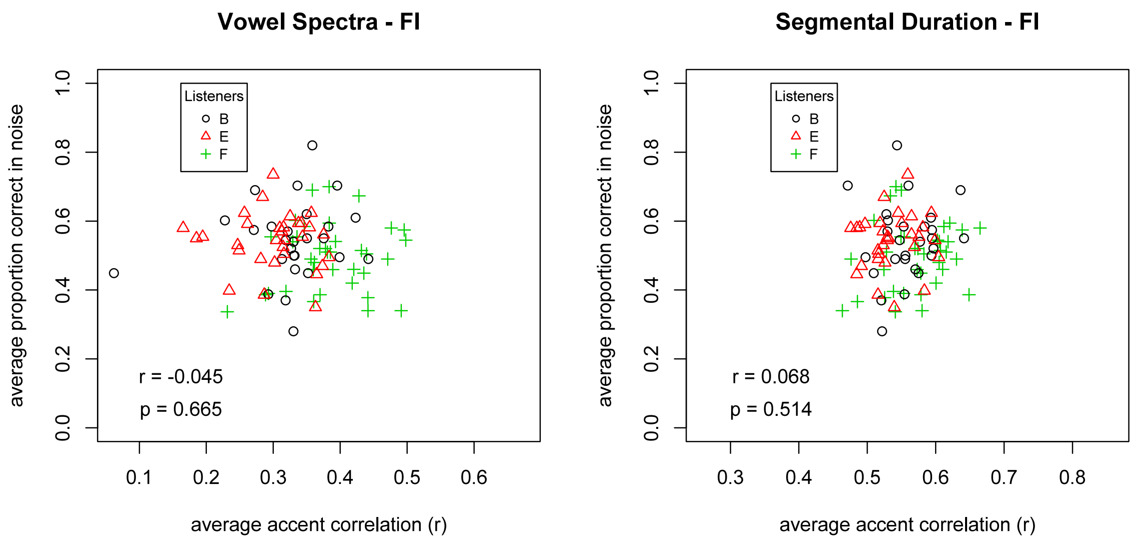


Figure 5.7: Scatterplots for French inexperienced accented English stimuli only (FI) of accent similarity (both spectra and duration) vs identification accuracy in noise, for each listener group.

There were no correlations for the spectral and durational measures. In this circumstance, accent similarity and proficiency work opposite ways and cancel out each other out because similarity should promote intelligibility, but the similarity to FI talkers would show that the listeners are low proficiency speakers.

5.3.2.2. French production data analysis

The ACCDIST measures based on MFCC spectra and vowel duration in relationship to the intelligibility for the French language stimuli is displayed in Figures 5.8 and 5.9. As for the English data, it was plotted how similar each listener's accent was to each of the 4 accents (i.e., averaged across talkers) and this is represented by 4 separate points per listener on the scatterplots in each language (i.e. 1 point for each of the 4 accents).

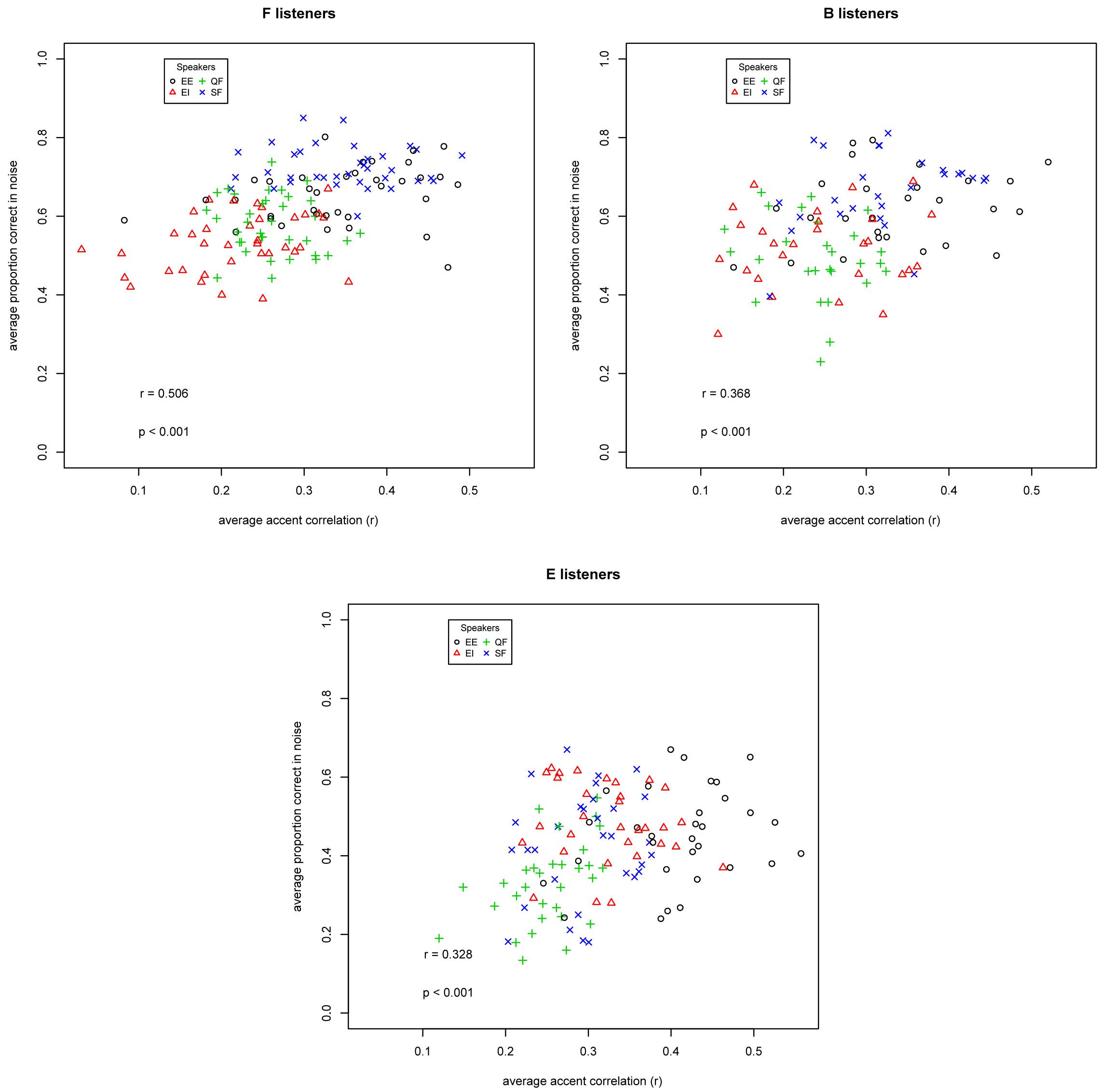


Figure 5.8: Scatterplots of accent correlations in French based on vowel spectra vs identification accuracy for each listener group. The *r* value represents the correlation between the two variables. The *p* value is taken from the mixed-factor analysis of the same relationship, with subject added as a random factor.

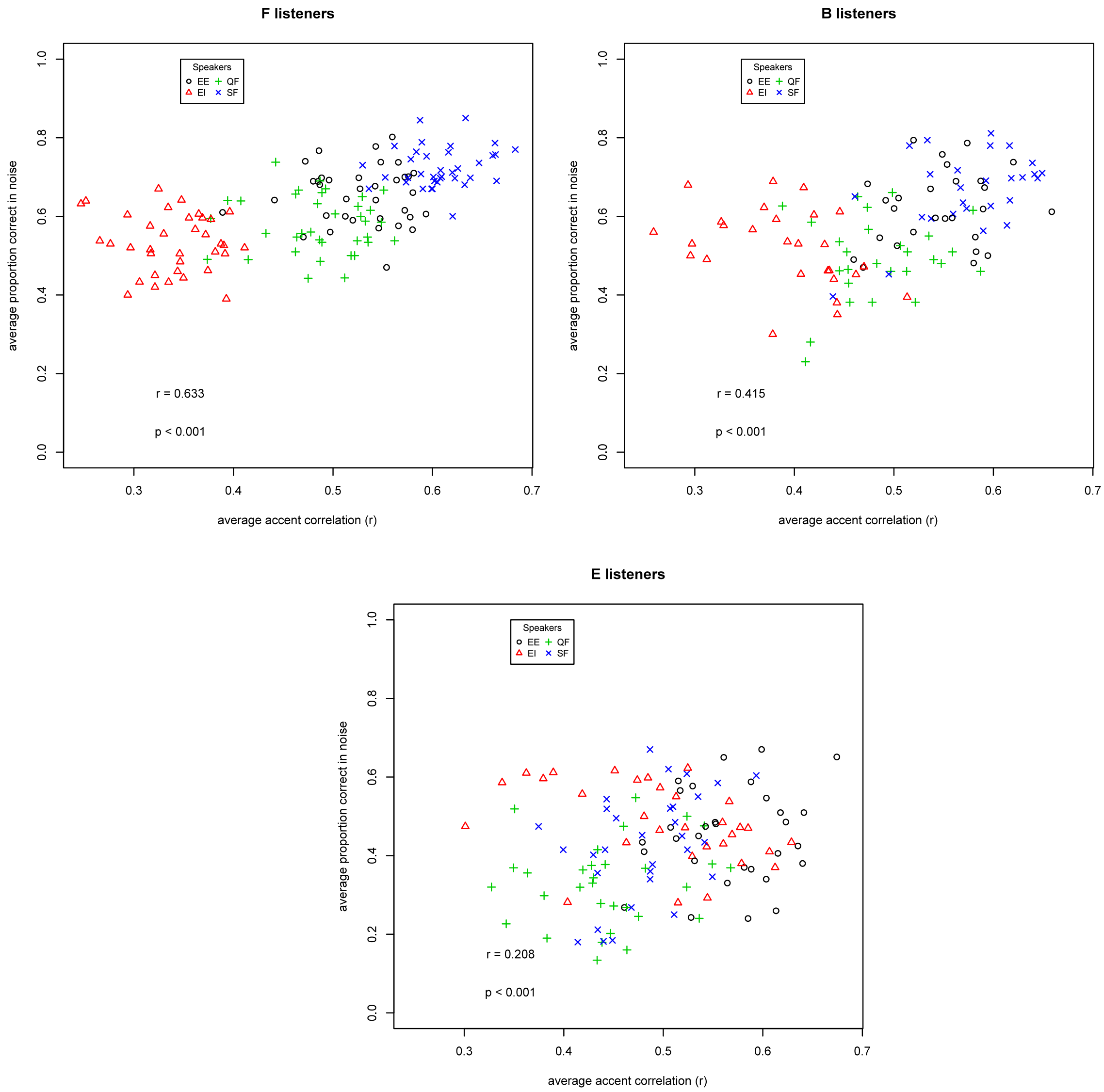


Figure 5.9: Scatterplots of accent correlations in French based on vowel duration vs identification accuracy for each listener group. The *r* value represents the correlation between the two variables. The *p* value is taken from the mixed-factor analysis of the same relationship, with subject added as a random factor.

There was a strong relationship between the F listeners' vowel duration measurements and recognition accuracy ($r = 0.633$, $p < .01$); their speech was closest to the SF talkers, closely followed by EE, QF talkers, and furthest away from the EI talkers, reflecting their accent recognition processes. The vowel spectra measures also showed a strong correlation ($r = 0.506$, $p < .01$) but with equidistant distances between the F listeners' speech and the SF and EE talkers'. The correlation for the duration measures for the B listeners was also significant but weaker than the F listeners' (r = 0.415, $p < .01$), with Figure 5.9 showing a comparable but less clear pattern of acoustic similarity. The vowel spectra measures were not as strong as the duration measures for the F and B listeners. For both measures, the E listeners' accent was almost equidistant to the EI and EE accents and furthest away from QF speech, paralleling their recognition patterns.

The data for the SF accent only is displayed in Figure 5.10.

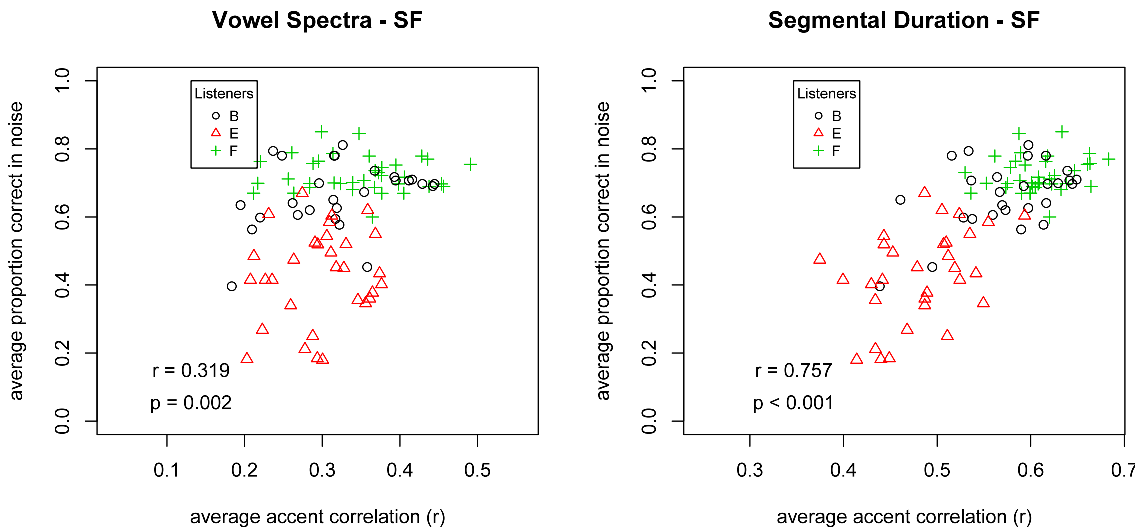


Figure 5.10: Scatterplots for Standard French stimuli only (SF) of accent similarity (both spectra and duration) vs identification accuracy in noise, for each listener group.

As for the SE accent only data, talker-listener accent similarity and French proficiency work in the same direction in this condition, with the more proficient speakers having an accent that is more similar to SF. The vowel spectra measure shows that both the F and B listeners were the closest to the SF talkers' accent, with more variance for the B than the F listener group, and the E listeners were furthest away from SF speech. The relationship was not as strong as for the SE accent only analysis (r =

0.319, $p < .05$), but the vowel duration measures revealed a much stronger correlation ($r = 0.757$, $p < .01$) compared to the spectra measure with the same pattern of acoustic distance, showing that duration is a strong indicator of proficiency. The analysis for the EI accent only displayed in Figure 5.11 revealed a significant negative correlation for the duration measures ($r = -0.399$, $p < .01$), but no correlation for the spectral measurements ($p > .05$), indicating, again, that proficiency is strongly dominated by durational patterns in French.

