The University of Lodz
Faculty of Philology
Institute of English Studies

Magdalena Zając

Phonetic convergence in the speech of Polish learners of English

Doctoral thesis written in the Institute of English studies under the supervision of prof. UŁ dr hab. Ewa Waniek-Klimczak

Lodz, 2015

Uniwersytet Łódzki
Wydział Filologiczny
Instytut Anglistyki

Magdalena Zając

Konwergencja fonetyczna u polskich użytkowników języka angielskiego

Praca doktorska napisana w Instytucie Anglistyki pod kierunkiem prof. UŁ dr hab. Ewy Waniek-Klimczak

Łódź, 2015

Contents

List of Tables

List of Figures

Introduction

This dissertation investigates variability in L2 pronunciation, focusing on speech convergence as a result of exposure to different phonetic varieties. The term speech convergence refers to a process during which speakers adapt their linguistic behaviour according to who they are talking or listening to. Previous studies have examined various aspects of this phenomenon; it has been investigated under different names (accommodation, imitation, alignment) and with the use of quite distinct methodological procedures (e.g. spontaneous conversational interactions, socially minimal laboratory settings). Regardless of methodological and terminological discrepancies, any types of linguistics adjustments that are made upon exposure to the speech of another person are viewed as instances of speech convergence in the current study. It should be emphasised that the term speech (or phonetic) convergence will be used to refer to different types of linguistic behaviour and will encompass shifts both towards and away from the speech of another individual. This application of the term may seem misleading as the word *convergence* suggests making one's speech more similar to that of another person (rather than making it different). However, the usage of the name speech convergence in the current study is carefully thought-out and deliberate. Other names that have been commonly used to discuss speech adjustments include accommodation and imitation. Neither of them was selected for the purposes of the present investigation since they are both associated with very specific types of methodology (interactional vs. non-interactional experimental design). The term convergence, on the other hand, is more neutral in the sense that it has not been used in one type of investigation exclusively. What is more, it can be found in almost all previous work on speech adjustments (both in those studies that refer to the process as accommodation and those that call it imitation). It is also important to note that the expression convergence strategies will be used to refer to three types of linguistic behaviour: convergence (making one's speech more similar to that of another person), divergence (moving away from the speech of another person) and maintenance (maintaining one's default linguistic behaviour in spite of exposure to the speech of another person). Therefore, the word *convergence* will have two slightly different meanings in this dissertation. When discussing the general phenomenon, the expressions speech convergence or phonetic convergence will denote a situation where speakers adapt their linguistic behaviour depending on who they are talking or listening to. When discussing particular types of linguistic behaviour, the name convergence will refer to the process of making one's

speech more similar to that of another person. Finally, it should be pointed out that the term imitation will be used to refer to convergence strategies in a non-interactional setting (i.e. one that does not include social interaction), whereas the name accommodation will be applied in relation to speech behaviour in an interactional setting (i.e. one that does entail some type of social interaction).

The aim of the experimental study that forms the central part of this dissertation is to examine L2 convergence strategies upon exposure to native and non-native pronunciation. The study concentrates on the speech behaviour of advanced Polish learners of English, who are exposed to two pronunciation varieties: Polish-accented English and native English. The issue seems worth investigating for two reasons. Firstly, previous research on convergence in non-native productions suggest that the process does take place in L2 speech. Some of the studies have examined convergence strategies upon exposure to native speech, others have focused on interactions between speakers who communicate in a shared second language. Nevertheless, it seems that the issue investigated in the current study, i.e. phonetic convergence towards native speakers as compared with convergence towards other learners of the target language, has not as yet been thoroughly examined. More generally, phonetic convergence in L2 speech seems to be an interesting and fruitful research area due to the complex nature of the L2 sound system. Research on second language acquisition suggests that learners' productions in a second language are a product of an independent, self-governed linguistic system, which does not correspond exactly either to the L1 or the L2 and restructures itself as the learner gains more knowledge of the target language. What is more, L2 phonetic performance has been found to be conditioned by a variety of social-psychological, psycholinguistic and linguistic factors. Thus, unlike linguistic behaviour in a speaker's first language, it could be hypothesised that L2 convergence strategies will not only differ as a function of the speech variety one is exposed to (e.g. native vs. non-native productions) but will also be affected by a number of factors specific to L2 speech only. The interaction between these two types of conditioning deserves further investigation.

The dissertation is organised into four chapters; the first two provide theoretical background, the next two describe the study and its findings. Chapter One reviews previous research on speech convergence. The chapter describes the methodology and approaches used in previous work; the factors that may condition convergence strategies are also discussed. It is important to note that all of the studies presented in Chapter One are concerned with L1 data. Although the dissertation concentrates on L2 speech convergence, findings pertaining to L1 convergence are considered relevant to the present investigation. Following Adjemian

(1976) and Tarone (1979) (see Chapter Two), the L2 linguistic system is viewed as equivalent to L1 system in the sense that it varies in different social contexts, depending on factors such as topic, focus on language form, interlocutor, etc. Accordingly, it is assumed that L2 speech convergence may be motivated and modified by similar social-psychological factors as convergence in L1 speech.

Chapter Two provides an overview of some of the previous work on L2 pronunciation. The structure and development of the L2 sound system and the various factors that may influence L2 phonetic performance are described. It should be stressed that Chapter Two focuses on the research considered particularly relevant to the present investigation. Selected findings are presented since L2 phonetics has become a very productive research area. Presenting the results of all of the previous studies did not seem feasible or necessary for the current investigation.

Chapter Three describes the study on speech convergence in the pronunciation of Polish learners of English. First, an overview of previous research on L2 speech convergence is provided. The following sections are concerned with the aims of the study, pilot work, hypotheses and methodology. The final section of the chapter presents the results. The discussion of pilot work constitutes and important part of the chapter; the findings of the pilot studies have revealed several methodological issues that were addressed in the current investigation and have lead to the development of a new experimental procedure.

In Chapter Four, the results of the study on phonetic convergence in the speech of Polish learners of English are analysed and discussed. Additionally, the chapter provides an evaluation of the experimental method. The section is included in the chapter since the introduction of a new experimental procedure constituted and important element of the current study. The final section of the chapter offers suggestions for further research.

The final section of the dissertation (Conclusions) reviews the experimental procedure and summarises the results of the study. The appendices contain the questionnaire and PowerPoint presentations that were used in the experiment.

Chapter One: Speech convergence

1.1. Introduction

The purpose of this chapter is to describe the process of speech convergence, review the most relevant studies concerned with the phenomenon and discuss their implications. Speech convergence will be taken to mean the speech adjustments that an individual makes as a result of exposure to the speech of another person or people. The term is used by the author to encompass adjustments of phonetic features such as fundamental frequency, vowel quality or the length of VOT, non-content speech behaviour such as the frequency of back-channel responses or laughter and, in the instance of one classical study on accommodation, shifts from one language to another. Studies concerned with the adjustments in the syntactic structure of utterances (e.g. Levelt and Kelter, 1982; Bock, 1986; Branigan, Pickering and Cleland, 2000) or the tone of voice (Neumann and Strack, 2000) and adjustments made upon exposure to lip-read speech (Miller, Sanchez and Rosenblum, 2010) are not included in this chapter, as they were considered to fall outside the scope of the current investigation and do not seem to afford additional insights into the mechanisms that underlie speech convergence.

The phenomenon of speech convergence has been explored under different names and with the use of various frameworks and methodological procedures. Earlier research regards it as a process that takes place in conversational interactions and has a social-psychological basis. There, it is usually termed accommodation or convergence. In many of the more recent studies, the process is examined in non-interactional, laboratory settings and is mostly referred to as imitation. Imitation is often treated as an automatic and unintentional reflex of the brain that develops when one is still an infant. The two seemingly conflicting approaches are merged in a number of recent studies on speech convergence, where the process is considered from a social-psychological viewpoint, while at the same time being investigated with the use of laboratory-based methodology.

The first two sections of Chapter One survey research carried out in conversational interactions and focus on the social-psychological aspect of speech convergence. Section 1.2. describes and exemplifies the tenets of Communication Accommodation Theory (CAT), while Section 1.3. pertains to the studies carried out outside the CAT framework. The next two sections are concerned with laboratory-based studies where speech convergence is investigated in socially-minimal settings. Section 1.4. reviews the studies in which the process

is mostly viewed as an automatic cognitive reflex. Section 1.5. describes the laboratory-based research that incorporates social-psychological elements. The final section provides a summary and interpretation of all the relevant findings.

1.2. Speech convergence in Communication Accommodation Theory

Communication Accommodation Theory (CAT), developed by Howard Giles and colleagues in the 1970s, proposes that some of the speech adjustments that individuals make in an interaction are driven by their personal and social identities. More specifically, as explained by Giles and Ogay (2007: 326), it "[...] provides a wide-ranging framework aimed at predicting and explaining many of the adjustments individuals make to create, maintain or decrease social distance in an interaction". Initially the theory was concerned with accent and bilingual shifts, later it broadened its focus and eventually evolved into an "integrated, interdisciplinary statement of relational processes in communicative interaction" (Giles, Coupland and Coupland, 1991: 2). Although language change remains the main subject of research in CAT, the framework can also be applied to examine nonverbal communication (e.g. smiling, gesticulating) and communicative symbols such as dress or hair style (Giles and Ogay, 2007).