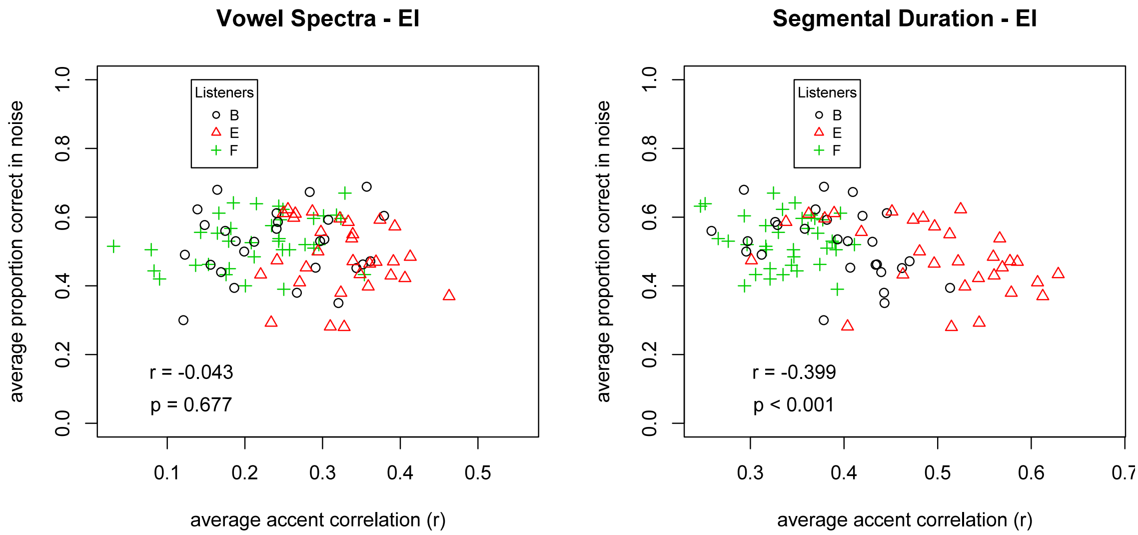


Figure 5.11: Scatterplots for inexperienced English-accented French stimuli only (EI) of accent similarity (both spectra and duration) vs identification accuracy in noise, for each listener group.

Therefore, the acoustic similarity measures for the English stimuli very much replicates the ones found in Chapter 3, even with a broader range of L2 proficiency for the L2 French listeners. However, the pattern was different for the French stimuli, with the results showing that the weight of vocalic cues is lesser for French listeners than it is for English listeners, with durational cues having a stronger impact on the talker-listener interaction than spectral cues. This could be due to the differences in the phonetic systems of the two languages, notably the fact that English has a greater vocalic inventory than French.

5.3.3. Principal component analysis

5.3.3.1. Accent perception PCA

One goal of the analysis was to examine whether the listeners' processing patterns for English accents in noise relate to their processing patterns in French. In order to conduct this large cross-language investigation of accent processing, the data had to be reduced in order to make the comparison simpler. For this, a principal component analysis (PCA) was conducted on the speech recognition data in each language separately with the data averaged across accents and noise levels (including quiet). The analysis produced several factors accounting for the variability in the data. Tables 1 and 2 display the loadings for the PCA for the English and French data respectively.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
SEn6	0.698	0.190	-0.107	-0.450		0.276	-0.242							-0.183			-0.136	0.123	-0.158	
SEn3	0.759	-0.312	0.125		0.168	-0.300			0.217	-0.123			-0.133			0.109	-0.176	-0.131	0.107	-0.152
SE0	0.779	-0.319		-0.135			0.147		-0.152	-0.140	0.100	0.137	-0.195	0.222	-0.180		0.151		-0.140	
SEp3	0.707	-0.489		-0.133					0.111				0.326	0.133	0.198				-0.159	
SEQ	0.676	-0.489		0.189	-0.147	0.121		-0.105		0.195	0.207			0.143	-0.120	0.146	-0.125	0.174	0.104	
IEn6	0.669	0.214	0.386	-0.191	-0.201		0.137	-0.205	0.310			-0.169			-0.139					0.168
IEn3	0.656	0.160	-0.487	-0.107	0.286		-0.208			-0.227						0.121	0.239	0.115	0.152	
IE0	0.757	-0.229			0.261		-0.337		-0.107	-0.105	0.150	0.101		-0.146		-0.205	-0.103			0.165
IEp3	0.773	-0.134	-0.179		-0.130		0.180		-0.116		-0.353	0.121	0.214	-0.142	-0.195					
IEQ	0.715	-0.291	-0.209		-0.298	0.169		-0.130	-0.195	0.206		-0.102	-0.103	-0.166	0.124		0.142	-0.195		
FEn6	0.361	0.198	0.599	-0.530	-0.110			-0.109		0.180		0.261			0.103	-0.112			0.138	
FEn3	0.585	0.604			-0.183		-0.134		-0.133		-0.119	-0.271		0.255		-0.206				
FE0	0.638	0.185	-0.299		-0.183		0.501			-0.233	0.225				0.204		-0.105			
FEp3	0.664	-0.124	-0.148	-0.153		-0.282	0.215	0.418	0.107	0.343		-0.143				-0.132		0.112		
FEQ	0.475	0.202		0.488	-0.432		-0.273	0.238	0.323			0.214								
FIn6	0.457	0.579	-0.337	-0.149	0.265	0.216		0.190		0.196	0.136		0.122	0.104	-0.128	0.108		-0.194		
FIn3	0.525	0.474	0.336	0.355	0.164	-0.164		-0.190			0.273		0.170	-0.133						-0.145
FI0	0.313		0.705					0.457	-0.320	-0.192						0.155				
FIp3	0.617	0.335		0.214	0.106	-0.401		-0.223	-0.295	0.224	-0.180	0.155			0.100	0.138				
FIQ	0.335		0.379	0.380	0.491	0.485	0.189		0.121	0.111	-0.161		-0.126							

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
SSL	7.808	2.113	1.845	1.232	0.985	0.803	0.742	0.681	0.597	0.544	0.446	0.361	0.322	0.306	0.258	0.256	0.221	0.188	0.165	0.127
PVar	0.390	0.106	0.092	0.062	0.049	0.040	0.037	0.034	0.030	0.027	0.022	0.018	0.016	0.015	0.013	0.013	0.011	0.009	0.008	0.006
CVar	0.390	0.496	0.588	0.650	0.699	0.739	0.776	0.810	0.840	0.868	0.890	0.908	0.924	0.939	0.952	0.965	0.976	0.985	0.994	1.000

Table 1: Principal component analysis loadings for the English dataset. SSL indicates SS loadings, PVar, Proportion Variance and CVar, Cumulative Variance.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
SFn6	0.690		-0.140	0.254	-0.163	0.400	0.219	0.262	-0.208			-0.156	0.195						0.121	
SFn3	0.846				-0.233		0.181			-0.200			0.120		-0.159			-0.191	-0.203	
SF0	0.882	-0.291										0.121		-0.112				0.122		0.222
SFp3	0.842	-0.115	0.130		-0.144	-0.164	0.217		0.122			0.137	0.112		0.127	-0.136		0.214		-0.108
SFQ	0.821		0.155		-0.104	-0.158	-0.232	-0.101	-0.189			-0.104	0.194				-0.303			
QFn6	0.617		-0.189	-0.157	0.618			0.139		0.227	-0.256		0.129						-0.106	
QFn3	0.725	-0.353	-0.283	0.233			-0.143				0.212	0.145		-0.120		0.288				
QF0	0.815	-0.253			0.116	-0.167	0.243					0.146	-0.120	0.142	0.142			-0.219	0.168	
QFp3	0.786	-0.104	0.186	0.203		0.261	-0.148					0.224	-0.169		-0.245	-0.150	-0.135			
QFQ	0.803	-0.132	0.163	-0.148					0.106	-0.205	-0.226	-0.189	-0.211	0.207		0.207				
EEn6	0.653	0.338	-0.221	-0.153	-0.257		0.335			0.317		-0.114	-0.231							
EEn3	0.719			0.106	0.415	-0.292			-0.156	-0.199	0.229	-0.154	-0.105		-0.146	-0.115				
EE0	0.801	-0.146	-0.139			0.229	-0.220		0.170			-0.130	-0.145	-0.176	0.247	-0.161				
EEp3	0.777		-0.156		-0.101		-0.247	-0.267	-0.116	0.238	0.110	-0.151		0.280			0.143			
EEQ	0.544	0.228	0.571	0.214		-0.246	-0.135	0.164	0.234	0.240		-0.114		-0.135						
EIn6	0.327	0.502	-0.105	0.650	0.197		0.149	-0.169	0.253					0.141			-0.110			
EIn3	0.436	0.690	0.118				-0.251	0.205	-0.354			0.200								
EI0	0.624	0.442	-0.157	-0.219				-0.475			-0.211			-0.196						
EIp3	0.491	0.372	-0.374	-0.462			-0.131	0.247	0.341		0.156									
EIQ	0.391	0.115	0.585	-0.393	0.284	0.391	0.162	-0.108			0.240									

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
SSL	9.749	1.596	1.209	1.156	0.905	0.686	0.656	0.590	0.558	0.446	0.382	0.336	0.319	0.312	0.248	0.224	0.206	0.179	0.158	0.087
PVar	0.487	0.080	0.060	0.058	0.045	0.034	0.033	0.030	0.028	0.022	0.019	0.017	0.016	0.016	0.012	0.011	0.010	0.009	0.008	0.004
CVar	0.487	0.567	0.628	0.685	0.731	0.765	0.798	0.827	0.855	0.878	0.897	0.913	0.929	0.945	0.957	0.969	0.979	0.988	0.996	1.000

Table 2: Principal component analysis loadings for the French dataset. SSL indicates SS loadings, PVar, Proportion Variance and CVar, Cumulative Variance.

The two first factors accounted for the majority of the variance in English and French and therefore likely highlighted the speech processing differences between the listener groups. The other factors were difficult to interpret for the purpose of the analysis and only accounted for a small amount of the variance; they were thus not considered further. The PCA shows that the first factor (PC1) accounts for the majority of the variance (39% in English; 48.8% in French). In both languages, all of the accent and noise variables were positively loaded on PC1, such that PC1 is a measure of average speech-in-noise accuracy. This indicates that this factor is an overall proficiency measure, with some of the L2 accents (i.e., EI in French, FI and FE in English) having lower positive loadings on this factor, suggesting that the more native-like accents tended to be better indicators of individual differences in overall proficiency.

The second factor (PC2) also accounts for a substantial amount of the variance in the PCA even though its contribution is smaller than PC1's (English: 10.6%; French: 8%; Tables 1 and 2). PC2 is a differential factor with, overall, the L1 accents variables positively loaded and the L2 accents negatively loaded on the factor for both English and French. In English, PC2 shows differences between L1 and L2 accents, while in French, the factor mainly extracts EI speech from the other accents. Thus, PC2 is more specifically tuned to accents, compared to PC1 which accounts for language proficiency.

A mixed model analysis was conducted for PC1 and PC2 to examine whether they varied for listener groups (i.e., whether these factors were sensitive to cross-language differences) with listener group as fixed factor and subject as a random factor. Figures 5.12 and 5.13 display boxplots for PC1 and PC2, respectively, for each listener group in

English and French.

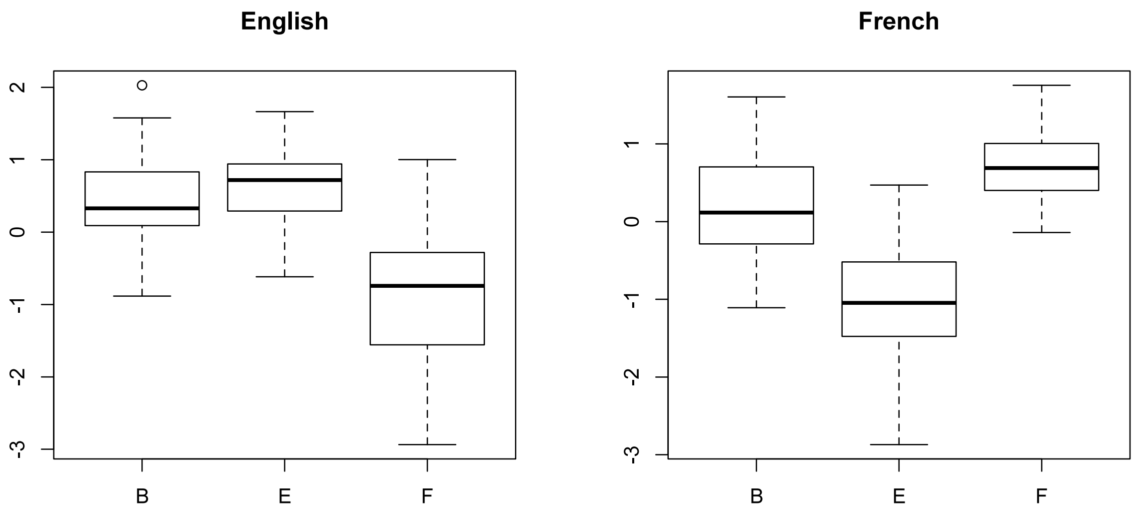


Figure 5.12: Boxplots showing group differences for PC1 (proficiency factor) for the English and French datasets.