The roots of CAT can be traced back to Giles's (1973) "accent mobility" model, which was introduced in response to Labov's (1966) proposition that shifts in pronunciation are triggered by changes in stylistic context. Giles postulated that one should also consider interpersonal aspects in the analysis of pronunciation change and designed an experiment to test the hypothesis that accent mobility may be person-based and depend on the social motivational tendencies of the speaker. The informants in the experiment were Bristol-born male teenagers who spoke with a Bristol accent and had a working-class background. They were interviewed under two conditions: by an older Received Pronunciation speaker and by another teenager born in Bristol. It was assumed that the subjects would perceive the first interlocutor as "of a higher prestige in terms of age, education and accent usage in relation to themselves" (Giles, 1973: 94) and view the second interviewer as "of equivalent prestige in terms of age, education and accent usage" (ibid.). The analysis was based on the judgments of naive listeners, who were asked to assess whether any accent and grammatical changes occurred in the speech of a given participant by listening to samples taken from the two interviews. The results of the experiment implied that the informants' pronunciation and

lexico-grammatical usage did change depending on which interlocutor they talked to. The author's interpretation of the data was that Bristol teenagers standardized their speech when they were interviewed by the RP speaker. The strategy they used was termed "convergence". As explained by Giles (1973: 90), "[...] if the sender in a dyadic situation wishes to gain the receiver's social approval then he may adapt his accent patterns towards that of this person, i.e. reduce pronunciation dissimilarities – accent convergence." It would appear that the Bristol-born teenagers accommodated their pronunciation towards that of the RP interviewer in order to make a favourable impression, thus lending support to Giles's (1973) postulate that apart from being conditioned by stylistic context, accent shifts may also vary as a function of interlocutor and the speaker's attitude towards them.

The design of Giles's (1973) study and his interpretation of the results call attention to an important aspect of CAT, namely that it draws extensively on concepts derived from social psychology, such as similarity attraction (Byrne, 1971). This theory posits that we are inclined to like people to whom we bear some kind of similarity more than those with whom we have little in common. It also implies that by reducing dissimilarities we may induce others to look upon us more favourably, which is visibly at the core of Giles's (1973) statements concerning Bristol teenagers and their attempts to gain social approval through speech convergence.

Another study which laid the foundations for CAT pertained not to changes in pronunciation within an L1 but to language shifts in bilingual speakers (Giles, Taylor and Bourhis, 1973). The subjects were English Canadian students from the province of Quebec, who heard a recording of a French Canadian describing a picture and were required to draw it while listening. The participants were divided into four groups, the first group heard the French Canadian talk about the drawing in French, the second heard him use a mixture of French and English, the third was exposed to fluent English speech, and the fourth heard the French Canadian speaking nonfluent English. The participants were told the French Canadian speaker was aware that his recording would be later played to English Canadians and that he could decide which language to use in this task. The next stage of the experiment involved asking the subjects to rate their reactions towards the French Canadian and evaluate his performance. Finally, the English Canadians recorded a description of another picture, supposedly for the French Canadian to draw later. The hypothesis formulated by the authors of the study was that the more effort the French Canadian was perceived to put into accommodation towards the English Canadians, the more favourably he would be viewed and the more effort would be put into accommodating back to him. The results revealed that some

of the participants did converge towards the French Canadian (by responding in French or in English and French) and that they were more likely to accommodate if he had previously accommodated to them (i.e. used English or a mixture of French and English). Building on the assumptions of similarity attraction, the authors concluded that "[...] the results of the study [...] do support the notion [...] that accommodating individuals induce their recipients to evaluate them more favourably (Giles et al., 1973: 187)". The authors also suggested that the phenomenon of speech convergence is "[...] a reflection of an individual's desire for social approval" (ibid.).

Coupland (1984) recorded the spoken interactions between a travel agency assistant and her 51 clients. The participants were all residents of Cardiff with different socioeconomic backgrounds. The phonological variables under investigation were the usage of H-dropping, T-voicing/tapping, G-dropping and simplification of final consonant clusters. The variables were selected on the basis of their function in marking social and stylistic variation in Cardiff English. The frequency of non-standard feature usage in the assistant's speech was first compared with the frequency of non-standard feature usage in the clients' speech and then contrasted across her interactions with representatives of different socioeconomic groups. The results of the study revealed that the assistant converged her pronunciation towards that of her clients. As regards the functions of convergence, Coupland mentions the desire for social approval and stresses the need for communication efficiency. Giles and Ogay (2007) argue that converging can serve as a means of decreasing uncertainty and anxiety that a person might experience during an interaction, as it renders it more predictable and facilitates communication. Communication efficiency as a motive for accommodation is also discussed by Gallois et al. (1995), who propose that convergence may result from a desire to make the interaction flow more smoothly.

When describing the sources of convergent behaviour, Giles et al. (1991) place much emphasis on the importance of power relations. The issue is taken up in a study on the talkers' F0 by Gregory and Webster (1996). The research is based on recordings of twenty five interviews between talk show host Larry King and his guests (politicians, well-known entertainers and athletes) on the CNN *Larry King Live* talk show. The goal of the study was to test the hypotheses that the participants of the study would converge in F0 towards their conversational partners and that the magnitude of convergence would be determined by the talkers' relative social status. The data supported the hypotheses, revealing that Larry King accommodated towards higher status guests, whereas lower status guests converged towards him. The findings testify to the claim that power or status relations can moderate the degree

and direction of convergence. Interestingly, it was also found that in the case of "deferent partners" accommodating towards their "dominant partner", convergence did not increase over time. In the case of lower status partners, on the other hand, the amount of accommodation did increase as time passed.

As described above, convergence consists in making one's communicative behaviour more similar to that of the interlocutor. A strategy with the opposite function is termed divergence and involves emphasising speech differences between the interactants. Adopted in order to dissociate oneself from the conversational partner, it might ensue when an individual finds their conversational partner's characteristics, attitudes or beliefs undesirable (Giles, 1973). Interestingly, as observed by Giles et al. (1991: 27), "[c]onvergence is a strategy of identification with the communication patterns of an individual internal to the interaction, whereas divergence is a strategy of identification with linguistic communicative norms of some reference group external to the immediate situation". Just as convergence is linked to similarity attraction, divergence is grounded in the social psychological theory of intergroup relations (Tajfel, 1974; Turner, 1975; Tajfel and Turner, 1979; Tajfel, 1981; in Tajfel, 1982). The theory hinges on the premise that social identity is an essential component of one's self image and predicts that when individuals interact with representatives of a different social group, they will search for characteristics that make them positively distinct from the members of the outgroup. The purpose of this process is to achieve a positive image of one's ingroup and thus enhance the positive evaluation of one's self image. Given that distinguishing oneself and one's ingroup from others may be realised through speech divergence, it is easy to see why the advocates of CAT embraced this theory.

An interesting and influential study that examines the mechanism of divergence was carried out by Bourhis and Giles (1977). The study is concerned with two groups of Welsh-born adults learning the Welsh language. The first group attached considerable importance to national group membership, the second attended Welsh language classes to further their careers. Participants from both groups were asked to help in a survey on language learning techniques in which they responded to an RP-speaking interviewer's pre-recorded questions. The interview began with emotionally neutral questions followed by a statement that Welsh was a "dying language with a dismal future", which was designed to threaten the subjects' feeling of ethnic identity. Participants' performance was evaluated by two raters who were naive to the experiment and were not linguistically trained. The results revealed that when the importance of learning Welsh had been challenged, the group of informants who exhibited a stronger sense of national identity diverged from the RP interviewer by broadening their

Welsh accent. This finding illustrates how divergent behaviour may be prompted by a desire to distance oneself from the conversational partner and lends support to the argument that the need to express one's social identity may motivate language shifts.

Identity-related accent shifts were examined more recently by Llamas, Watt and Johnson (2009), who focused on the pronunciation of native English participants from Berwick-upon-Tweed. The subjects were recorded while conversing with representatives of different varieties of English and a non-native speaker of the language. The analysed phonetic variables were the quality of /r/, the quality of the lettER vowel and vowel length. One of the goals of the study was to investigate participants' potential divergence from the members of the outgroup (the different interlocutors). Unfortunately, the findings concerning accommodation proved inconclusive.