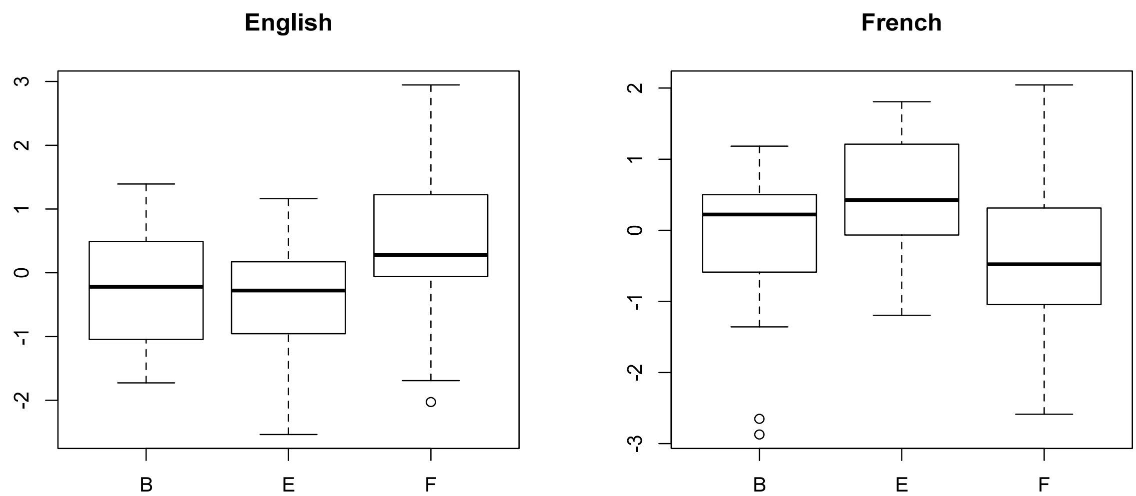


Figure 5.13: Boxplots showing group differences for PC2 (accent factor) for the English and French datasets.

For PC1, Figure 5.12 shows some variation between the listener groups, validating that this component is sensitive to proficiency differences. The analysis revealed significant main effects of listener group in both English, $F(2, 91) = 40.27$, $p < .01$, and French, $F(2, 91) = 54.61$, $p < .01$. Tukey tests conducted on all pairs of listener groups showed that for English, only the F listeners were significantly different from the other groups, $p < .01$ (i.e., significantly worse), while in French, all groups were significantly different from one another. The analysis thus confirms PC1's sensitivity to group difference in terms of language abilities by showing variability in L2 listeners.

Figure 5.13 also shows differences between the listener groups, validating that PC2 is tuned to accent-specific effects. The analysis revealed significant main effects of listener groups for English, $F(2, 91) = 40.27$, $p < .01$, and French, $F(2, 91) = 54.61$, $p < .01$. Tukey tests revealed that, in English, only the F listeners group was significantly different from the other groups, $p < .01$, indicating that this group is more affected by the L1-L2 accent differences than the other listeners. In French, all groups were significantly different from one another, $p < .01$, with a smaller difference in magnitude for the F-B listener pair, $p < .05$. This confirms PC2's sensitivity to group difference in terms of accent processing in both languages.

5.3.3.2. Cross-language interactions

The next step was to investigate how the components highlighting proficiency and accent group differences in English interact with the components in French. First, a Pearson's correlation analysis was conducted across all listener groups in order to observe any relationships across languages. For PC1, the correlation analysis revealed a significant relationship in terms of proficiency between the two languages ($r = -0.214$, $p < .05$), indicating that listeners who were more proficient in English were less proficient in French, and vice versa. There was no significant correlation for PC2 ($p > .05$), indicating no accent processing interaction across languages.

Correlation analyses were conducted separately for each group of listeners. For the E listeners, there was a significant correlation for PC1 ($r = 0.646$, $p < .01$) but no significant correlation for PC2 ($p > .05$), indicating that listeners with a high proficiency in French had a better recognition of English speech. This suggests that the highly proficient in French E listeners benefit from an overall speech processing facilitation effect without L1-L2 accent interference. The results for the F and B listeners also had correlations for PC1 (FE: $r = 0.373$, $p < .05$; FB: $r = 0.427$, $p < .05$) and no correlations for PC2 ($p > .05$), again indicating a general facilitation effect to process speech in the L1 for listeners with a high proficiency in their L2 (or similar levels of high proficiency in both languages for the B listeners). Therefore, the PCA revealed a cross-language speech processing facilitation effect in which listeners with a high proficiency in their L2 benefitted from elevated speech processing abilities in both their L1 and L2.

5.3.3.3. Production PCA

The PCA for the accent perception data revealed clear, interpretable components with strong cross-language interactions within listeners groups. The same analysis was applied to the production data in order to observe parallel trends of cross-language interactions. Tables 3 to 6 display the PCA loadings for the English and French vowel spectra and duration data. However, the PCA generated factors that were not easily interpretable. For example, there were some positively loaded averages similar to the weighted averages in the perception data, but in the context of this analysis, it is less clear how to interpret them. This could be driven by global fluency factors making the listeners more similar to all the accents, but it is unclear what would cause this. Likewise, there were some accent differential factors separating L1 from L2 accents in the data, but these factors were not generally as interpretable as those in the perception data.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
vowel_FEfa	-0.472	0.274	0.553		0.127	0.553		-0.164	-0.169							
vowel_FEfb	-0.354	0.557	0.513				-0.413		0.305	0.107						
vowel_FEma	0.220	0.364	0.524	-0.184	-0.555		0.401	-0.177								
vowel_FEmb	-0.330	0.481	0.588	0.255		-0.378		0.178	-0.196	-0.119		-0.113				
vowel_FIfa	0.366	0.718	-0.202	0.207	0.178	-0.175	-0.137	-0.396		-0.117						
vowel_FIfb	-0.379	0.673			0.444		0.364	0.207		0.106					0.105	
vowel_FIma	0.344	0.733	-0.202	-0.395		0.133		0.169		-0.213	0.156	-0.123				
vowel_FImb	0.216	0.895	-0.219		0.164		0.119	-0.112		0.100					-0.175	
vowel_IEfa	0.835	-0.194	0.175	-0.291	0.281	-0.112		-0.108								0.158
vowel_IEfb	0.317	-0.559	0.593		0.379		0.114				0.208					
vowel_IEma	0.818	0.298	-0.138		-0.104		-0.139	0.159	-0.244	0.278	0.102					
vowel_IEmb	0.845	0.165		-0.345			-0.182		-0.175	-0.101		0.133	0.160			
vowel_SEfa	0.860		0.183	0.149		0.263		0.145		-0.190	-0.221		-0.130			
vowel_SEfb	0.714	0.145		0.610	-0.124	0.108					0.152		0.133			
vowel_SEma	0.888	-0.216	0.210	0.122					0.121	0.160	-0.125	-0.139		0.124		
vowel_SEmb	0.954		0.121									-0.146		-0.171		

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
SSL	6.074	3.592	1.827	0.918	0.842	0.604	0.580	0.400	0.294	0.279	0.184	0.131	0.096	0.072	0.059	0.048
PVar	0.380	0.225	0.114	0.057	0.053	0.038	0.036	0.025	0.018	0.017	0.012	0.008	0.006	0.004	0.004	0.003
CVar	0.380	0.604	0.718	0.776	0.828	0.866	0.902	0.927	0.946	0.963	0.975	0.983	0.989	0.993	0.997	1.000

Table 3: Principal component analysis for the English vowel spectra measures. SSL indicates SS loadings, PVar, Proportion Variance and CVar, Cumulative Variance.

The durational data for the English stimuli is displayed in Table 4.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
t_FEfa	0.282	0.795	0.105	0.225	-0.272	-0.359			0.125							
t_FEfb	0.880			-0.133	-0.178	-0.128			-0.378							
t_FEma	0.752	0.416	-0.295				-0.322			-0.233						
t_FEmb	0.714	0.362		0.469		0.252		-0.254								
t_FIfa	-0.211	0.850		0.137	0.262	0.140		0.345								
t_FIfb	0.277	0.587	0.667	-0.222	0.228			-0.161								
t_FIma	0.598	0.546		-0.284	-0.357	0.345										
t_FImb	0.448	0.602	-0.491	-0.233	0.192	-0.109	0.261	-0.145								
t_IEfa	0.954	-0.172											0.122		0.144	
t_IEfb	0.937	-0.199					-0.108					-0.142				
t_IEma	0.932						-0.133			0.255	-0.109					
t_IEmb	0.916	-0.226								0.105	0.268					
t_SEfa	0.864	-0.255	0.207	0.134			0.272			-0.159						
t_SEfb	0.938	-0.254														-0.116
t_SEma	0.940	-0.219										0.164			-0.103	
t_SEmb	0.951	-0.178											-0.150	0.132		

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
SSL	9.517	3.007	0.851	0.530	0.416	0.369	0.304	0.259	0.203	0.178	0.113	0.080	0.060	0.051	0.043	0.020
PVar	0.595	0.188	0.053	0.033	0.026	0.023	0.019	0.016	0.013	0.011	0.007	0.005	0.004	0.003	0.003	0.001
CVar	0.595	0.783	0.836	0.869	0.895	0.918	0.937	0.953	0.966	0.977	0.984	0.989	0.993	0.996	0.999	1.000

Table 4: Principal component analysis for the English vowel duration measures. SSL indicates SS loadings, PVar, Proportion Variance and CVar, Cumulative Variance.

The PCA for the French vowel spectra data is displayed in Table 5.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
vowel_EEfa	0.848	-0.234	0.126		0.271	0.120	0.191		-0.160		-0.140			-0.124	-0.101	
vowel_EEfb	0.936	-0.185					0.112	-0.124								-0.174
vowel_EEma	0.892		-0.268			-0.114	-0.169		-0.111			-0.138	0.155	-0.113		
vowel_EEmb	0.744	-0.398	0.261		-0.170	-0.241		0.119		0.309						
vowel_EIfa	0.754		0.220	-0.329	0.128	0.385		-0.239	-0.131		0.120					
vowel_EIfb	0.825	-0.185	0.298	-0.226		-0.145	-0.110		0.200			0.190		-0.104		
vowel_EIma	0.679	-0.454	0.158	0.323		0.175	-0.330				-0.136		-0.120			
vowel_EImb	0.419		0.783	-0.101	0.301	-0.104		0.177	0.181			-0.121				
vowel_QFfa	0.776		-0.530			-0.115				-0.110	0.175		-0.182			
vowel_QFfb	0.539	0.527	-0.443	-0.197		0.286			0.275	0.148						
vowel_QFma	0.144	-0.680	-0.311	0.446		0.326	0.161	0.238	0.104							
vowel_QFmb	0.371	0.545	0.394		-0.486	0.285		0.263								
vowel_SFfa	0.589	0.686	-0.106	-0.133	0.236			0.123		0.115	-0.154	0.126			0.118	
vowel_SFfb	0.193	0.460	0.254	0.699	-0.217		0.300	-0.216								
vowel_SFma	0.845		-0.333		-0.121	-0.260				-0.118	-0.110			0.131		
vowel_SFmb	0.321	0.499		0.660	0.308		-0.283				0.111					

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
SSL	7.083	2.476	1.854	1.474	0.688	0.657	0.421	0.334	0.267	0.189	0.149	0.115	0.096	0.077	0.072	0.048
PVar	0.443	0.155	0.116	0.092	0.043	0.041	0.026	0.021	0.017	0.012	0.009	0.007	0.006	0.005	0.005	0.003
CVar	0.443	0.597	0.713	0.805	0.848	0.890	0.916	0.937	0.953	0.965	0.974	0.982	0.988	0.993	0.997	1.000

Table 5: Principal component analysis for the French vowel spectra measures. SSL indicates SS loadings, PVar, Proportion Variance and CVar, Cumulative Variance.

Table 6 displays the French vowel duration data.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
t_EEfa		0.457	0.461	-0.295	0.665	-0.171										
t_EEfb	0.159	0.830		-0.324		0.111		0.186		0.212	-0.152			-0.107		
t_EEma	0.470	0.578	0.425		-0.314			0.194		-0.310						
t_EEmb	0.241	0.447	0.716		-0.241	-0.130		-0.348								
t_EIfa	0.796	0.388	-0.198			-0.132			0.113	0.206	0.198		0.184			
t_EIfb	0.837	0.403						0.132			0.126	0.108			0.168	-0.178
t_EIma	0.845	0.227		0.228		0.135		-0.122		0.130	-0.284	0.107			0.109	0.136
t_EImb	0.797	0.240		0.279	0.229	0.174	0.194			-0.111	0.166	0.172		-0.123	-0.121	
t_QFfa	-0.393	0.716	-0.274	-0.117		-0.230	0.231		0.304				-0.166			
t_QFfb	-0.429	0.716	-0.193	-0.324			-0.234		-0.133		0.151			-0.119	0.109	0.112
t_QFma		0.639	-0.483	0.292	0.162	0.347	-0.194	-0.148		-0.110		-0.198				
t_QFmb	-0.269	0.457		0.660		-0.476		0.163								
t_SFfa	-0.806	0.409		0.109					-0.276			0.111		0.214		
t_SFfb	-0.789	0.420	-0.197				0.181				-0.132	0.141	0.256			
t_SFma	-0.747	0.171	0.277	0.251		0.243	-0.290	0.135	0.244	0.136		0.161				
t_SFmb	-0.692	0.153	0.435	0.175		0.329	0.294	0.112			0.120	-0.171			0.103	

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
SSL	5.715	3.892	1.621	1.085	0.714	0.703	0.423	0.336	0.284	0.273	0.261	0.210	0.157	0.126	0.103	0.099
PVar	0.357	0.243	0.101	0.068	0.045	0.044	0.026	0.021	0.018	0.017	0.016	0.013	0.010	0.008	0.006	0.006
CVar	0.357	0.600	0.702	0.770	0.814	0.858	0.884	0.905	0.923	0.940	0.957	0.970	0.979	0.987	0.994	1.000

Table 6: Principal component analysis for the French vowel duration measures. SSL indicates SS loadings, PVar, Proportion Variance and CVar, Cumulative Variance.

Finally, the within language group correlations for these factors were generally not significant with only a moderately significant correlation for F listeners for PC1 across languages for the F listener group, $r = 0.367$, $p < .05$. This indicates that the F listeners who produced speech in a more native-like manner in a language reproduced the same pattern in the other. The correlation is difficult to interpret but could be reflective of individual aspects of speaking styles instead of indicators of native-like patterns of production. Therefore, the overall picture of the PCA indicates that there is no strong evidence of interaction between English and French in terms of production in a way that might have been predicted by the SLM. However, it is probably more broadly true that the PCA is less effective for the production than the perception analysis.

5.4. Discussion

Overall, the results demonstrated that the talker-listener interactions in both French and English were strongly driven by accent, L2 proficiency and acoustic similarity effects. There was strong evidence showing that L1 French listeners process accents in noise differently from L1 English listeners, having less of a strong advantage for Standard French and more sensitivity to the other accents. In addition, the L1 French listeners had higher intelligibility for high-proficiency L2 English accents over Quebecois French, an L1 French accent. This type of accent processing is different from the one displayed by the L1 English speakers in the present study and Chapter 2 where they showed a strong selective tuning for their own accent (Southern British English) and poor, undifferentiated performance for all other accents.