A framework that nicely complements the tenets of CAT is Bell's (1984) theory of audience design. One of its fundamental assumptions is that although speakers accommodate primarily to their addressees, third persons may also cause an individual to shift their speech patterns. In other words, speakers are believed to design their talks for the audience of their utterances. According to Bell, several audience types may be identified. For example, apart from the person who is directly addressed in an interaction, Bell also discriminates between auditors, who are known and present in an interaction but not directly addressed by the speaker, and overhearers, who are known to be there by the speaker but are not ratified participants. As support for his hypothesis, Bell mentions a study by Douglas-Cowie (1978, in Bell, 1984) that is concerned with linguistic code-switching in a northern Irish village. The informants were recorded when interacting one-on-one with a fellow-villager, when talking to a fellow-villager with an English outsider as auditor, and when directly addressing the English outsider. It was found that the subjects adjusted some features of their pronunciation towards the English interviewer both when he acted as addressee and when he acted solely as auditor. The results validate Bell's claims that speakers that are not directly involved in an interaction may to some extent affect an individual's speech patterns. Bell also refers to earlier studies (Bell, 1977; Bell, 1982a; Bell, 1982b, in Bell, 1984), in which he examined the pronunciation shifts in the speech of New Zealand radio newscasters. It was found that when the newscasters read news on a station with higher-status audience, they used more standard pronunciation features than when reading news on a station with lower-status audience, thus accommodating towards different addressees. The results of the study show that interacting partners do not need to participate in a face-to-face, two-way interaction in order for accommodation to occur.

Finally, it should be mentioned that apart from convergence and divergence, Giles and colleagues recognise one other type of linguistic behaviour, i.e. maintenance. When a person adopts this strategy, their communicative behaviour remains unchanged during an interaction. In other words, a person continues to use a given style irrespective of the style of his or her interlocutor (Giles and Ogay, 2007). As an example, Giles and Sachdev (2008) mention an Arab diplomat giving a speech addressed to international audiences in Arabic instead of English, which had previously been the case during such events. Bourhis (1984) observed some instances of using this strategy in his work on Francophones and Anglophones in Montreal, whom he asked for directions either in English or in French. When addressed in French, 30% of Anglophones responded in English, which may could treated as an example of maintenance. Importantly, the usage of this strategy does not necessarily stem from an inability to adjust to different language varieties. According to Giles and Ogay (2007), maintenance can be employed as a method of asserting one's identity in a more unobtrusive manner.

1.3. Speech convergence in other socio-psychologically based studies

There exist a number of studies carried out outside the accommodation framework, which, nonetheless, are based on much the same principles and treat convergence as a socially motivated phenomenon. For instance, some interesting research on speech accommodation was conducted by Welkowitz and Feldstein (1969; 1970, in Welkowitz, 1972) and Welkowitz Finklestein, Feldstein and Aylesworth (1972). The studies are based on recordings of volunteers from a psychology course, who were divided into same-sex dyads after completing a set of personality tests. Some of the pairs of participants were told that the personality tests revealed that they were very similar to each other, some were informed that the tests showed that they were dissimilar, and some were told that they were randomly paired. The informants met three times in one-week intervals and talked to each other for an hour on each occasion. The independent variables under investigation were pause durations (Welkowitz and Feldstein, 1969; Welkowitz and Feldstein 1970, in Welkowitz, 1972) and vocal intensity (Welkowitz et al., 1972). The results indicated that the group of informants who believed themselves to have similar attitudes and personalities tended to accommodate to each other. Just as with the results of some of the research carried out within Communication Accommodation Theory, the observed tendency can be explained using the assumptions of

similarity attraction and illustrates that strengthening the sense of solidarity within a given pair of participants may induce them to converge towards their conversational partner. Welkowitz et al. (ibid.) and Welkowitz and Feldstein (ibid.) also found that convergence was not immediate and occurred only after a longer period of interpersonal contact, indicating that the magnitude of accommodation may be conditioned by the degree of familiarity with one's conversational partner.

Two interesting studies on convergence of vocal intensity and temporal patterns were also carried out by Natale (1975a; 1975b). The study concerned with shifts in vocal intensity (Natale, 1975a) consisted of two experiments. The subjects in the first experiment (male students recruited from Ohio University) conversed with an interviewer whose vocal intensity was experimentally controlled and fluctuated between different levels of loudness. The interviewer and the interviewees were seated in separate booths and could hear each other through speakers. The topic of the conversations was fixed and the interview was structured. Natale found that participants' vocal intensity increased as the interviewer's voice grew louder, supporting the hypothesis that the informants would converge towards their conversational partner. The second experiment consisted of seating same-sex dyads (25 females and 25 males recruited from Ohio University) on opposite sides of a curtain (so that they would not see each other) and asking them to converse freely on a topic of their own choosing. Convergence was defined as a reduction in the difference between the mean vocal intensities of the participants in a given dyad. Prior to the conversation task, the participants completed the Marlowe-Crowne Social Desirability Scale (Marlowe and Crowne, 1961, in Natale, 1975a), which gauges the degree to which an individual feels the need for social approval. The test was used to verify the hypothesis that subjects who seek social acceptance are more likely to accommodate towards their interlocutor. The results of the study showed that the informants who scored higher on the Marlowe-Crowne scale converged in loudness to a greater extent than those who obtained lower scores in the test. Interestingly, the results also indicated that several informants exhibited no convergent behaviour (both in the second and the first experiment). As argued by Natale (ibid.), the findings of the study support the idea that convergence of non-content speech behaviour is prompted by a desire for effective communication and intelligibility. The author also suggests that "differences in an individual's empathy, rapport, social desirability, or other relevant personal characteristics are related to the degree of non-content speech convergence in various dyads" (Natale, 1975a: 801). Similarly as in the previously cited studies (Welkowitz and Feldstein, 1969, 1970, in

Welkowitz, 1972; Welkowitz et al., 1972), Natale observed that the amount of convergence on the part of the participants increased over time.

The findings concerning accommodation of non-content speech behaviour were confirmed in Natale's (1975b) second study, which focused on speech convergence on pause duration. Twenty six participants recruited from an introductory psychology course were paired into same-sex dyads and instructed to converse freely for thirty minutes. The informants were in full view of each other and the conversations took place twice, in an interval of one week. The subjects took the Marlowe-Crowne Social Desirability Scale test before participating in the speaking tasks. Similarly as in the first study (Natale, 1975a), the magnitude of convergence corresponded to a participant's score on the social desirability scale, indicating that the extent to which and individual feels the need for social approval may have an impact on phonetic imitation. However, only the data from the second conversation task yielded significant results, validating the previous findings that the amount of convergence might be a function of the length of interpersonal contact.

Another study concerned with the imitation of non-content speech behaviour was carried out by Gregory and Hoyt (1982), who investigated the accommodation of vocal intensity, pauses and utterance frequency. The subjects were five airmen who participated in informal interviews designed to elicit their views on race relations in the air force. The decibel level, the frequency of sound events and the silent period location and duration obtained from the interviews were analysed using Fourier series. Fragments of the participants' and the interviewer's productions were contrasted with each other to create actual paired matches (i.e. the real conversation pairs) and virtual paired matches (i.e. conversations that did not occur). If the actual pairs had a better match value than the virtual pairs, the participants' behaviour in a given conversation was treated as convergence. The findings of the study suggested that the informants accommodated towards each other. In addition, some participants were found to converge to a greater extent than others. A more detailed analysis of the situational context of the interviews and the subjects' background revealed that "cultural homogeneity" might have facilitated accommodative behaviour between some of the conversational partners. Once again, the effect seems to be related to the notion of similarity attraction and points to the importance of a shared sense of solidarity in speech convergence.

In a popular study on convergence patterns between men and women, Bilous and Krauss (1988, in Pardo, 2010) sought to challenge the stereotype that men dominate conversational interactions. The experiment involved recording conversations between same-sex and mixed-sex dyads. Accommodation was measured by comparing a given participant's speech

produced in the same-sex condition with the speech produced by him/her in the mixed-sex condition. The dependent variables under investigation were the total number of words, average utterance length, frequency of interruption, frequency of short pauses, frequency of long pauses, frequency of back-channel responses and frequency of laughter. The obtained data yielded quite complex results. For example, the researchers discovered that female participants converged to male partners in the total number of words and the frequency of interruptions, while men converged to female partners in the frequency of back channels and the frequency of laughter. It was also observed that women diverged from men in the frequency of back channels and the frequency of laughter, whereas male participants did not diverge from women in any of the investigated features. Another finding was that both male and female participants accommodated towards their partners in average utterance length and frequency of short and long pauses. Thus, the results of the study suggest that there might exist some correlation between gender and imitation strategies, although the relationship appears not to be entirely straightforward. As argued by Bilous and Krauss (1988, in Pardo, 2010: 185):

> Any generalizations about the ways that men and women accommodate to each other when they interact must take into account the relevant properties of the situation in which the interaction takes place and the goals of the participants in those situations.

The effect of gender on speech accommodation was also touched upon by Hannah and Murachver (1999), who found subtle differences in the way female and male participants responded to the speech of their interlocutors. The speech-related variables under investigation were the amount of speaking time, frequency of interruptions, and frequency of back-channels.

More recently, Pardo (2006) examined convergence between talkers in conversational interactions using perceptual judgements from an “AXB task”, a tool which was first used to study phonetic convergence in a seminal study by Goldinger (1998). In an AXB task, a group of listeners judges the similarity of speech samples in order to gauge the amount of convergence. Sets of three audio samples are presented to a listener on each trial. X is the sample taken from the interlocutor (called the model talker in laboratory studies on convergence). A given participant’s samples are presented as A and B, one is the sample produced before exposure to the interlocutor’s speech, the other is the sample produced after or during exposure to the interlocutor’s speech. The listener’s task is to rate which sample, A

or B, sounds like a better imitation of X. The participants in Pardo's (ibid.) study were 6 male and 6 female native speakers of American English who were grouped into same-sex dyads and instructed to complete a map task (Anderson et al., 1991). The informants were also required to read a list of landmark label phrases from the map task in carrier phrases prior to and following the conversational task. The subjects' productions of the landmark label phrases under the three conditions were contrasted with speech samples taken from their conversational partners and presented to a separate group of listeners in the AXB task. The listeners were instructed to focus solely on the pronunciation of the informants, i.e. "[...] the way that the talkers were articulating the consonants and vowels [...] (Pardo, 2006: 2384)". This was done in order to draw the listeners' attention away from nonphonetic features of the participants' speech. Statistical analysis of the data showed that the subjects accommodated towards their partners during conversational interaction. Convergence was found to start early in the conversation, increase over the course of the interaction and persist until the post-map task reading. Pardo (ibid.) also found that the talker's role in the interaction (either giving or receiving instructions in the map task) and the speaker's gender had an impact on convergence patterns. Overall, men accommodated more than women and givers accommodated more than receivers. However, givers were found to converge to receivers in the female dyads, while in male pairs the opposite pattern was observed. Pardo's (ibid.) findings suggest that the speaker's role in an interactions may have a bearing on the magnitude of convergence. Her results corroborate Bilous and Krauss's (1988) findings in the sense that although gender appears to affect accommodation, the effect seems to be rather complex.

Pardo (2010) re-examined the results of her 2006 study using acoustic measurements to establish which phonetic features might have contributed to listeners' judgements of convergence. The variables under investigation were utterance duration and F0 of landmark label phrases from different stages of the experiment and vowel spectra in hVt words, which were included in the pre-task and post-task reading phase of the experiment. Statistical analysis revealed that perceived convergence was weakly related to pitch and speaking rate and that the correlation was only present for female pairs of talkers. As regards vowel quality, statistical analysis of the data suggested that participants converged towards their conversational partners in the realisation of high vowels. Interestingly, it was also found that givers diverged from receivers in the realisation of low vowels. The results imply that linguistic factors may also play a significant role in convergence, as the magnitude of the observed accommodation depended on a given pronunciation feature. Pardo's (ibid.) findings

seem to accord with the observation that power relations have an important bearing on convergence (Giles et al., 1991).

Pardo et al. (2013) delved into the issue of how a given participant's role in an interaction influences the direction and extent of convergence. The participants, 4 female and 4 male pairs of native English speakers, were instructed to complete a number of map tasks. The subjects switched roles from giving to receiving instructions several times during the experiment. The study was concerned with the duration of filled and unfilled pauses, articulation rate and the total time talking, which were analysed using acoustic measurements. Speech samples taken from the interacting participants were also presented to a separate group of listeners in an AXB task. The results of the study confirmed the previous findings that a talker's role in a conversation may have a bearing on his/her accommodation strategies, although the relationship proved to be quite complex. Pardo et al. (ibid.) also discovered that the subjects' initial roles in the interaction affected their subsequent accommodative behaviour. More specifically, it was found that participants who originally acted in the role of givers dominated in the amount of time spent talking even when they switched roles to givers.

Pardo et al. (2012) examined convergence between pairs of talkers following long-term exposure to each other's speech patterns. The participants were 5 pairs of previously unacquainted college roommates (all native speakers of American English), whose pronunciation was analysed at four different periods in time: before exposure to each others' speech, after a 1.5 month period of cohabitation, after a 3.5 month period of cohabitation and after a 4.5 month period of cohabitation. The subjects provided American English vowels embedded in hVd/t words in carrier phrases and read two sentences which included phonetic features that exhibit variation across US dialect regions. A few key phrases were extracted from the sentence recordings and presented to a separate group of listeners (30 native speakers of American English) in an AXB classification task. In addition to perceptual similarity judgments, Pardo et al. (ibid.) collected acoustic measurements of item duration and vowel quality. The participants were also required to complete a survey designed to evaluate the strength of their relationship. Statistical analysis of the data suggested that the informants converged towards each other to some extent. Nonetheless, accommodation patterns were found to be rather variable, both across different pairs of talkers and across different utterances. Pardo et al. (2012: 196) remark that the findings point to the possibility that "[...] each individual talker might converge on a unique set of acoustic-phonetic attributes while diverging, varying randomly, or remaining neutral on others." The researchers also report that the magnitude of convergence was moderately related to reported closeness

between the pairs of participants, thus providing some evidence for the claims that increased familiarity and a shared feeling of solidarity between a pair of talkers may bring about greater levels of accommodation.

Following Bilous and Krauss (1988), Schweitzer and Lewandowski (2012) focused on the accommodation of back-channel responses. The subjects were 8 female speakers who participated in spontaneous conversations with 6 different female interlocutors and talked on topics of their own choice. All subjects were native speakers of German, who conversed with each other in their native language (N. Lewandowski, personal communication, January 7, 2014). Following each conversation, the subjects were required to rate their partners' likeability and competence. Statistical analysis of the results showed that the speakers accommodated the frequency of backchannels towards their interlocutors (however, the selected statistical model did not indicate whether the participants converged or diverged). Schweitzer and Lewandowski (ibid.) report that the participants used backchannels more frequently when they found their interlocutors to be competent or friendly, which seems to lend further support to the claim that social factors contribute to the magnitude of speech convergence.

The following two studies do not focus specifically on the social or psychological aspects of pronunciation shifts. However, they provide some interesting insights into the parameters that govern speech convergence in communicative interactions. Kim, Horton and Bradlow (2011) examined the effect of language distance on phonetic accommodation. The participants in the experiment were paired according to their native language and dialect. Eight pairs of informants comprised two speakers with the same L1 (four native English pairs and four native Korean pairs). Half of the native-native pairs shared the same variety of L1, half used different dialects. Another eight pairs of participants comprised two talkers with different L1s. These were native speakers of English conversing either with a native speaker of Korean or a native speaker of Chinese. Each member of a pair received a picture, which differed slightly from the picture given to his/her conversational partner (diapix task, Van Engen et al., 2010). The subjects' task was to talk to each other in order to find all of the differences. The degree of convergence was measured with the use of an AXB perceptual similarity task that was completed by a separate group of listeners. The A and B stimuli were speech samples of a given member of the pair taken from early and late portions of the recorded conversations. The stimuli were contrasted with the interlocutor's speech sample from either early or late stage of the interaction (X). The data suggested that it was only the pairs of speakers with the same L1 who accommodated towards each other. The finding could

imply that increased language distance between conversational partners inhibits speech convergence. Kim et al. (ibid.) ascribe the apparent lack of accommodation in the case of the remaining pairs of talkers to high attentional demands and processing load involved in cross-dialect and native-non-native communication. Nevertheless, one needs to bear in mind that the audio samples presented in the AXB task contained different words. It is quite possible that the participants converged on some fine-grained phonetic features that were not present in the selected samples or were present only in some of them.

Lewandowski (2012) examined convergence in spoken interactions between native and non-native speakers of the same language. The participants, two native English speakers and twenty native speakers of German, were required to complete a diapix task. Although the native English speakers were specifically asked not to converge their pronunciation towards their partners, acoustic analysis of amplitude in target words extracted from the conversations revealed that the subjects accommodated towards their German partners. When asked about it after the experiment, the native speakers stated that they felt they managed not to shift their pronunciation towards the German talkers. The results stand in contrast with the assumption that a desire to modify social distance or increase the effectiveness of communication constitute the primary reasons for the occurrence of speech convergence. Instead, Lewandowski's findings suggest that imitating the speech of one's conversational partner is to some extent an automatic tendency that may take place irrespective of the speaker's conscious decisions.