Why are there differences in the way L1 speakers process accents in French and English? Is there something particular about French? One hypothesis is that this is caused by the phonological differences between French and English. For example, English has more vowels than does French, and vowel differences are the major cause of accent differences within English. As suggested earlier, it may thus be that accent variation affects L1 French listeners differently because they process vowel variation differently, or that accent variation in French does not involve differences in vowels to the same extent as in English. Another explanation for this phonological variation is the way in which vowel duration contributes to speech intelligibility in French. Indeed, there is evidence from the production data that vowel duration brings a stronger contribution to the talker-listener accent interaction in French than vowel spectra. In English, however, both vowel spectra and duration contribute more evenly to accent intelligibility. In addition, the French listeners displayed high intelligibility for EE accented speech and the accent was acoustically close to the F listeners’, notably in terms of vowel duration. This indicates that English speakers acquiring French vowels may have less difficulty than French speakers learning English vowels, and so easily achieve a native-like command of French. Indeed, the fact that the EE-accented talkers were significantly more intelligible than the QF talkers was surprising considering that Quebecois French is a L1 accent. The opposite effect was not found for English, but instead, the English listeners found FE and IE-accented speech to be similarly intelligible, indicating that the FE talkers had not acquired the same level of intelligibility in English as the EE talkers in French. This suggests that there could be a discrepancy in L2 learning between English and French, with accent variation in French easier to process because of its vocalic characteristics.

Another hypothesis is that the differences in accent processing may not arise solely from language-dependent factors, but instead could be accounted for in terms of listener-dependent factors such as accent familiarity and language experience. For example, the majority of the French listeners were residing in London, and it is possible that they had much more exposure to English- over Quebecois-accented French. Perhaps the English exposure was great enough to make English accents nearly as intelligible as standard L1 French. In addition, it is possible that the French listeners' language experience (i.e., residency in the L2 speaking country) might have led them to become more attuned to accent differences, particularly English accents in French, despite the fact that they reported little familiarity with them. This strong sensitivity to accent differences was not replicated for English listeners, even though a large portion of the listeners had spent some time abroad in French speaking countries, but in the past. The English listeners did show some sensitivity to accents, but when listening in their L2, as did the French listeners when processing English speech. It is thus possible that the experience of living in the L2 speaking country affects listeners' sensitivity to accents, including in their L1 (even though some of the French listeners had been residing in the UK for only a limited time; e.g., four weeks). A broader investigation of language experience, notably length of residency, needs to be undertaken to examine these speech processes, including groups of French and British English listeners residing in France.

One of the major findings was the lack of evidence for inter-lingual interferences for L2 listeners as would have been predicted by the SLM (Flege, 1995, 2003). Indeed, according to the SLM, L1 dominant bilinguals and L2 speakers are not able to suppress the influence of their L1 on their L2 because their languages share a common phonological space and will unavoidably interfere. Instead, in the present study, the perception data showed strong facilitation effects for general speech processes in highly proficient L2 listeners. Indeed, strong correlations were shown for language ability and proficiency in English and French for all listener groups, indicating that listeners with a high degree of proficiency in their L2 benefitted from better speech recognition processes in both their languages. In addition, the data revealed no correlations with secondary accent effects, which is a strong indicator that overall language abilities are independent of talker-listener accent interactions. Therefore, the broad level of language interaction found in the present study differs from the inter-lingual interference described by Flege et al.'s work. The facilitation effects offered by a high degree of L2 proficiency resembles our previous research (Pinet and Iverson, 2010) where highly proficient L2 French listeners were able to take advantage of the presence of both their L1 and L2 cues to overcome degraded listening conditions, suggesting an effect of flexibility of processing rather than interference. However, it is fair to say that the cross-language interferences predicted by the SLM are mainly observed on micro scale investigations of L2 speakers' production of L1 speech within studies of phonetic contrasts (e.g., /eɪ/-/e/ production in Italian speakers of English, Flege et al., 2003). In this study, the interactions were, however, observed on a macro scale, within a broad study of accent and the cross-language production component analysis was rather inconclusive. It is possible, therefore, that inter-lingual interferences are better observed on a micro level scale of production.

One of the main goals of this study was to examine bilingual listeners' accent processing patterns in both their languages in order to observe whether they would display accent processing patterns typical of monolingual listeners in both French and English, or instead show a different pattern of accent processing (e.g., particular advantages for L2 accents of French and English, strong inter-lingual interferences, flexibility of processing). The results showed that, while the bilingual listeners' accent processing pattern was very similar to the English listeners' in their L1, paralleling Chapter 2's results, it differed from the French listeners' processing pattern in their L1.

Indeed, there were consistent differences between the French and bilingual listeners in terms of language ability in French and sensitivity to accents. They had lower overall recognition levels for French and showed less of a strong advantage for SF speech compared to the French listeners, with a reduced difference in intelligibility between SF and EE speech, as well as similar intelligibility levels for QF and EI speech. These differences thus point towards a language dominance effect in English for these listeners, even though every effort was made to recruit balanced bilinguals; the participants were raised speaking both languages and had a native-like command of the languages on interview. It is possible that, similar to the French listeners, their language experience may have affected their processing patterns in French and brought about their dominance in English as they were all residing in the UK at the time of testing and consequently were using English more often.

Furthermore, the bilinguals' language dominance was revealed through combined objective measures of speech perception and production compared to previous studies involving accent ratings. For instance, in Yeni-Komshian et al. (2000), the bilingual speakers' degree of accentedness was assessed by native listeners, and the authors note that it is possible that a fine-grained acoustic analysis of sentences produced by the English-dominant bilinguals would reveal subtle divergences from English phonetic norms that went undetected by the listeners who rated the sentences for foreign accent. The present study addressed this issue with measures of language proficiency, both in terms of perception (overall performance of accent recognition) and production (acoustic similarity to L1 speech), and thus offered a reliable assessment method of language dominance in bilingual listeners instead of relying on accent ratings. This assessment thus revealed more subtle patterns of production that could go unnoticed to the naïve listener.

To conclude, this investigation of cross-linguistic talker-listener accent interactions in noise has revealed several important findings. First, facilitation effects were observed in highly bilingual listeners, indicating that learning a L2 to high level of proficiency does not cause phonological interferences but instead, benefits the listeners in terms of general speech processing. Subsequently, it was suggested that communicating daily in a L2 might lead to higher sensitivity to accents both in the L1 and the L2 and provide the listeners with raised phonological awareness. The results also showed that strong measures of both perception and production, including acoustic analysis, are needed to detect effects of language dominance in from-birth bilinguals that would have been otherwise undetected by judgements of accentedness. Exploring the differences in accent processing is important not only because almost all of the published research on accent processing has focused on English, but also because little is known about how speech communication compares to other languages. The present study raised the importance of such investigations by showing that accent processing in French differs considerably from English. The investigation of accent variation and speech in noise in French is thus relevant to researchers who are specifically interested in French, but also gives a better understanding of general, language-independent, speech recognition processes.

6. Chapter six: Discussion

Overall, the findings that emerged from this research bring a substantial contribution to our understanding of how accent affects the intelligibility of speech in noise for normal-hearing L1 and L2 listeners. Several underlying factors were shown to affect the talker-listener interaction with different magnitudes. L2 experience was shown to play a major role in the interaction, with both the talkers' and the listeners' degree of L2 proficiency having strong effects on speech intelligibility, modulating accent processing patterns and general speech in noise processes, with language processing facilitation effects shown for highly bilingual speakers. Acoustic similarity in the accents of the talkers and the listeners was shown to account for a substantial amount of the variance in accent intelligibility. Not only was this relationship shown for speech processing patterns in English, but the metric was also extended to account for accent processing in French, with vowel duration being a strong indicator of proficiency and acoustic similarity. The cross-language research work revealed how different patterns of accent processing varies across languages, with strong selective tuning processes for standard accents displayed by L1 English listeners and more sensitivity to accent differences shown by L1 French listeners. Finally, adaptation to a novel accent in British English listeners was shown to be small in magnitude, even with the added presence of social interaction, and there was a persistent selective tuning for the listeners' own accent.

One of the major findings that emerged from this research is the evidence in support of a perception-production link, reflected in the relationship between speech recognition in noise and acoustic similarity in the accents of the talkers and the listeners (e.g., motor theory; Liberman and Mattingly, 1985). These interactions were shown for both L1 and L2 speakers across languages, even though the magnitude of the effect was relatively small in French. For L2 listeners, the strength of a perception-production link was very much modulated by their language experience. For instance, facilitation effects were much stronger in highly bilingual listeners. The link was evidenced by strong similarity in terms of vowel production, but it must be noted that this effect was observed on a macro scale investigation of the speakers' phonetic details of accent production. That is, the production analysis involved vowel spectra and duration measurements, and was applied to general measures of accent, as they represent good indicators of accent differences. However, it is possible that because the measures were conducted on the entire vowel space, other factors may have contributed to the interaction such as overall language proficiency and other speech processes, thus reinforcing the interaction. Conducting measures on a micro scale (e.g., a single phonetic contrast) may not have revealed this perception-production link. Further investigation is needed to understand the underlying nature of the perception-production link in this particular accent interaction.

A significant part of this research work has been focused on the implementation of ACCDIST in order to investigate a broad range of L1 and L2 accent interactions in both English and French, and the research has brought a contribution to the literature in terms of assessment methods of language experience (L2 proficiency, language dominance in bilinguals) and accent similarity. Such methods constitute robust and objective measures to evaluate the various factors involved in the interaction as opposed to techniques relying on listeners' judgement of degree of accentedness (e.g., accent classification methods, Clopper, 2008). In particular, computational techniques based on acoustic measurements of speech have been shown to be a useful and clear indicator of

L2 proficiency across languages. The ACCDIST metric is also a useful tool for sociophonetic research, for instance, by establishing the acoustic contribution to making standard accents in Britain intelligible to speakers of other L1 accents. Most importantly, implementing the metric to a broader range of acoustic measures (consonantal, supra-segmental) would enable us to build a more detailed picture of the talker-listener interactions involved in a perception-production link.

The fact that the work is interdisciplinary implies that it has relevance to various areas of research. Communicating in a noisy environment is comparable to mild hearing loss (e.g., ~6dB, Dubno et al., 1984) and L2 listeners' speech in noise recognition has also been compared to that of hearing-impaired listeners (e.g., 15dB raised reception threshold, Rogers et al., 2006). The findings here thus have direct implications for the hearing-impairment research by broadly contributing to our understanding of how hearing-impaired listeners function in a linguistically diverse society (e.g., when communicating with L2 English speakers, or when they have learned English as an L2). The work is also relevant to L2 learning and development issues of bilingualism. For instance, by revealing the benefits of an acoustic match in accent to maximise speech communication in noise, the work is relevant to understanding how L1, L2 and bilingual children cope with classroom noise and accented teachers/carers and how this may affect their learning. Furthermore, bilingual children have been shown to be more affected by noise than monolingual children (e.g., reverberant noise in classrooms, Rogers et al., 2006), which differs from the facilitation effects observed for the bilingual adults in the present work. This suggests that bilingual listeners may only gain the benefits of being highly proficient in two languages for general speech in noise processing later in life. The facilitation effects observed in highly bilingual speakers also indicated that learning a L2 offers the speakers with better communication abilities, which are relevant to language teaching. Finally, the work also has direct implications for forensic research, notably in terms of the processing of accented speech in distorted signals where the phonetic information can be reduced and masked by the distortion (e.g., babble noise, street noise). Implementing strong acoustic measures of accent would enable a prompt recognition of the accented speech for talker identification purposes.

The outcomes of the work presented here call for further investigation. There are two main lines of research that need immediate investigation. First, the work on accent processing described in Chapter 5 could be extended to a broader examination of accent interactions in French by carrying out a thorough investigation of the various factors likely to affect the L1-L2 talker-listener accent interaction in French (i.e., accent familiarity and exposure, acoustic similarity and L2 experience). Monolingual French listeners residing in France (in two geographical areas) and the UK will be tested on their recognition of L1 and L2 accents in noise that vary in acoustic similarity to French, and accent familiarity. Phonological differences will be assessed by using the same acoustic similarity metric described in Chapters 3 and 5 (ACCDIST; Huckvale, 2004, 2007a, b). Accent exposure will be assessed by determining whether ratings of accent familiarity and experience better account for the variation in speech in noise performance. This would not only reveal more characteristic accent processing patterns that would be free of continuous L2 use (i.e., monolingual French listeners tested in France and not in the UK), but would also substantially contribute to our general understanding of speech in noise perception. Furthermore, this would have direct implications for models of speech perception and general human language processing.

Accent effects on the recognition of speech in noise

The next line of research would be focused on investigating the effects of accented speech on individuals with hearing loss. Indeed, even though individuals with hearing impairments are expected to be affected by accent, it has not been established exactly how and to what extent this occurs. For example, hearing-impaired individuals may demonstrate the same highly selective tuning for L1 accents described in Chapter 1 and 4 for normal-hearing L1 English speakers, or their lower-levels of speech recognition performance may produce more graded patterns of performance such as those found here for inexperienced L2 speakers. This issue would be investigated by testing L1 speakers with moderate, flat hearing losses, in quiet and noise using an accent recognition task, in order to assess how their speech recognition abilities are affected by accent and subsequently examine their abilities to tune into L2 accents. L1 and L2 normal hearing speakers of English will also be tested as control groups. In addition, accent processing has been shown to become more difficult for older listeners (e.g., Adank and Janse, 2010) and this work could be extended to test the effects of age-related accent processing in older listeners, in order to get a broader knowledge of how accent processing is affected by different types of hearing losses and age-related factors. For this, older listeners with and without age-related hearing loss would be tested on their accent recognition abilities. This work will therefore give more insight into the different listening strategies used by both normal and hearing-impaired listeners, and how this can be compared to the listening strategies used by inexperienced L2 speakers.

7. References

Adank, P., Evans, B., G., Stuart-Smith, J. and Scott, S., K. (2009). “Comprehension of familiar and unfamiliar native accents under adverse listening conditions”, Journal of Experimental Psychology: Human Perception and Performance 35, 520-529.