1.4. Speech convergence in laboratory settings

Some of the more recent research on phonetic convergence is conducted in socially-minimal, laboratory settings and views the phenomenon as an automatic reflex of the human brain rather than a process predetermined by social-psychologically factors. For instance, the influential study by Goldinger (1998) is concerned with phonetic convergence not so much as the object of the investigation but as a tool for testing a hypothesis about spoken word representation, perception and production. The author advocates an exemplar model of speech perception and supports his claims by using data from a series of experiments in which the participants listened to and then listened and repeated pre-recorded realisations of single words. Prior to the listening and imitation blocks of the experiments, the subjects were required to read all of the investigated words so that their baseline productions could be

recorded (the term baseline production refers to a speaker's 'regular' pronunciation of a given item, i.e., the realisation that is not affected by exposure to another talker's speech). In the listening trials, the words were presented with different levels of repetition (from zero up to twelve repetitions). The next phase included the shadowing (i.e. imitation) task, which was subdivided into immediate and delayed shadowing. The former consisted of repeating the words immediately after hearing them, while the latter involved waiting 3-4 seconds before speaking. The stimuli heard in the listening and shadowing trials varied in terms of frequency of occurrence, ranging from high frequency to low frequency words. To gauge the amount of imitation, a different group of participants completed the AXB classification task. As referred to in the previous section, the AXB task is a perceptual test in which a group of listeners evaluates the similarity between the productions of a given informant and the productions of the model talker (the person whose voice in being imitated). In Goldinger's study, participants' responses to the AXB classification task indicated that imitation was more likely to occur in the immediate shadowing condition as compared with the delayed shadowing condition and that it increased with the number of repetitions. It was also discovered that low frequency words were imitated to a greater extent than high frequency words. The effects of word frequency and the number of repetitions on the magnitude of imitation were confirmed in a study by Goldinger and Azuma (2004). The major difference between the study by Goldinger (1998) and the study by Goldinger and Azuma (2004) was that in the latter, the participants did not shadow the audio stimuli immediately but were asked to produce them a week after exposure to the model talker's speech.

Taken together, the findings of Goldinger (1998) and Goldinger and Azuma (2004) suggest that phonetic imitation is to some extent an automatic cognitive reflex and that social motivations are not a prerequisite for some degree of imitation to occur. These observations seem to be in tune with the findings of Lewandowski (2012) and Chartrand and Bargh (1999). The latter examined facial expressions and mannerisms in pairs of interacting partners and found evidence that imitation effects in humans may often be automatic and unintentional. They also noted that merely perceiving an action being performed by another may induce one to behave in a similar manner. Kuhl and Meltzoff (1996) observed that infants as young as 12 weeks of age tend to imitate speech, providing more support for the claim that imitation is, to some extent, a natural and automatic process in humans. A similar view is espoused by Pickering and Garrod (2004), whose interactive alignment account proposes that "[...] in dialogue, the linguistic representations employed by the interlocutors become aligned at many levels, as a result of a largely automatic process (Pickering and Garrod, 2004: 169).

Inspired by Goldinger's (1998) research, Shockley, Sabadini and Fowler (2004) carried out two similar experiments. In the first experiment, the participants (8 undergraduate students) listened to pre-recorded single words produced by the model talkers (2 graduate students) and were instructed to "identify the word [they heard] by speaking it into the microphone quickly but clearly (Shockley et al., 2004: 424)". The tokens under investigation were 80 bisyllabic English words beginning with the voiceless stops /p, t, k/. The shadowed productions were contrasted with the subjects' productions from the baseline condition (where the participants were asked to read the investigated words from a computer screen) in an AXB task. The data collected in the first experiment revealed that shadowed words were rated as better imitations of the model talker's productions as compared with the baseline condition, which implies that the informants imitated the pronunciation of the model talkers and corroborates Goldinger's (1998) findings. However, Shockley et al. (2004) did not find evidence for Goldinger's claim that phonetic convergence increases with the number of repetitions. It was found that the number of prior exposures to the model talker's productions of a given word (zero vs. six) did not significantly affect the magnitude of perceived imitation.

The second experiment carried out by Shockley et al. was designed to expand on Goldinger's (1998) findings and attempted to evaluate which phonetic features are being imitated in a shadowing task. The same procedure was used as in the first experiment, the difference being that the VOTs in the model talker's productions were extended to twice their original duration. In addition to examining the listeners' judgements from an AXB task, the researchers analysed the participants' VOT duration in the shadowed and baseline productions using acoustic measurements. Statistical analysis of the data revealed that shadowed productions were reported to be better imitations of the model talker's productions as compared with the baseline condition. More importantly, it was also found that the subjects increased VOT duration upon exposure to the extended VOTs in the model talkers' speech. Similarly as in Goldinger's (1998) study, the results of the two experiments indicate that phonetic imitation is to some degree a spontaneous and automatic process.

A study concerned specifically with the automaticity of imitative behaviour was carried out by Delvaux and Soquet (2007), who argue that speakers tend to converge towards ambient speech automatically and unintentionally. The study comprises two experiments, the first of which was conducted on four native speakers of Belgian French. Two of the participants were representatives of Liège regiolect, the other two used a Brussels variety of French. First, the participants saw numbers and ideograms on a computer screen and were instructed to name them within carrier sentences. Next, the subjects were asked to perform

the same auditory naming task for the second time. In this task, they could also hear model talkers' voices name the numbers and ideograms over loudspeakers. Native speakers of the Liège regiolect could hear a model talker who used the Brussels variety and vice versa. Interestingly, the informants were never explicitly instructed to repeat or imitate what they heard, nor were they required to listen to the model talkers' voices. The dependent variables under investigation were the spectral and durational characteristics of /o/ and the duration of /i/, which differ between the two regiolects. The results revealed statistically significant convergence towards the model talkers' speech.

The second experiment followed the procedure of Experiment 1. The participants were eight native speakers of the Mons regiolect of Belgian French, who could hear the voice of a Liège French user in the second block of the experiment. The investigated dependent variables were the length of /ɛ/ and the quality of /o/, which differ across the two varieties. A post-test condition was added in Experiment 2, in which the subjects were instructed to perform the naming task for the third time, without the model talker's voice. It was found that the informants tended to converge towards the voice they could hear in the second block of the experiment and that the effect of exposure to the model talker's speech persisted until the post-test task. Delvaux and Soquet also report that when asked about it after the experiment, the participants were not aware of having imitated the model talkers. The findings of both Experiment 1 and 2 suggest that imitation is an unintentional and automatic process and accord with the data collected by Goldinger (1998) and Shockley et al. (2004). Moreover, the imitative effect appears to reduce gradually. Delvaux and Soquet's results imply that listening to the model talker's speech may leave a memory trace which affects the speaker's productions up to several minutes after exposure. This observation corroborates Pardo's (2006) and Goldinger and Azuma's (2004) findings that the effect of convergence may be carried over to the speech produced several minutes or several days after an interaction (or exposure to the model talker's speech in the case of Goldinger and Azuma).

Another study that employs phonetic imitation to examine the link between perception and production was conducted by Mitterer and Ernestus (2008). The participants were 18 native speakers of Dutch who were required to complete a shadowing task in which they repeated various Dutch nonwords. Some of the investigated nonwords contained initial stops with different degrees of prevoicing, while others included two variants of /r/ (alveolar and uvular), which occur as free allophones in Dutch. Mitterer and Ernestus found that the majority of the participants did not systematically imitate the /r/-stimuli and used their

habitual variant of this sound in most cases. As regards prevoicing in initial stops, the data showed that "[t]he phonologically relevant difference between presence and absence of pre-voicing was imitated, while the phonologically irrelevant amount of pre-voicing was not" (Mitterer and Ernestus, 2008: 171). The findings of the study indicate that the automaticity of phonetic convergence may be restricted by the linguistic structure of the language in question. More specifically, it would appear that it is the phonologically relevant pronunciation features that are more likely to be imitated.

Some interesting research on phonetic imitation that also makes reference to speech perception theories was carried out by Nielsen (2011). The study consists of two experiments and examines the imitation of reduced and extended VOT values by 25 native speakers of American English. The first experiment was divided into three stages: baseline recording, target exposure (listening) and post-exposure recording. In the first stage, the informants were asked to read a list of 150 words. 30 of these were filler words, 100 were words with an initial /p/ consonant, while the remaining 20 were words beginning with /k/. In the next stage of the experiment, the participants listened to the model talker's (also a native speaker of American English) realisations of the investigated lexical items with artificially extended VOT values. 20 of the /p/-initial and all of the /k/-initial words were purposefully excluded from the listening phase. In the final stage of the experiment, the subjects were required to read the word list from the baseline recording block for the second time. Statistical analysis of the data showed that the informants converged their pronunciation towards the model talker's extended VOTs even though the post-exposure phase took place several minutes after the listening task, thus validating Delvaux and Soquet's (2007) claims that exposure to the model talker's pronunciation may cause imitation effects even when the production is delayed. Interestingly, Nielsen discovered that VOT values increased also in the subjects' productions of the /p/- and /k/-initial words that were not included in the listening block. This signifies that phonetic imitation may operate both below word and phoneme level. Finally, the results of the first experiment indicated that lexical frequency had a significant effect on the magnitude of imitation, thus lending support to Goldinger's (1998) and Goldinger and Azuma's (2004) finding that low frequency words are imitated to a greater extent than high frequency words.