Adank, P. and Janse, E. (2010). “Comprehension of a novel accent by young and older listeners”, Psychology and Ageing 25(3), 736-740.

Baker, R. and Hazan, V. (2011). “DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs”, Behaviour Research Methods 43(3), 761-770.

Bench, J., Kowal, A. and Bamford, J. (1979). “The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children”, British Journal of Audiology 13, 108-112.

Bent, T., and Bradlow, A. R. (2003). “The interlanguage speech intelligibility benefit”, Journal of the Acoustical Society of America 114, 1600-1610.

Best, C. T., McRoberts, G. W., Lafleur, R. and Silver-Isenstadt, J. (1995). “Divergent developemental patterns for infants perception of 2 nonnative consonant contrasts”, Infant Behavior and Development 18 (3), 339 –350.

Bradlow, A. R. and Pisoni, D. B. (1999). “Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors”, Journal of the Acoustical Society of America 106 (4), 2074-2085.

Bradlow, A. R., Kraus, N. and Hayes, E. (2003). “Speaking clearly for children with learning disabilities: Sentence perception in noise”, Journal of Speech, Language and Hearing Reserach 46, 80-97.

Bradlow, A.R. and Bent, T. (2007). “Perceptual adaptation to non-native speech”, Cognition 106, 707-729.

Cincarek, T., Gruhn, R., Hacker, C., Nöth, E. and Nakamura, S. (2009). “Automatic pronunciation scoring of words and sentences independent from the non-native's first language”, Computer Speech & Language 23 (1), 65-88.

Clarke, C. M. (2000). “Perceptual adjustment to foreign-accented English”, Journal of the Acoustical Society of America 107, 2856(A).

Clarke, C. M. and Garrett, M. F. (2004). “Rapid adaptation to foreign-accented English”, Journal of the Acoustical Society of America 116, 3647-3658.

Clopper, C.G., Levi, S.V. and Pisoni, D.B. (2005a). “Perceptual similarity of regional dialects of American English”, Journal of the Acoustical Society of America 119, 566-574.

Clopper, C.G., Conrey, B. and Pisoni, D.B. (2005b). “Effects of talker gender on dialect categorization”, Journal of language and social psychology 24(2), 182-206.

Clopper, C.G. and Bradlow, A.R. (2007). “Native and non-native perceptual dialect similarity spaces”, Proceedings of ICPhS XVI, Saarbrucken, 6-10 August 2007, 665-668.

Clopper, C.G. and Bradlow, A.R. (2008). “Perception of dialect variation in noise: Intelligibility and classification”, Language and Speech 51, 175-198.

Clopper, C.G. (2008). “Auditory free classification: Methods and analysis”, Behaviour Research Methods 40, 575-581.

Cucchiarini, C., Strik, H. and Boves, L. (2000). “Different aspects of expert pronunciation quality ratings and their relation to scores produced by speech recognition algorithms”, Speech Communication 30 (2-3), 109-119.

Cutler, A., Mehler, J., Norris, D. and Segui, J. (1989). “Limits on bilingualism”, Nature (London) 340, 229-230.

Cutler, A., Mehler, J., Norris, D. and Segui, J. (1992). “The monolingual nature of speech segmentation by bilinguals”, Cognitive Psychology 24, 381-410.

Davis, S.B. and Mermelstein, P. (1980). “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences”, IEEE Transactions on Acoustics, Speech and Signal 28, 357-366.

Davis, M. H., Johnsrude, I. S., Hervais-Ademan, A., Taylor, K. and McGettigan, C. (2005). “Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise vocoded sentences”, Journal of Experimental Psychology: General 134(2), 222-241.

Dubno, J.R., Dirks, D.D. and Morgan, D.E. (1984). “Effects of age and mild hearing-loss on speech recognition in noise”, Journal of the Acoustical Society of America 76(1), 87-96

Eisner, F. and McQueen, J. M. (2006). “Perceptual learning in speech: Stability overtime”, Journal of the Acoustical Society of America, 119(4), 1950–1953.

Evans, B. G. and Iverson, P. (2004). “Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences,” Journal of the Acoustical Society of America 115, 352-361.

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

Evans, B.G. and Taylor, E. (2010). “Investigating the time course of perceptual adaptation to unfamiliar accented speech”, British Associations of Academic Phoneticians colloquium, London, UK (03/10)

Flege, J. E. (1995). “Second Language speech learning: Theory, findings and problems”, in Speech Perception and Linguistic Experience: Issues in Cross-Language Research, edited by W. Strange, York Press, Baltimore MD, 233-277.

Flege, J., Munro, M. and MacKay, I. (1995). “The effect of age of second language learning on the production of English consonants”, Speech Communication 16, 1-26.

Flege, J. E., MacKay, I. R. A. and Meador, D. (1999). “Native Italian speakers’ perception and production of English vowels”, Journal of the Acoustical Society of America 106, 2973-2987.

Flege, J., MacKay, I. and Piske, T. (2002). “Assessing bilingual dominance”, Applied Psycholinguistics 23, 567-598.

Flege, J. E. (2003). “Assessing constraints on second-language segmental production and perception”, in Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities, edited by A. Meyer and N. Schiller, Mouton de Gruyter, Berlin, Germany.

Flege, J., Schirru, C. and MacKay, I. (2003). “Interaction between the native and second language phonetic subsystems”, Speech Communication 40, 467-491.

Floccia, C., Goslin, J., Girard, F. and Konopczynski, G. (2006). “Does a regional accent perturb speech processing?”, Journal of Experimental Psychology: Human Perception and Performance 32, 1276–1293.

Floccia, C., Butler, J., Goslin, J. and Ellis, L. (2009). “Regional and foreign accent processing in English: Can listeners adapt?”, Journal of Psycholinguistic Research 38, 378-412.

Florentine, M., Buus, S., Scharf, B. and Canevet, G. (1984). “Speech reception thresholds in noise for native and non-native listeners (A)”, Journal of the Acoustical Society of America 75, S84.

Franco, H., Abrash, V., Precoda, K., Bratt, H., Rao, R. and Butzberger, J. (2000). “The SRI EduSpeak System: Recognition and Pronunciation Scoring for Language Learning”, Proceedings of InSTIL 2000 (Integrating Speech Technology in Language Learning), Dundee, Scotland.

Gass, S. and Varonis, E. (1984). ‘‘The effect of familiarity on the comprehensibility of nonnative speech’’, Language Learning 34, 65-89.

Gordon-Salant, S., Yeni-Komshian, G. H. and Fitzgibbons, P. J. (2010). “Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss”, Journal of the Acoustical Society of America 128 (1), 444-455

Hayes-Harb, R., Smith, B. L., Bent, T., and Bradlow, A. R. (2008). “The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts,” Journal of Phonetics 36, 664-679.

Hazan, V. and Markam, D. (2004). “Acoustic-phonetic correlates of talker intelligibility for adults and children”, Journal of the Acoustical Society of America 116, 3108-3118.

Hazan, V., Sennema, A., Iba, M. and Faulkner, A. (2005). “Effect of audio-visual perceptual perception and production of training on the consonants by Japanese learners of English”, Speech Communication 47(3), 360-378

Hazan, V., Kim, J. and Chen, Y. (2010). “Audio-visual perception in adverse conditions: language, speaker and listener effects”, Speech Communication 52, 996-1009

Hazan, V. and Baker, R. (2011). “Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions”, Journal of the Acoustical Society of America 130(4), 2139-2152

HTK Hidden Markov Modelling toolkit (1989). http://htk.eng.cam.ac.uk/

Huckvale, M. (2004). “ACCDIST: a metric for comparing speakers' accents”, ICSLP 2004, Kim, S.H., Young, D.H. (ed.) Jeju, Korea, 29-32.

Huckvale, M. (2007a). “Hierarchical clustering of speakers into accents with the ACCDIST metric”, International Congress of Phonetic Sciences, Saarbrücken, Germany.

Huckvale, M. (2007b). “ACCDIST: an accent similarity metric for accent recognition and diagnosis”, in Müller,C (ed.) Speaker Classification II. Lecture Notes in Computer Science series. Series edited by Carbonell,J., Siekmann,J.. Berlin: Springer, 4441st edition, 258-275.

Imai, S., Flege, J.E. and Walley, A. (2003). “Spoken word recognition of accented and unaccented speech: lexical factors affecting native and non-native listeners”, Proceedings of the International Congress on Phonetic Science, Barcelona, Spain.

Imai, S., Walley, A. and Flege, J. (2005). “Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented by native English and Spanish listeners”, Journal of the Acoustical Society of America 117, 896-907.

Iverson, P. and Pinet, M. (in preparation). “Individual differences in talker intelligibility in noise as a function of talker-listener similarity”.

Kuhl, P. K., Tsao. F. M. and Liu, H. M. (2003). "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning", Proceedings of the National Academy of Sciences 100, 9096-9101.

Kruskal, J.B. (1964). "Nonmetric multidimensional scaling: a numerical method", Psychometrika 29, 115-129.

Lane, H. (1963). "Foreign Accent and Speech Distortion", Journal of the Acoustical Society of America 35, 451-453.

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

Markam, D. and Hazan, V. (2002). "Speaker intelligibility of adults and children", Proceedings of ICSLP, 2002.

Mattys, S. L., White, L. and Melhorn, J. F. (2005). "Integration of multiple speech segmentation cues: A hierarchical framework", Journal of Experimental Psychology. General 134, 477-500.

Mayo, L. H., Florentine, M. and Buus, S. (1997). "Age of second-language acquisition and perception of speech in noise", Journal of Speech, Language and Hearing Research 40, 686-693.

Meador, D., Flege, J. and MacKay, I. (2000). "Factors affecting the recognition of words in a second language", Bilingualism: Language and Cognition 3, 55–67.

Munro, M. (1998). "The effects of noise on the intelligibility of foreign-accented speech", Studies in Second Language Acquisition 20, 139-154.

Naigles, L. R. and Mayeux, L. (2001). "Television as incidental language teacher" in Handbook of Children and the Media, Singer, D. G. and Singer, J. L. (eds), Sage, Thousand Oaks, CA, pp. 135–152.

Neumeyer, L., Franco, H., Digalakis, V. and Weintraub, M. (2000). "Automatic Scoring of Pronunciation Quality", Speech Communication 30, 83-93.

Oliver, G. and Iverson, P. (2010). "Individual variation in the production and perception of second-language phonemes: French speakers learning English /i -I/", Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010, Poznań, Poland.

Pinet, M. and Iverson, P. (2010). "Talker-listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience", Journal of the Acoustical Society of America 128, 1357-1365.

Pisoni, D., Nusbaum, H. and Greene, B. (1985). "Perception of synthetic speech generated by rule". Proceedings of the IEEE 73, 1665-1676.

Plomp, R. (1978). "Auditory handicap of hearing impairment and limited benefit of hearing aids", Journal of the Acoustical Society of America 63 (2), 533-549

Rogers, C. L., Dalby, J. and Nishi, K. (2004). "Effects of noise and proficiency level on intelligibility of Chinese-accented English", Language and Speech 47, 139-154

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M. and Abrams, H. B. (2006). "Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing", Applied Psycholinguistics 27, 465-485.

Selinker, L. (1972). "Interlanguage", Iral-International Review of Applied Linguistics in Language Teaching 10 (3), 209-231.

Smiljanic, R. and Bradlow, A.R. (2007). "Clear speech intelligibility: Listener and talker effects", Proceedings of ICPhS XVI, Saarbrucken, 6-10 August 2007.

Smiljanic, R. and Bradlow, A.R. (2008). "Speaking and hearing clearly: Talkers and listeners factors in speaking style changes", Language and Linguistics Compass 3/1, 236-264.

Stibbard, R. M. and Lee, J. I. (2006). “Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis”, Journal of the Acoustical Society of America 120, 433-442.

Stuart-Smith, J. (2007). “The influence of the media”, in: Llamas, C., Mullany, L. and Stockwell, P. (eds.) The Routledge Companion to Sociolinguistics. Routledge, London, UK, pp. 140-148

Thompson, M. and Hazan, V. (2010). “The impact of visual cues and lexical knowledge on the perception of a non-native consonant contrast for Colombian adults”, Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010, Poznań, Poland, 493-498.

van Wijngaarden, S. J. (2001). “Intelligibility of native and non-native Dutch speech”, Speech Communication 35, 103-113.

van Wijngaarden, S. J., Steeneken, H. J. M. and Houtgast, T. (2002). “Quantifying the intelligibility of speech in noise for non-native listeners”, Journal of the Acoustical Society of America 111, 1906-1916.

Weil, S. A. (2001). “Foreign-accented speech: Encoding and generalization”, Journal of the Acoustical Society of America 109, 2473 (A).

Yeni-Komshian, G. H., Flege, J. E. and Liu, S. (2000) “Pronunciation proficiency in the first and second languages of Korean-English bilinguals”, Bilingualism: Language and Cognition 3, 131-149.

8. Appendix

8.1. Appendix 1: BKB sentences, English

The capitalized words are the keywords on which the listeners were scored. The first 31 sentences (in bold) were used for the production tasks and acoustic analyses in Chapters 3 and 5.