Experiment 2 of Nielsen's study follows the procedure of Experiment 1 with the exception that VOT values of the model talker were reduced, not extended. Contrary to the previously obtained results, statistical analysis of the data revealed that the participants did not imitate the reduced VOT values that were present in the stimulus. As argued by Nielsen (2011: 139),

the observed discrepancy between the participants' imitative behaviour in the two experiments may be explained by the fact that "[...] imitating reduced VOT can introduce phonological ambiguity with the voiced stop, while there are no such consequences in imitating extended VOT". Both Nielsen's and Mitterer and Ernestus's (2008) findings imply that phonetic imitation in a laboratory setting is not an entirely automatic process, as it appears to be sensitive to linguistic structure. In addition, Nielsen reports that her data exhibited a wide range of variability in the degree of imitation, indicating that individual speaker differences have an impact on phonetic convergence.

Honorof, Weihing and Fowler (2011) conducted experiments on phonetic imitation whose purpose was to assess the validity of several competing theories of speech perception. The three experiments were all concerned with the imitation of 'dark' and 'light' allophones of /l/ by native speakers of American English and followed the same experimental procedure. The subjects listened to a model talker (also an American English speaker) pronounce nonsensical V.CV sequences containing [l], [ɫ], /ɹ/ and /w/ and were instructed to repeat what they hear as quickly as possible. In the first experiment, the model talker produced the /l/ allophones in a manner typical of his native accent. In the second experiment, the 'lightness' and 'darkness' of the /l/ stimuli were enhanced in order to intensify the perceptible difference between the two variants. As explained by Honorof et al. (2011: 24):

> [...] the model's goal was to de-emphasize the retraction of the tongue body for [l] tokens to make them sound 'lighter' than the [l]s from Experiment 1. For the [ɫ] variant, the model's goal was to de-emphasize the tongue-tip gesture while nonetheless retracting the post-dorsal region of the tongue midline into the oropharynx, without making medial contact with the rear wall of the pharynx [...]

Experiment 3 differed from the previous two in that it included magnetometric analysis, which was employed to examine the participants' articulation. Statistical analysis of the data showed that although the subjects displayed a tendency to imitate the modelled speech in all three experiments, the magnitude of imitation was small. In other words, the acoustic difference between /l/ variants in the informants' pronunciation was never close to the difference exhibited by the model talker. Honorof et al. (2011: 24) offer a plausible explanation for these results:

> We ascribe this pattern to two competing tendencies. One is the disposition to imitate (even without being instructed to do so explicitly) [...]; the second is the tendency to persist in habitual ways of producing phonetic segments.

Similarly as in some of the previously cited research, Honorof et al. (ibid.) report that some participants did not converge towards the model talker, whereas others converged towards the model to very different degrees. Overall, their findings seem to provide further evidence for the notion of phonetic imitation being to some extent an automatic process. At the same time, their results suggest that the magnitude of convergence may be inhibited by individual speech habits and other individual speaker differences.

Brouwer, Mitterer and Huettig (2010) set out to investigate whether speakers imitate reduced speech in a shadowing task. The speech stimuli were sentences extracted from the Spoken Dutch Corpus (Oostdijk, 2000). Each sentence contained one target word; half of the target words were produced canonically, half were produced in a reduced form. The participants, 16 native speakers of Dutch, were instructed to listen to the sentences and repeat them as fast as possible. The variables under investigation were the duration and segment realisation of the target words. The results showed that the participants converged on the duration of both canonical and reduced forms, however, the effect was relatively weak. More specifically, the difference in duration between the canonical production and the reduced production was greater in the stimuli than in the shadowed responses. It was also found that the subjects imitated duration more closely in the case of the canonical forms. As regards the segmental realisation of the target words, Brouwer et al. (ibid.) report that the subjects mostly used segments characteristic of canonical forms, even when presented with the reduced realisations. On the whole, the results appear to support the claim that phonetic imitation may be susceptible to language structure.

Kim (2011) concentrated on phonetic convergence in native speakers of English after exposure to native and non-native speech. In the baseline condition, the participants were instructed to read two sets of words, one containing words beginning with bilabial stops and one comprising words with initial alveolar stops. In the exposure condition, the subjects heard the target words from one of the sets realised by the model talkers (a native speaker of American English and a native speaker of Korean). On each trial, the subjects could see a number of English words displayed on the computer screen and their task was to identify which of the words was produced by the model talker. The participants read both sets of words again in the post-exposure condition. The phonetic variable under investigation was

CV duration, which was considerably smaller for the non-native model talker than for the native model talker. The results showed that the informants maintained their baseline CV durations after exposure to the American English speaker and reduced CV length after listening to the Korean speaker. The findings indicate that the subjects converged only towards the non-native speaker, which contradicts Kim et al.'s (2011) observation that smaller language distance facilitates accommodation. This could be explained by the fact that Kim et al. (2011) used a considerably different experimental procedure than Kim (2011). Also, it is highly likely that the speakers in Kim's (2011) study did not imitate the length of the native model talker's CV sequences because they already exhibited similar duration values in the baseline condition. In other words, it is possible that the informants had no room to accommodate in the case of the native American speaker. Kim also reports that the convergence effect was carried over to the realisation of the word set that the participants did not hear during exposure, which accords with Nielsen's (2011) findings. On the whole, Kim's (2011) results support the notion that speakers are naturally predisposed to imitate the speech they are exposed to.

1.5. Laboratory-based convergence with social-psychological motivations

In the studies discussed in this section, phonetic convergence is viewed from a social-psychological standpoint. At the same time, the process is examined in settings where social interaction is severely limited. Thus, the research discussed in this section combines laboratory-based methodology with the theoretical approaches first adopted by Giles and colleagues in Communication Accommodation Theory. One of such studies was carried out by Namy, Nygaard and Sauerteig (2002), who sought to expand on Bilous and Krauss's (1988) observations that gender differences may have some bearing on convergence and divergence patterns. A group of 8 men and 8 women (native speakers of American English) were asked to take part in a shadowing task in which they repeated the speech of two other female and two other male model talkers (native speakers of American English). The words produced by the participants in the shadowing task were contrasted with the productions from the baseline condition in an AXB task where 32 female and 32 male listeners judged which of these two productions sounded more like a given model talker's pronunciation. The data revealed that women were more likely to converge than men and that the subjects accommodated more to male than female model talkers. It was also discovered that the latter

tendency might have been driven by one particular male model talker, since the informants were found to accommodate more to him than to any other speaker. These findings seem to substantiate the claim that gender differences may affect phonetic convergence. More importantly, they suggest that social factors may moderate the magnitude of convergence even in the case of very restricted social interaction. In addition, the results obtained by Namy et al. (ibid.) imply that some voices evoke more imitation than others.

Babel (2009) set out to examine the social and linguistic factors that affect the magnitude of phonetic convergence in a laboratory setting. The stimuli for imitation, /i æ ɑ o u/ embedded in fifty monosyllabic English words, were provided by two male native speakers of California English, one African American, one Caucasian American. As many as 178 informants took part in a shadowing task and were assigned to one of four conditions. One group of participants was presented only with the model talker's voice, while the other group could also see a still digital image of the model talker that they were listening to. The two groups were further subdivided into informants who were assigned to the white model talker and those who listened to the black model talker. The group of participants who were exposed to the visual stimuli were also required to rate the attractiveness of the model talker they heard. All subjects completed an Implicit Association Task (Greenwald, McGhee and Schwartz, 1998) in order to measure their implicit racial bias. The results of the study revealed that the participants imitated /æ ɑ/ to a greater extent than /i o u/, lending support to the claim that phonetic convergence may be selective from a linguistic perspective. Babel also discovered that the subjects who were assigned to the visual stimuli condition and scored with a pro-black bias on the Implicit Association Task were more likely to converge towards the black model talker. It also transpired that the model talkers' attractiveness ratings had some impact on the magnitude of phonetic imitation. The more attractive a given model talker was considered, the more the female participants were likely to converge. In case of the male participants, an opposite trend was observed. Taken together, Babel's findings suggest that phonetic imitation may be mediated both by linguistic and social factors. The results support Namy et al.'s (2002) findings that some degree of socially motivated convergence may take place even in socially minimal settings.

Babel (2010) aimed to replicate the study by Bourhis and Giles (1977) in a laboratory setting. The subjects were 44 native speakers of New Zealand English, who participated in a shadowing task. The stimuli were monosyllabic English words containing KIT, DRESS, TRAP, START, STRUT and THOUGHT vowels produced by a native Australian English

speaker. Prior to production, the participants were asked to complete the Implicit Association Task, which was intended to gauge their inherent bias towards Australia. The informants were also divided into two groups. One group was presented with a text designed to dispose them favourably towards the Australian model talker and Australia as a whole. The other group read a text whose purpose was to make them look upon the model talker and Australia in a negative light. Following the assumptions of Communication Accommodation Theory, Babel (ibid.) hypothesised that participants who expressed positive feelings towards Australia would converge towards the model talker in order to decrease social distance. Overall, the data showed that the subjects imitated the speech of the native Australian English speaker. Similarly as in Babel's previous study (Babel, 2009), it was found that participants who exhibited a pro-Australia bias were more likely to converge towards the model talker. The texts designed to affect the participants' feelings towards Australia and the model talker, on the other hand, did not have a bearing the magnitude of imitation.

Babel and Bulatov (2012) examined the imitation of fundamental frequency by native speakers of American English. Two groups of participants completed a shadowing task in which they were exposed to the speech of a male model talker. One of the groups listened to stimuli which had been high-pass filtered at 300 Hz (thus eliminating the fundamental frequency), while the second group listen to unaltered speech. The magnitude of convergence was measured both acoustically and using an AXB similarity task. Acoustic analysis of the data revealed that the subjects converged towards the model talker when the speech signal was unaltered and tended to diverge from the model talker when they heard the filtered speech. Similarity judgements obtained in the AXB task corroborated the results of acoustic analysis. However, further statistical analysis showed no significant correlation between the perceptual judgements of convergence and f0 measurements. As succinctly put by Babel and Bulatov (2012: 16), "[t]hese results suggest that there is not one single feature that serves as the only, or even as the primary, imitable feature". Interestingly, as opposed to Namy et al.'s (2002) observations, acoustic analysis of the data indicated that male participants accommodated to a greater extent than female participants.

Babel et al. (2012) focused on how perceived attractiveness and perceived typicality of a model talker's voice influence phonetic imitation. Drawing on Goldinger's (1998) findings about the effect of word frequency on convergence, Babel et al. predicted that unique voices would be imitated to a greater extent than typical voices. In the first stage of the experiment, 15 monosyllabic words with /i *a* u/ were produced by 30 male and 30 female native speakers of American English. The productions were presented to a group of 30 listeners (also native

speakers of American English), whose task was to rate the voices' attractiveness and typicality. The most attractive, unattractive, typical and atypical voices served as model talkers in a shadowing task (8 voices in total, 4 for each gender). In the shadowing task, 20 subjects (10 males and 10 females) produced baseline tokens of the 15 investigated words and then shadowed the model talkers' realisations of these words. Phonetic convergence was measured by analysing similarity judgements of a separate group of 35 listeners who completed an AXB task. The results of the study indicated that the participants imitated all of the model talkers. The listeners in the AXB task perceived the greatest amount of imitation in the case of the least typical male model talker. It was also found that among the female model talkers, the most attractive female voice induced most convergence. Finally, it was discovered that the participants imitated words containing /u/ more than words containing the other two investigated vowels. On the whole, the results of all the cited studies by Babel and colleagues suggest that phonetic imitation in a laboratory setting should not be viewed solely as an automatic reflex of the language system. It would appear that sociolinguistic factors such as inherent social biases or perceived attractiveness can also play an important role in phonetic convergence, even in circumstances where no apparent social interaction can be found. Finally, the data obtained by Babel and colleagues support the previous findings that the magnitude of imitation may be constrained by language-internal factors.

Further support for the observation that speech accommodation may be mediated by social-psychological factors in contexts with restricted interpersonal interaction can be found in the study by Yu, Abrego-Collier and Sonderegger (2013). Over 80 participants produced several dozen /p t k/-initial English words in baseline and post-exposure blocks of the experiment. In the exposure phase, the subjects heard a first-person narrative read out by the model talker. The narrative contained the /p t k/-initial words, whose VOT values were artificially extended by 100%. The subjects were assigned to one of four different conditions. One group of informants heard a narrative designed to dispose them favourably towards the model talker, the other group heard a narrative whose purpose was to make them view the model talker in a negative light. The two narratives were further subdivided into one where the model talker appeared to be heterosexual and one where the model talker appeared to be homosexual. Following the experiment, the participants completed a battery of tests and surveys devised to shed some light on their attitude towards the model talker and their sexual orientation as well as to examine their neurocognitive abilities and personality traits. Statistical analysis of the results revealed that participants who expressed positive feelings towards the model talker extended their VOT values more than those who regarded the model

talker with disfavour. In contrast, Yu et al. (ibid.) did not observe significant effects of speaker gender or perceived sexual orientation of the model talker on the magnitude of convergence. It was also discovered that subjects who obtained high scores on openness and attention focus were more likely to imitate the extended VOTs than those with the opposite traits. Interestingly, Yu et al. (ibid.) report no overall imitation effect, which provides further evidence for the observation that speech convergence is highly variable and related to individual speaker-differences. At the same time, the results of the study run counter to Nielsen's (2011) findings about extended VOT imitation. Yu et al. (2013: 11) offer a likely explanation for the observed discrepancy, which also underscores the impact of experimental design on the magnitude of convergence,

> [...] [T]he exposure material in Nielsen's study were English words presented in isolation, while our exposure materials were embedded in a meaningful narrative. The marked difference in experimental results might be partly attributable to the decontextualization of the exposure materials in Nielsen's study; imitation might be more automatic (i.e., they can occur without the speaker's intention or control) in a context where the words are presented in isolation devoid of social significance. The narrative in the present study, in contrast, allows participants to make evaluative judgements on the narrator [...].

Another interesting observation made by Yu et al. is that native speakers of English may identify /t/ with extended VOT values with a number of personal qualities such as articulateness, elegance or prissiness. As argued by the authors, "[w]hile the indexical meanings associated with released /t/ are not intrinsically positive or negative, some subjects might nonetheless resist extending their VOTs in order to avoid projecting an articulate persona (Yu et al., 2013: 11)." Thus, it seems perfectly possible that the social meaning associated with a given phonetic variable may constitute another factor that modulates potential convergence or divergence.

1.6. Summary

Previous studies on speech convergence illustrate that speakers may tend to adjust a number of phonetic and quasi-phonetic variables following exposure to another person's speech. Participants have been found to modify non-content speech behaviour such as vocal intensity (Welkowitz et al., 1972; Natale, 1975a; Natale, 1975b; Gregory and Hoyt, 1982; Lewandowski, 2012), laughter (Bilous and Krauss, 1988) and back-channel responses (Bilous

and Krauss, 1988; Hannah and Murachver, 1999; Schweitzer and Lewandowski, 2012), temporal parameters such as pause duration (Gregory and Hoyt, 1982), speaking rate (Pardo, 2010), vowel duration (Delvaux and Soquet, 2007), word duration (Brouwer et al., 2010) and VOT (Shockley et al., 2004; Nielsen, 2011) as well as various other pronunciation features such as fundamental frequency (Gregory and Webster, 1996; Babel and Bulatov, 2012), vowel quality (Delvaux and Soquet, 2007; Babel, 2009; Babel, 2010; Pardo, 2010; Babel et al., 2012; Pardo et al., 2012), /r/ allophones (Mitterer and Ernestus, 2008), /l/ allophones (Honorof et al., 2011) and consonant elision (Coupland, 1984). The variables have been analysed with the use of both acoustic measurements (e.g. Delvaux and Soquet, 2007; Babel, 2009; Pardo, 2010; Nielsen, 2011; Babel and Bulatov, 2012) and perceptual judgements from naive listeners (e.g. Giles, 1973; Bourhis and Giles, 1977; Namy et al., 2002; Pardo, 2006; Babel and Bulatov, 2012). As regards the process of data collection, speech convergence has been analysed both in conversational interactions between pairs of talkers (e.g. Giles, 1973; Natale, 1975a, 1975b; Coupland, 1984; Gregory and Webster, 1996; Pardo, 2006) and by asking speakers to repeat single words or utterances after a pre-recorded voice (e.g. Goldinger, 1998; Namy et al., 2002; Shockley et al., 2004; Babel, 2009; Brouwer et al., 2010; Nielsen, 2011). Speech convergence has also been found to occur when speakers are not in a direct or face-to-face interaction with their interlocutors, but are simply aware of their presence (Bell, 1977; Douglas-Cowie, 1978; Bell, 1982a; Bell, 1982b; Bell, 1984).

The occurrence of speech convergence has been ascribed to different factors. In laboratory-based research, speech convergence is often studied with no reference to its social purpose and is simply viewed as an automatic and unintentional reflex of the human brain. Some evidence for this claim has been provided by Goldinger (1998), Goldinger and Azuma (2004), Shockley et al. (2004), Delvaux and Soquet (2007), Kim (2011) and Lewandowski (2012). Nonetheless, it seems worth pointing out that the obtained results may have been partly driven by the experimental procedure of choice. Speech convergence has mostly been found to be automatic in studies where participants are required to repeat single word productions. This type of experimental setting draws the speaker's attention to speech form and thus may result in more robust imitation effects. As observed by Yu et al. (2013), imitation may also seem more automatic when speech samples are presented in a context that is devoid of social significance. Finally, convergence is presumably more easily detectable (both acoustically and perceptually) in isolated, single word productions than in whole strands of conversation.

From a social-psychological perspective, convergence strategies serve as a tool for mediating social distance and/or facilitating communication in an interaction. This interpretation of the phenomenon has found empirical support in the studies by Welkowitz and Feldstein (1969), Giles (1973), Bourhis and Giles (1977), Gregory and Hoyt (1982), Coupland (1984), Bilous and Krauss (1988), Gregory and Webster (1996), Pardo (2006), Pardo (2012) and others. Importantly, a socially rich setting seems not to be a precondition for the occurrence of socially or psychologically based speech convergence. The data collected by Namy et al. (2002), Babel (2009), Babel (2010), Babel et al. (2012) and Yu et al. (2013) suggest that certain personality traits, social group membership and a speaker's attitude towards a given talker or a particular social group may affect the magnitude of phonetic convergence even in laboratory settings.