The CLOWN had a FUNNY FACE
The CAR ENGINE'S RUNNING
SHE CUT with her KNIFE
CHILDREN LIKE STRAWBERRIES
The HOUSE had NINE ROOMS
THEY'RE BUYING some BREAD
The GREEN TOMATOES are SMALL
HE PLAYED with his TRAIN
The POSTMAN SHUT the GATE
THEY'RE LOOKING AT the CLOCK
The BAG BUMPS on the GROUND
The BOY DID a HANDSTAND
A CAT SITS ON the BED
The LORRY CARRIED FRUIT
The RAIN CAME DOWN
The ICE CREAM was PINK
The LADDER'S NEAR the DOOR
THEY had a LOVELY DAY
The BALL WENT INTO the GOAL
The OLD GLOVES are DIRTY
HE CUT his FINGER
The THIN DOG was HUNGRY
The BOY KNEW the GAME
SNOW FALLS at CHRISTMAS
SHE'S TAKING her COAT
The POLICE CHASED the CAR
A MOUSE RAN DOWN the HOLE
The LADY'S MAKING a TOY
Some STICKS were UNDER the TREE
The LITTLE BABY SLEEPS
THEY'RE WATCHING the TRAIN
The SCHOOL FINISHED EARLY
The GLASS BOWL BROKE
The DOG PLAYED with a STICK
The KETTLE'S QUITE HOT

The FARMER KEEPS a BULL
THEY SAY some SILLY THINGS
The LADY WORE a COAT
The CHILDREN are WALKING HOME
HE NEEDED his HOLIDAY
The MILK CAME in a BOTTLE
The MAN CLEANED his SHOES
THEY ATE the LEMON JELLY
The BOY'S RUNNING AWAY
FATHER LOOKED at the BOOK
SHE DRINKS from her CUP
The ROOM'S GETTING COLD
A GIRL KICKED the TABLE
The WIFE HELPED her HUSBAND
The MACHINE was QUITE NOISY
The OLD MAN WORRIES
A BOY RAN down the PATH
The HOUSE had a NICE GARDEN
SHE SPOKE TO her SON
THEY'RE CROSSING the STREET
LEMONS GROW on TREES
HE FOUND his BROTHER
Some ANIMALS SLEEP ON STRAW
The JAM JAR was FULL
THEY'RE KNEELING DOWN
The GIRL LOST her DOLL
The COOK'S MAKING a CAKE
The CHILD GRABS the TOY
The MUD STUCK on his SHOE
The BATH TOWEL was WET
The MATCHES LIE on the SHELF
THEY'RE RUNNING PAST the HOUSE
The TRAIN had a BAD CRASH
The KITCHEN SINK'S EMPTY
A BOY FELL from the WINDOW
SHE USED her SPOON
The PARK'S NEAR the ROAD
The COOK CUT some ONIONS
The DOG MADE an ANGRY NOISE
HE'S WASHING his FACE
SOMEBODY TOOK the MONEY
The LIGHT WENT OUT
THEY WANTED some POTATOES
The NAUGHTY GIRL'S SHOUTING
The COLD MILK'S in a JUG
The PAINT DRIPPED on the GROUND
The MOTHER STIRS the TEA
THEY LAUGHED at his STORY

MEN WEAR LONG TROUSERS
The SMALL BOY was ASLEEP
The LADY GOES TO the SHOP
The SUN MELTED the SNOW
The FATHER'S COMING HOME
SHE had her POCKET MONEY
The LORRY DROVE up the ROAD
HE'S BRINGING his RAINCOAT
A SHARP KNIFE'S DANGEROUS
THEY TOOK some FOOD
The CLEVER GIRLS are READING
The BROOM STOOD in the CORNER
The WOMAN TIDIED her HOUSE
The CHILDREN DROPPED the BAG
The DOG CAME BACK
The FLOOR LOOKED CLEAN
SHE FOUND her PURSE
The FRUIT LIES on the GROUND
MOTHER FETCHES a SAUCEPAN
THEY WASHED in COLD WATER
The YOUNG PEOPLE are DANCING
The BUS WENT EARLY
THEY had TWO EMPTY BOTTLES
A BALL'S BOUNCING ALONG
The FATHER FORGOT the BREAD
The GIRL has a PICTURE BOOK
The ORANGE was QUITE SWEET
HE'S HOLDING his NOSE
The NEW ROAD'S on the MAP
The BOY FORGOT his BOOK
A FRIEND CAME for LUNCH
The MATCH BOXES are EMPTY
HE CLIMBED his LADDER
The FAMILY BOUGHT a HOUSE
The JUG STOOD on the SHELF
The BALL BROKE the WINDOW
THEY'RE SHOPPING for CHEESE
The POND WATER'S DIRTY
THEY HEARD a FUNNY NOISE
POLICE are CLEARING the ROAD
The BUS STOPPED SUDDENLY
SHE WRITES to her BROTHER
The FOOTBALLER LOST a BOOT
The THREE GIRLS are LISTENING
The COAT LIES ON a CHAIR
The BOOK TELLS a STORY
The YOUNG BOY LEFT HOME
THEY'RE CLIMBING the TREE

SHE STOOD near her WINDOW
The TABLE has THREE LEGS
A LETTER FELL on the MAT
The FIVE MEN are WORKING
HE LISTENS TO his FATHER
The SHOES were VERY DIRTY
THEY WENT on HOLIDAY
BABY BROKE his MUG
The LADY PACKED her BAG
The DINNER PLATE'S HOT
The TRAIN'S MOVING FAST
The CHILD DRANK some MILK
The CAR HIT a WALL
A TEA TOWEL'S by the SINK
The CLEANER USED a BROOM
SHE LOOKED IN her MIRROR
The GOOD BOY'S HELPING
THEY FOLLOWED the PATH
The KITCHEN CLOCK was WRONG
The DOG JUMPED ON the CHAIR
SOMEONE'S CROSSING the ROAD
The POSTMAN BRINGS a LETTER
THEY'RE CYCLING ALONG
HE BROKE his LEG
The MILK was by the FRONT DOOR
The SHIRTS HANG in the CUPBOARD
The GROUND was TOO HARD
The BUCKETS HOLD WATER
The CHICKEN LAID some EGGS
The SWEET SHOP was EMPTY
The DOGS GO for a WALK
SHE'S WASHING her DRESS
The LADY STAYED for TEA
The DRIVER WAITS by the CORNER
THEY FINISHED the DINNER
The POLICEMAN KNOWS the WAY
The LITTLE GIRL was HAPPY
HE WORE his YELLOW SHIRT
THEY'RE COMING for CHRISTMAS
The COW GAVE some MILK
The BOY GOT INTO BED
The TWO FARMERS are TALKING
MOTHER PICKED some FLOWERS
A FISH LAY on the PLATE
The FATHER WRITES a LETTER
The FOOD COST a LOT
The GIRL'S WASHING her HAIR
The FRONT GARDEN was PRETTY

HE LOST his HAT
The TAPS are ABOVE the SINK
FATHER PAID AT the GATE
SHE'S WAITING for her BUS
The BREAD VAN'S COMING
THEY had some COLD MEAT
The FOOTBALL GAME'S OVER
THEY CARRY some SHOPPING BAGS
The CHILDREN HELP the MILKMAN
The PICTURE CAME from a BOOK
The RICE PUDDING was READY
The BOY had a TOY DRAGON
A TREE FELL on the HOUSE
The FRUIT CAME in a BOX
The HUSBAND BRINGS some FLOWERS
THEY'RE PLAYING in the PARK
SHE ARGUED with her SISTER
A MAN TOLD the POLICE
POTATOES GROW in the GROUND
HE'S CLEANING his CAR
The MOUSE FOUND the CHEESE
THEY WAITED for ONE HOUR
The BIG DOG was DANGEROUS
The STRAWBERRY JAM was SWEET
The PLANT HANGS ABOVE the DOOR
The CHILDREN are ALL EATING
The BOY has BLACK HAIR
The MOTHER HEARD her BABY
The LORRY CLIMBED the HILL
The ANGRY MAN SHOUTED
The DOG SLEEPS in a BASKET
THEY'RE DRINKING TEA
MOTHER OPENS the DRAWER
An OLD WOMAN was at HOME
HE DROPPED his MONEY
THEY BROKE ALL the EGGS
The KITCHEN WINDOW was CLEAN
The GIRL PLAYS with the BABY
The BIG FISH GOT AWAY
SHE'S HELPING her FRIEND
The CHILDREN WASHED the PLATES
The POSTMAN COMES EARLY
The SIGN SHOWED the WAY
The GRASS is GETTING LONG
The MATCH FELL on the FLOOR
A MAN'S TURNING the TAP
The FIRE was VERY HOT
HE'S SUCKING his THUMB

The SHOP CLOSED for LUNCH
The DRIVER STARTS the ENGINE
The BOY HURRIED to SCHOOL
Some NICE PEOPLE are COMING
SHE BUMPED her HEAD
THEY MET SOME FRIENDS
FLOWERS GROW in the GARDEN
The TINY BABY was PRETTY
The DAUGHTER LAID the TABLE
THEY WALKED ACROSS the GRASS
The MOTHER TIED the STRING
The TRAIN STOPS at the STATION
The PUPPY PLAYS with a BALL
The CHILDREN WAVE at the TRAIN
MOTHER CUT the CHRISTMAS CAKE
HE CLOSED his EYES
The RAINCOAT'S VERY WET
A LADY BUYS some BUTTER
THEY CALLED an AMBULANCE
SHE'S PAYING for her BREAD
The POLICEMAN FOUND a DOG
Some MEN SHAVE in the MORNING
The DRIVER LOST his WAY
THEY STARED at the PICTURE
The CAT DRANK from a SAUCER
The OVEN DOOR was OPEN
The CAR'S GOING TOO FAST
The SILLY BOY'S HIDING
The PAINTER USED a BRUSH
The APPLE PIE'S COOKING
HE DRINKS from his MUG
The SKY was VERY BLUE
THEY KNOCKED on the WINDOW
The BIG BOY KICKED the BALL
PEOPLE are GOING HOME
The BABY WANTS his BOTTLE
The LADY SAT on her CHAIR
THEY had some JAM PUDDING
The SCISSORS are QUITE SHARP
SHE'S CALLING her DAUGHTER
Some BROWN LEAVES FELL off the TREE
The MILKMAN CARRIED the CREAM
A GIRL RAN ALONG
The MOTHER READS a PAPER
The DOG CHASED the CAT
The CAKE SHOP'S OPENING
THEY LIKE ORANGE MARMALADE
The MOTHER SHUT the WINDOW

HE'S SKATING WITH his FRIEND
The CHEESE PIE was GOOD
RAIN FALLS from CLOUDS
SHE TALKED to her DOLL
THEY PAINTED the WALL
The TOWEL DROPPED on the FLOOR
The DOG'S EATING some MEAT
A BOY BROKE the FENCE
The YELLOW PEARS were LOVELY
The POLICE HELP the DRIVER
The SNOW LAY on the ROOF
The LADY WASHED the SHIRT
The CUP HANGS on a HOOK
The FAMILY LIKE FISH
SUGAR'S VERY SWEET
The BABY LAY on a RUG
The WASHING MACHINE BROKE
THEY'RE CLEARING the TABLE
The CLEANER SWEPT the FLOOR
A GROCER SELLS BUTTER
The BATH WATER was WARM
HE'S REACHING for his SPOON
SHE HURT her HAND
The MILKMAN DRIVES a SMALL VAN
The BOY SLIPPED ON the STAIRS
THEY'RE STAYING for SUPPER
The GIRL HELD a MIRROR
The CUP STOOD on a SAUCER
The COWS WENT to MARKET
The BOY GOT into TROUBLE
THEY'RE GOING OUT
The FOOTBALL HIT the GOALPOST
HE PAID his BILL
The TEACLOTH'S QUITE WET
A CAT JUMPED OFF the FENCE
The BABY has BLUE EYES
THEY SAT on a WOODEN BENCH
MOTHER MADE some CURTAINS
The OVEN'S TOO HOT
The GIRL CAUGHT a COLD
The RAINCOAT'S HANGING UP
SHE BRUSHED her HAIR
The TWO CHILDREN are LAUGHING
The MAN TIED his SCARF
The FLOWER STANDS in a POT
The PEPPER POT was EMPTY
The DOG DRANK from a BOWL
A GIRL CAME into the ROOM

THEY'RE PUSHING an OLD CAR
The CAT CAUGHT a MOUSE
The ROAD GOES UP a HILL
SHE MADE her BED
BANANAS are YELLOW FRUIT
The COW LIES on the GRASS
The EGG CUPS are on the TABLE
HE FRIGHTENED his SISTER
The CRICKET TEAM'S PLAYING
The FATHER PICKED some PEARS
The KETTLE BOILED QUICKLY
The MAN'S PAINTING a SIGN
THEY LOST some MONEY

8.2. Appendix 2: BKB sentences, French

The capitalized words are the keywords on which the listeners were scored. The first 31 sentences (in bold) were used for the production tasks and acoustic analysis in Chapter 5. The translations were made from the English version paying particularly attention to overall sentence length and keyword numbers to keep the two versions as balanced as possible.

Le CLOWN avait un VISAGE RIGOLO
Le MOTEUR de la VOITURE TOURNE
Elle COUPE avec SON COUTEAU
Les ENFANTS AIMENT les FRAISES
La MAISON avait NEUF PIECES
Ils ACHÈTENT DU PAIN
Les TOMATES VERTES sont PETITES
Il JOUAIT avec SON TRAIN
Le FACTEUR FERME la PORTE
Ils REGARDENT L'HORLOGE de la CUISINE
Le SAC REBONDIT sur le SOL
Le GARÇON FAIT le POIRIER
Un CHAT est ASSIS SUR le LIT
Le CAMION TRANSPORTAIT des FRUTS
La PLUIE TOMBAIT du CIEL
La GLACE ÉTAIT ROSE
L'ÉCHELLE est PRÈS de la PORTE
Ils ont PASSÉ une BONNE JOURNÉE
Le BALON est ALLÉ DANS le BUT
Les VIEUX GANTS sont SALES
Il a COUPÉ SON DOIGT
Le CHIEN MAIGRE a FAIM

Le GARÇON CONNAISSAIT le JEU
La NEIGE TOMBE à NOËL
Elle PREND SON MANTEAU
La POLICE POURSUIVIT la VOITURE
La SOURIS se CACHE DANS le TROU
La DAME FABRIQUE un JOUET
Des BRANCHES ÉTAIENT sous l'ARBRE
Le PETIT BÉBÉ PLEURE
Ils REGARDENT le TRAIN PASSER
L'ÉCOLE FINIT plus TÔT
Le VASE se CASSA en DEUX
Le CHIEN JOUAIT avec un BÂTON
La BOUILLOIRE est ASSEZ CHAUDE
Le FERMIER a un TAUREAU NOIR
Ils DISENT des CHOSES BÊTES
La FEMME PORTAIT un MANTEAU
Les ENFANTS RENTRENT chez eux en MARCHANT
Il AVAIT BESOIN de VACANCES
Le LAIT ÉTAIT dans une BOUTEILLE
L'HOMME NETTOYAIT ses CHAUSSURES
Ils ONT MANGÉ de la GELÉE de CITRON
Le JEUNE GARÇON S'ENFUYAIT
Le PÈRE LIT le LIVRE
Elle BOIT DANS sa TASSE
La PIÈCE se REFROIDIT VITE
Une FILLE a TAPÉ dans la TABLE
La FEMME AIDE son MARI
La MACHINE était ASSEZ BRUYANTE
Le VIEIL HOMME S'INQUIÈTE
Un GARÇON COURAIT dans le CHEMIN
La MAISON avait un BEAU JARDIN
Elle PARLAIT AVEC SON FILS
Ils TRAVERSENT LA RUE
Les CITRONS POUSSENT sur des ARBRES
Il RETROUVE SON FRÈRE
CERTAINS ANIMAUX DORMENT sur la PAILLE
Le POT de CONFITURE était PLEIN
ILS se METTENT à GENOUX
La FILLE a PERDU sa POUPÉE
Le PATISSIER PRÉPARE un GÂTEAU
L'ENFANT PREND le JOUET
La BOUE COLLAIT sous sa CHAUSSURE
La SERVIETTE de BAIN était MOUILLÉE
Les ALLUMETTES sont POSÉES sur l'ÉTAGÈRE
Ils PASSENT à CÔTÉ de la MAISON en COURANT
Le TRAIN a eu un TERRIBLE ACCIDENT
L'ÉVIER de la CUISINE est VIDE
Un GARÇON est TOMBÉ de la FENÊTRE

Elle PREND sa PETITE CUILLÈRE
Le PARC est PRÈS de la ROUTE
Le CHEF DÉCOUPE des OIGNONS
Le CHIEN ABOIE MÉCHAMMENT
Il NETTOIE SON VISAGE
QUELQU'UN a VOLÉ l'ARGENT
La LUMIÈRE S'EST ÉTEINTE
Ils VOULAIENT des POMMES de TERRE
La petite FILLE MÉCHANTE CRIAIT
Le LAIT FROID est dans une CRUCHE
La PEINTURE GOUTTAIT sur le SOL
La MÈRE REMUAIT le THÉ
Ils RIAIENT à SON HISTOIRE
Les HOMMES PORTENT des PANTALONS LONGS
Le PETIT GARÇON était ENDORMI
La FEMME ENTRE DANS le MAGASIN
Le SOLEIL fait FONDRE la NEIGE
Le PÈRE RENTRE à la MAISON
Elle AVAIT son ARGENT de POCHE
Le CAMION MONTE la CÔTE
Il EMPORTE son MANTEAU de PLUIE
Un COUTEAU POINTU est DANGEREUX
ILS PRENNENT de la NOURRITURE
Les FILLES INTELLIGENTES LISENT
Le BALAIS ÉTAIT dans le COIN
La FEMME RANGEAIT sa MAISON
Les ENFANTS ont fait TOMBER le SAC
Le CHIEN EST REVENU
Le SOL avait l'AIR PROPRE
Elle a RETROUVÉ SON PORTE-MONNAIE
Le FRUIT est ÉTALÉ sur le SOL
MAMAN PREND une CASSEROLE
Ils se LAVÈRENT DANS de l'EAU FROIDE
Les JEUNES GENS DANSENT
Le BUS est PARTI TÔT
Ils AVAIENT DEUX BOUTEILLES VIDES
Une BALLE REBONDIT sur le SOL
Le PÈRE a OUBLIÉ le PAIN
La FILLE a un LIVRE d'IMAGES
L'ORANGE était PLUTOT SUCREÉ
IL se BOUCHE le NEZ
La NOUVELLE ROUTE est sur la CARTE
Le GARÇON a OUBLIÉ son LIVRE
Un AMI est PASSÉ DÉJEUNER
Les BOITES d'ALLUMETTES sont VIDES
Il GRIMPE à SON ÉCHELLE
La FAMILLE a ACHETÉ une MAISON
La CRUCHE est SUR l'ÉTAGÈRE

La BALON a CASSÉ la FENÊTRE
ILS ACHÈTENT du FROMAGE
L'EAU de la MARE est SALE
ILS ont ENTENDU un DRÔLE de BRUIT
La POLICE ÉVACUE la RUE
Le BUS s'est ARRÊTÉ SOUDAINEMENT
Elle ÉCRIT à SON FRÈRE
Le JOUEUR de foot a PERDU une CHAUSSURE
Les TROIS filles ÉCOUTENT BIEN
Le MANTEAU est POSÉ SUR une CHAISE
Le LIVRE RACONTE une HISTOIRE
Le JEUNE GARÇON est PARTI de la MAISON
ILS MONTENT à l'ARBRE
Elle se TENAIT à CÔTÉ de sa FENÊTRE
La TABLE a TROIS PIEDS
Une LETTRE est TOMBÉE sur le TAPIS
Les CINQ HOMMES TRAVAILLENT
Il ÉCOUTE SON PÈRE
Les CHAUSSURES ÉTAIENT très SALES
Ils SONT ALLÉS en VACANCES
Le BÉBÉ a CASSÉ sa TASSE
La FEMME PRÉPARAIT son SAC
L'ASSIETTE du DÎNER était CHAUDE
Le TRAIN ROULE VITE
L'ENFANT BUVAIT du LAIT
La VOITURE est RENTRÉE dans le MUR
Un TORCHON était à CÔTÉ de l'ÉVIER
La FEMME de MÉNAGE UTILISAIT un BALAI
Elle se REGARDAIT DANS le MIROIR
Le GENTIL GARÇON AIDE
ILS SUIVAIENT le CHEMIN
L'HORLOGE de la CUISINE n'était pas à l'HEURE
Le CHIEN a SAUTÉ sur la CHAISE
QUELQU'UN TRAVERSE la RUE
Le FACTEUR APPORTE une LETTRE
ILS se PROMENENT à VÉLO
Il S'EST CASSÉ la JAMBE
Le LAIT était à CÔTÉ de la PORTE
Les CHEMISES sont ÉTENDUES dans le PLACARD
Le SOL était TROP DUR
Les SEAUX sont REMPLIS d'EAU
Les POULES PONDENT des OEUFS
Le MAGASIN de BONBONS était VIDE
Les CHIENS sont ALLÉS se PROMENER
ELLE ATTEND sa ROBE
La FEMME est RESTÉE pour le THÉ
Le CHAUFFEUR attend au COIN de la RUE
Ils ONT FINIT de DÎNER

Le POLICIER CONNAIT le CHEMIN
La PETITE FILLE était CONTENTE
IL PORTAIT sa CHEMISE JAUNE
ILS VIENNENT pour NOËL
La VACHE DONNE du LAIT
Le GARÇON S'EST MIS au LIT
Les DEUX FERMIERS se PARLENT
La MÈRE RAMASSE des FLEURS
Le POISSON est POSÉ sur l'ASSIÈTTE
Le PÈRE ÉCRIT une LETTRE
La NOURRITURE COUTE beaucoup d'ARGENT
La FILLE LAVE ses CHEVEUX
Le JARDIN DEVANT la maison est JOLI
IL a PERDU son CHAPEAU
Les ROBINETS sont AU-DESSUS de l'ÉVIER
Le PÈRE a PAYÉ A la BARRIÈRE
Elle ATTEND SON BUS
Le CAMION du BOULANGER est PASSÉ
Ils ont MANGÉ de la VIANDE FROIDE
Le MATCH de FOOT est FINI
ILS PORTENT des SACS de COURSES
Les ENFANTS AIDENT le LAITIER
La PHOTO est TIRÉE d'un LIVRE
Le GÂTEAU de RIZ était PRÊT
Le GARÇON avait un DRAGON en JOUET
Un ARBRE est TOMBÉ sur la MAISON
Le FRUIT ÉTAIT dans une BOÎTE
Le MARI APPORTE des FLEURS
Ils JOUENT DANS le PARC
ELLE se DISPUTE avec sa SŒUR
Un HOMME a PRÉVENU la POLICE
Les POMMES de terre POUSSENT dans le SOL
IL NETTOIE sa VOITURE
La SOURIS a TROUVÉ le FROMAGE
ILS ont ATTENDU PENDANT une HEURE
Le GROS CHIEN était DANGEREUX
La CONFITURE de FRAISES était SUCRÉE
La PLANTE est PENDUE AU-DESSUS de la PORTE
TOUS les ENFANTS MANGENT
Le GARÇON a les CHEVEUX NOIRS
La MAMAN a ENTENDU son BÉBÉ
Le CAMION est MONTÉ en haut de la CÔTE
L'HOMME en COLÈRE CRIA
Le CHIEN DORS dans son PANIER
ILS BOIVENT du THÉ
La MÈRE OUVRE le TIROIR
Une VIEILLE FEMME était à la MAISON
Il a FAIT TOMBER son ARGENT

Ils ONT CASSÉ TOUS les OEUFS
La FENÊTRE de la CUISINE était PROPRE
La FILLE JOUE avec le BÉBÉ
Le GROS POISSON S'EST ÉCHAPPÉ
Elle AIDE SES AMIS
Les ENFANTS ont LAVÉ les ASSIETTES
Le FACTEUR est PASSÉ TÔT
La PANCARTE INDIQUE le CHEMIN
L'HERBE est TROP LONGUE
L'ALLUMETTE est TOMBÉE sur le SOL
Un HOMME OUVRE le ROBINET
Le FEU était TRÈS CHAUD
Il SUCE SON POUCE
Le MAGASIN est fermé PENDANT MIDI
Le CONDUCTEUR DÉMARRE le MOTEUR
Le GARÇON s'est DÉPÊCHÉ d'aller à l'ÉCOLE
Des GENS SYMPATHIQUES sont VENUS
ELLE s'est COGNÉE la TÊTE
Ils ONT RENCONTRÉ LEURS AMIS
Les FLEURS POUSSENT dans le JARDIN
Le PETIT BÉBÉ était JOLI
La FILLE MET la TABLE
ILS ont MARCHÉ SUR l'HERBE
La MÈRE fait un NŒUD avec la FICELLE
Le TRAIN S'ARRÊTÉ à la GARE
Le CHIOT JOUE avec une BALLE
Les ENFANTS SALUENT le TRAIN
La MÈRE COUPE le GÂTEAU de NOËL
IL FERME les YEUX
Le MANTEAU de PLUIE était MOUILLÉ
Une FEMME ACHÈTE du BEURRE
Ils ONT APPELÉ une AMBULANCE
Elle PAYE SON PAIN
Le POLICIER a TROUVÉ un CHIEN
Certains HOMMES se RASENT le MATIN
Le CONDUCTEUR a PERDU son CHEMIN
ILS ADMIRAIENT la PHOTO
Le CHAT BOIT dans une SOUCOUPE
La PORTE du FOUR était OUVERTE
La VOITURE ROULE TROP VITE
Le petit GARÇON BÊTE se CACHE
Le PEINTRE UTILISE un PINCEAU
La TARTE aux POMMES CUIT
Il BOIT DANS sa TASSE
Le CIEL était TRÈS BLEU
Ils ONT TAPÉ à la FENÊTRE
Le GRAND GARÇON TAPE dans la BALLE
Les GENS RENTRENT chez EUX

Le BÉBÉ VEUT son BIBERON
La FEMME est ASSISE sur sa CHAISE
Ils ont MANGÉ un GÂTEAU à la CONFITURE
Les CISEAUX sont ASSEZ POINTUS
ELLE APPELLE sa FILLE
DES FEUILLES sont TOMBÉES de l’ARBRE
Le LAITIER APPORTE de la CRÈME
Une FILLE court le LONG du CHEMIN
La MAMAN LIT le JOURNAL
Le CHIEN POURSUIVAIT le CHAT
La PÂTISSERIE s'est OUVERTE ce MATIN
ILS AIMENT la CONFITURE d'ORANGES
La MÈRE FERME la FENÊTRE
Il FAIT du PATIN AVEC son AMI
La TARTE au FROMAGE était BONNE
La PLUIE TOMBE des NUAGES
ELLE PARLAIT à sa POUPÉE
ILS ont PEINT le MUR
La SERVIETTE est TOMBÉE par TERRE
Le CHIEN MANGE de la VIANDE
Un GARÇON a CASSÉ la BARRIÈRE
Les POIRES JAUNES étaient BONNES
La POLICE AIDE le CONDUCTEUR
La NEIGE RECOUVRAIT le TOIT
La FEMME NETTOYAIT la CHEMISE
La TASSE est PENDUE à un CROCHET
La FAMILLE AIME le POISSON
Le SUCRE est en PLUSIEURS MORCEAUX
Le BÉBÉ est ALLONGÉ sur un TAPIS
La MACHINE à LAVER est CASSÉE
ILS DÉBARASSENT la TABLE
La femme de MÉNAGE BALAYE par TERRE
Un ÉPICIER VEND du BEURRE
L'EAU du BAIN est TIÈDE
Il PREND SA CUILLÈRE
ELLE s'est fait MAL à la MAIN
Le LAITIER CONDUIT un PETIT CAMION
Le GARÇON est TOMBÉ DANS les ESCALIERS
ILS RESTENT pour SOUPER
La FILLE TENAIT un MIROIR
La TASSE est POSÉE sur une SOUCOUPE
Les VACHES sont PARTIES au MARCHÉ
Le GARÇON a EU des ENNUIS
ILS SORTENT ce SOIR
Le BALON de FOOT a touché le BUT
IL a PAYE son ADDITION
Le TORCHON est ASSEZ MOUILLÉ
Un CHAT a SAUTÉ DE la BARRIÈRE

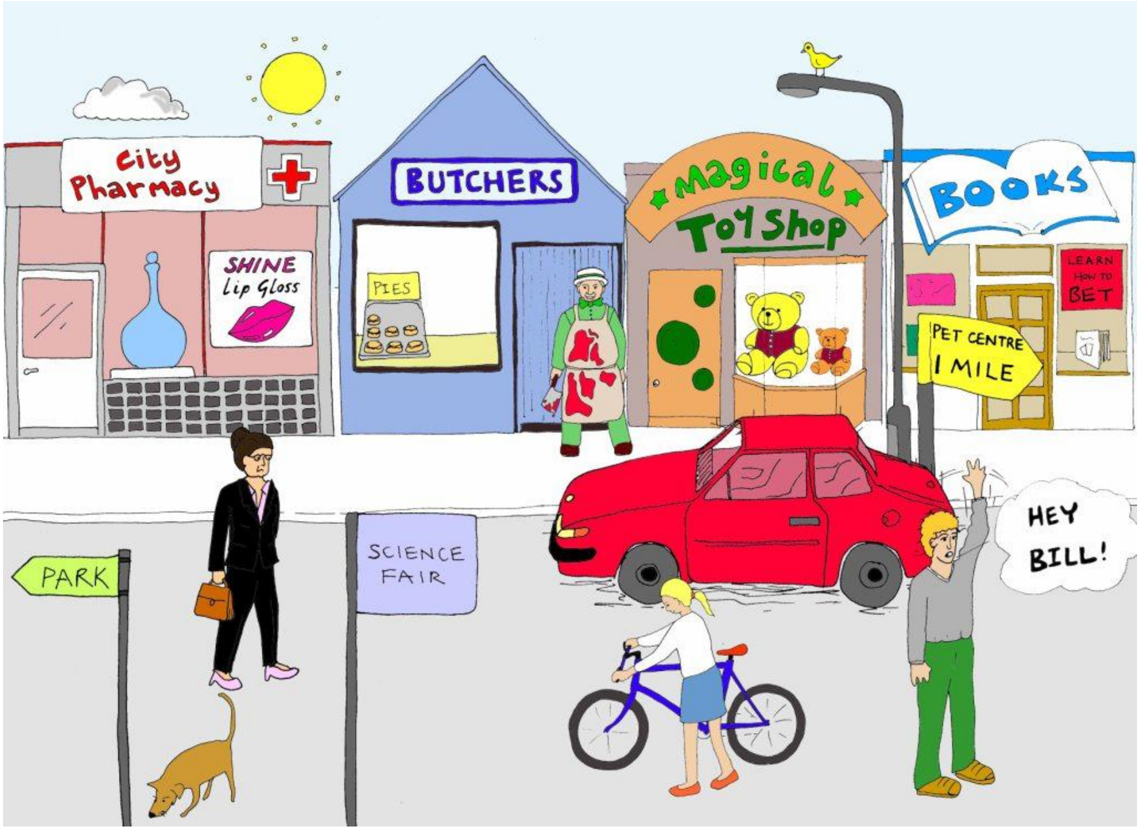
Le BÉBÉ a les YEUX BLEUS
Ils se SONT ASSIS sur un BANC en BOIS
La MÈRE a FAIT des RIDEAUX
Le FOUR est TROP CHAUD
La FILLE a ATTRAPÉ un RHUME
Le MANTEAU de PLUIE est ÉTENDU
ELLE s'est BROSSÉE les CHEVEUX
Les DEUX ENFANTS RIENT
L'HOMME MET son ÉCHARPE
Les FLEURS sont DANS un POT
Le POT de POIVRE était VIDE
Le CHIEN BOIT dans un BOL
Une FILLE ENTRE dans la PIÈCE
ILS POUSSENT une VIEILLE VOITURE
Le CHAT a ATTRAPÉ une SOURIS
La ROUTE MONTE en HAUT d'une COLLINE
Elle a FAIT SON LIT
Les BANANES sont des FRUITS JAUNE
La VACHE est ALLONGÉE dans l'HERBE
Les COQUETIERS SONT sur l'ÉTAGÈRE
Il a FAIT PEUR à sa SŒUR
L'ÉQUIPE de CRICKET JOUE
Le PÈRE RAMASSE des POIRES
La bouilloire à VITE CHAUFFER l'EAU
L'HOMME a PEINT une INDICATION
ILS ont PERDU de l'ARGENT

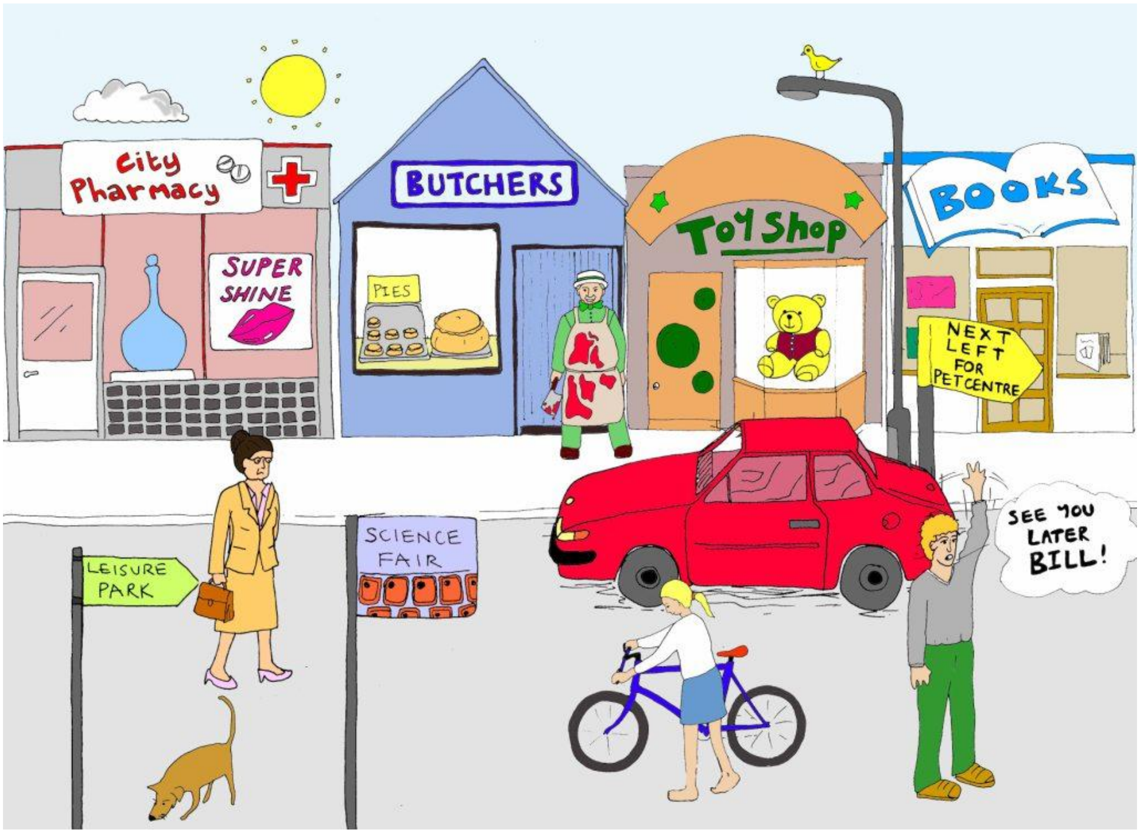
8.3. Appendix 3: Diapix materials

Diapix pictures used for the 'spot the difference' task in Chapter 4 to elicit dialogue between the French L2 and SSBE speaker. Courtesy of Baker and Hazan (2011), materials available online at: http://www.springerlink.com.libproxy.ucl.ac.uk/content/r3w63v3243m61g75/13428_2011_Article_75_ESM.html

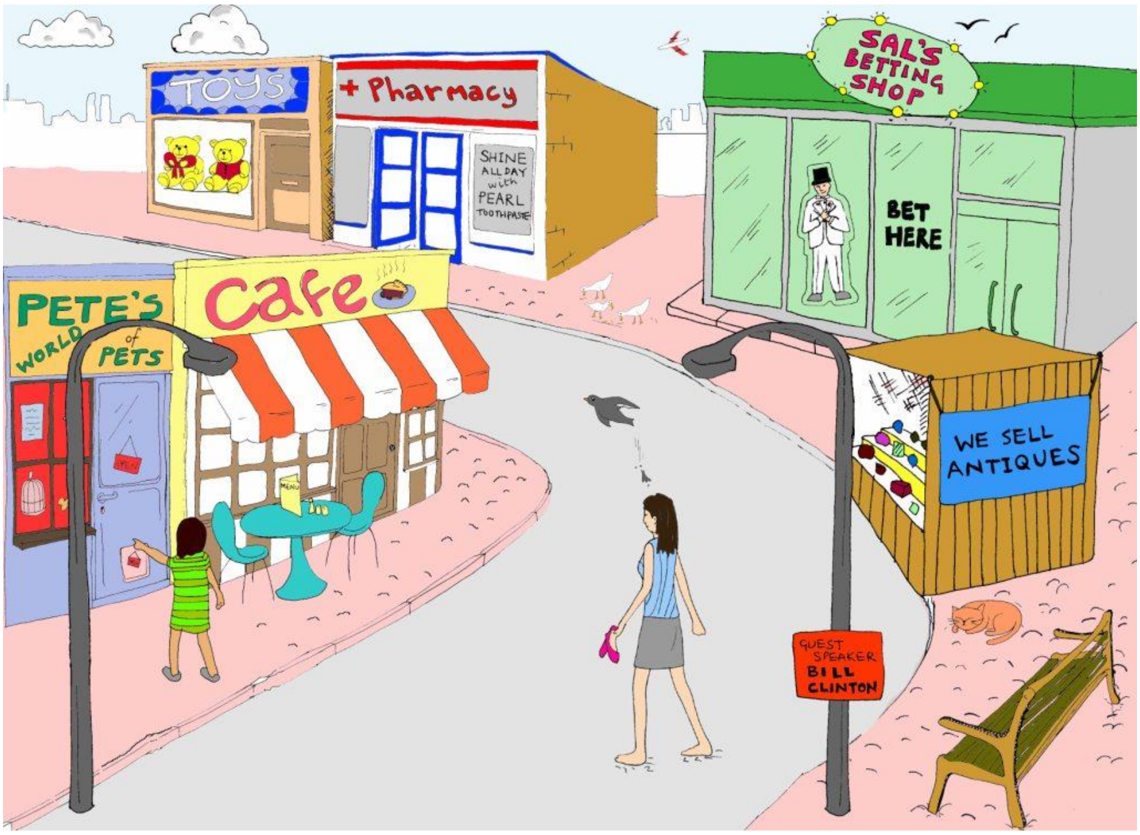

PAUL & KIM'S cafe
Shack TO RENT
PUSH ONCE
BEACH VOLLEYBALL LESSONS
SHORE FORECAST SUN

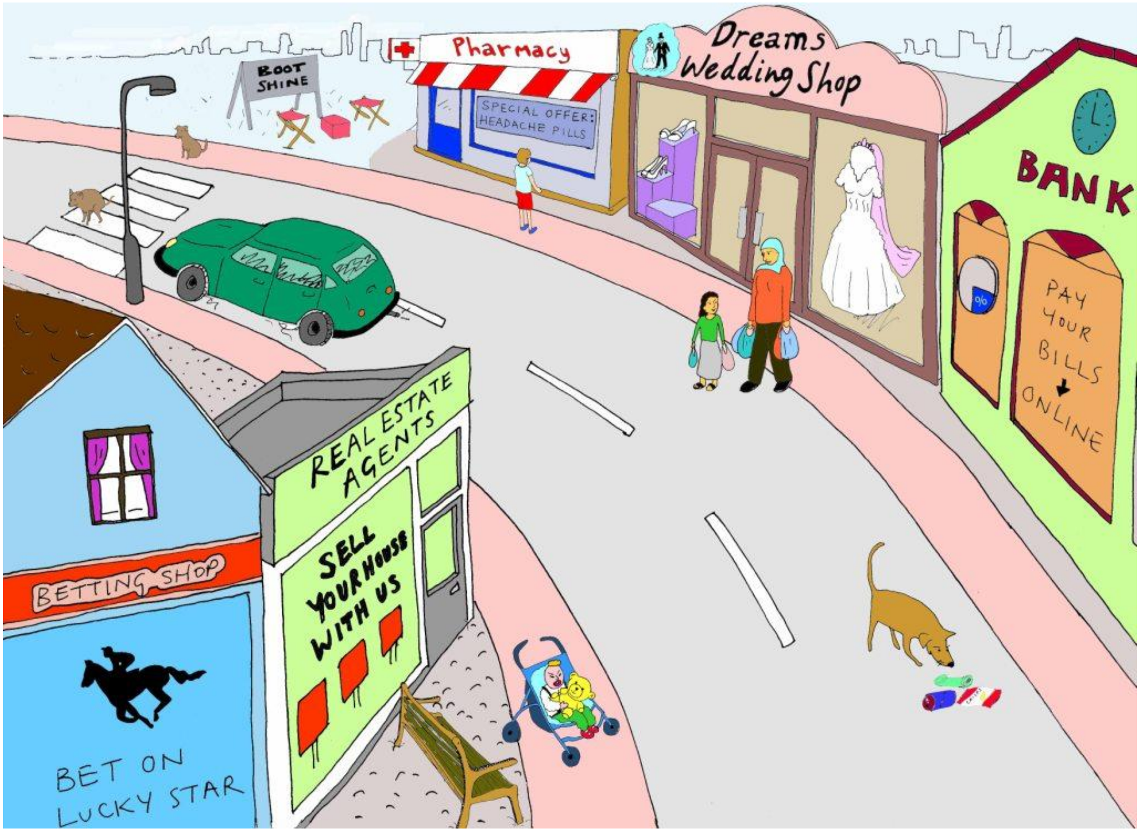

TOM & PAUL'S cafe
PEACH JUICE
PUSH HERE
SHORE FORECAST RAIN


City Pharmacy
SHINE Lip Gloss
BUTCHERS
PIES
Magical Toy Shop
BOOKS
LEARN HOW TO BET
PET CENTRE 1 MILE
HEY BILL!
PARK
SCIENCE FAIR


City Pharmacy
SUPER SHINE
BUTCHERS
PIES
Toy Shop
BOOKS
NEXT LEFT FOR PET CENTRE
SEE YOU LATER BILL!
LEISURE PARK
SCIENCE FAIR


TOYS
+ Pharmacy
PILL BOXES ON SALE
EXTRA SHINE with PEARL TOOTHPASTE
PLACE YOUR BETS
PETE'S PET SHOP
Cafe
SELL YOUR SECOND HAND GOODS
DOCTOR


SAL'S BETTING SHOP
TOYS
+ Pharmacy
SHINE ALL DAY with PEARL TOOTHPASTE
BET HERE
PETE'S WORLD of PETS
Cafe
WE SELL ANTIQUES
GUEST SPEAKER BILL CLINTON


BOOT SHINE
Pharmacy
SPECIAL OFFER: HEADACHE PILLS
Dreams Wedding Shop
BANK
PAY YOUR BILLS ONLINE
REAL ESTATE AGENTS
SELL YOUR HOUSE WITH US
BETTING SHOP
BET ON LUCKY STAR


SHINE AND POLISH
Pharmacy
HEADACHE PILLS: 2 FOR 1
BANK
PHONE BANKING: PAY BILLS
DISCOUNT PET INSURANCE
REAL ESTATE AGENTS
LET US SELL YOUR HOME
BETTING SHOP