The automaticity of speech convergence is also called in question by the results of several studies which illustrate that imitation is sensitive to language structure (Mitterer and Ernestus, 2008; Babel, 2009; Babel, 2010; Brouwer et al., 2010; Honorof et al., 2011; Nielsen, 2011) and may be to some extent mediated by language distance (Kim et al., 2011). Additionally, both the research carried out in the laboratory and the studies which examine speech convergence in conversational interactions (e.g. Natale, 1975a; Mitterer and Ernestus, 2008; Pardo et al., 2012, Yu et al., 2013) suggest that individual speaker differences have considerable impact on the magnitude of speech convergence.

Contrary to the view advocated in the current study, it has been sometimes argued that imitation should be kept apart from accommodation, each notion representing speech behaviour generated by very different processes. However, taken together, the findings reviewed in this chapter indicate that the cognitive and the social-psychological aspects of convergence form a complex pattern of interaction. On the one hand, there appears to be a strong tendency for speakers to imitate the speech they are exposed to. On the other hand, the predisposition to converge appears to be constrained both by social and linguistic factors and is characterised by a high level of speaker- and context-related variability. In other words, previous findings suggest that speakers have an inherent tendency to imitate and that this tendency will either be impeded or reinforced by social-psychological and linguistic factors.

One could also argue that the findings of imitation- and accommodation-based studies should not be combined because of their use of radically different methodological tools. Admittedly, the experimental procedures used in the former may produce primed responses, while the methods employed in the latter result in more spontaneous and naturally-occurring speech behaviour. However, laboratory-based studies on phonetic convergence offer valuable

and detailed insights into the type of phonetic features that are more susceptible to imitation, the knowledge of which may be used to form predictions about accommodation patterns in more spontaneous interactions. Overall, it seems legitimate to argue that the seemingly conflicting approaches should not be viewed in terms of an either/or dichotomy but considered complementary (provided that one avoids sweeping generalisations).

Chapter 2: Formation of the L2 sound system

2.1. Introduction

The chapter provides an overview of the studies concerned with L2 pronunciation that are relevant to the current investigation; it describes the construction and development of the L2 sound system and the factors affecting L2 phonetic performance. The chapter is organised to reflect the evolution of research into L2 phonetics, the scientific field in which the current study is embedded. The emergence of L2 phonetics as a fully-fledged research area was a gradual process. The early studies on non-native pronunciation were predominantly concerned with uncovering the factors that enable successful second language acquisition and did not attempt to explain the internal processes responsible for attainment. The scope of L2 pronunciation research began to broaden with the introduction of the interlanguage model of second language acquisition in the 1970s (Selinker, 1972). The central assumption of the framework is that a speaker's productions in a second language are generated by an independent linguistic system that is separate from both the learner's L1 and L2. The establishment of interlanguage as an autonomous, self-governed linguistic system was accompanied by the emergence of novel approaches towards the study of L2 phonetic performance, which began to be treated as a research subject in its own right. The studies on non-native pronunciation gradually became more multidimensional, investigating different factors that contribute to the formation of the L2 sound system as well as the complex relationships between them. Due to an increasingly large number of studies that have been conducted on L2 pronunciation, the chapter focuses on the research considered particularly relevant to the present investigation.

The studies reported in this chapter are, for the most part, discussed in a chronological order; they are divided into sections on the basis of their subject matter and approach towards the study of pronunciation. First, the early studies on L2 pronunciation are discussed (Section 2.2.). It is important to note that the findings described in that section are referred to as early not only because of the time of the their publication but also because of their focus of interest and their approach towards the L2 sound system. The common denominator in these early studies is that their main interest lies in uncovering the factors that enhance or prevent the successful acquisition of foreign-language pronunciation. The next section (2.3.) describes the interlanguage model and its implications for the formation of the L2 sound system; as

opposed to early studies on L2 phonetics, this strand of research concentrates on describing and explaining the internal processes that are responsible for second-language acquisition. Section 2.4. reviews relevant sociolinguistic research on non-native pronunciation, which focuses on examining and explaining variability in non-native phonetic performance. What distinguishes the studies discussed in this section is the underlying assumption that interlanguage pronunciation shares many characteristics with the pronunciation of an L1 and can be treated and examined similarly. Section 2.5. surveys studies and frameworks concerned with learners' perception of L2 sounds and the relationship between perception and production in non-native speech; these studies adopt a psycholinguistic approach to the study of L2 pronunciation and are interested in how the brain processes L2 sounds. Section 2.6. is concerned with studies conducted in institutional or classroom settings which deal with the influence of pronunciation instruction on the acquisition of the L2 sound system. Although not a major research area within the field of L2 phonetics, the effect of conscious phonetics and phonology knowledge on the production of L2 sounds is relevant to the current study. Section 2.7. discusses previous work on learner attitudes towards L2 pronunciation. Unlike the majority of the studies described in this chapter, this particular strand of research does not examine actual phonetic performance of L2 learners. However, similarly as in the case of pronunciation instruction, it is included in the chapter because attitudinal factors were considered relevant to the current investigation. The final section of the chapter summarises the findings presented in this chapter.

2.2. Early studies on L2 pronunciation

The main purpose of the research reviewed in this section was to examine the factors that enhance or prevent successful acquisition of L2 pronunciation. One of the first factors associated with pronunciation accuracy that received scientific attention was the age at which the learning of a second language commences (often referred to as the age of L2 learning and abbreviated to AOL) and, more specifically, the existence of a critical period for human speech learning. The matter was studied by many researchers and derived mostly from the work of Lenneberg (1967), Scovel (1969) and others, who posited a correlation between the biological maturation of the brain and the ability to master the pronunciation of a foreign language. One of the first such studies was carried out by Asher and García (1969), who tested the hypothesis that children, as opposed to adult learners, are biologically predisposed

to achieve native-like pronunciation in a foreign language. The participants were Cuban immigrants between the ages of seven and nineteen, who had lived in the United States for several years. The subjects were required to read a few sentences in English and were evaluated according to the degree of foreign-accent in their speech by native speakers of American English. The results revealed that none of the children or teenagers whose pronunciation was investigated were rated as native speakers of English. Nonetheless, many of the subjects who came to the United States between one and six years of age and had lived in this country for at least five years were judged as having a near-native pronunciation. A similar study was conducted by Oyama (1976), who focused on the pronunciation of Italian immigrants to the United States and found that age of arrival (ranging from 6 to 20 years) was a strong predictor of the degree of foreign accent. More specifically, participants who came to the US at a younger age were rated higher in terms of pronunciation accuracy by native-speaker judges.

The claim that the earlier in life one learns an L2, the more native-like it will be pronounced was validated in a number of subsequent studies that examined the phonetic performance of speakers who had learned a language in a predominantly L2-speaking country. For instance, Flege (1988) investigated the pronunciation of Chinese speakers of English (who produced test sentences that were rated for foreign accent by a group of native speakers) and found that the speakers who began learning English as children (i.e. arrived in the United States at an average age of 7.6 years) were rated significantly higher than speakers who started learning later in life (i.e. arrived in the United states as adults). At the same time, both early and late learners were rated significantly lower than a control group of native speakers. Similar findings were also obtained by Suter (1976), Tahta et al. (1981), Piper and Cansin (1988), Thompson (1991) and Piske, MacKay and Flege (2001).

In a few other studies, on the other hand, an opposite effect of age on pronunciation accuracy was observed, demonstrating that early learners do not outperform late learners in all learning environments. For instance, Snow and Hoefnagel-Höhle (1977) asked native speakers of English to imitate Dutch words in a laboratory experiment (the participants had no previous knowledge of the language) and observed that younger groups were rated lower than older groups in terms of pronunciation accuracy. Olson and Samuels (1973) examined the effect of age on the acquisition of L2 pronunciation in a setting that closely resembled a normal foreign language classroom and found that it was the adults rather than children that were rated as superior in terms of foreign language pronunciation.

Rather than focus on one particular factor, some early studies investigated a range of different variables associated with pronunciation accuracy. For example, Suter (1976) analysed as many as 20 variables suspected of being significantly correlated with the degree of foreign accent in L2 pronunciation. The participants were over 60 non-native speakers of English with different L1 backgrounds, whose phonetic performance was rated by a panel of 14 native English-speaking judges. Statistical analysis of the results demonstrated that the factors most strongly related to pronunciation accuracy were: the native language of the speaker, the strength of the speaker's concern about their pronunciation and the amount of native English input they received (operationalised as the amount of time a given informant spent speaking English with native speakers at home, at work or at school). The variables that were found to have a negligible effect on the level of accentedness in the subjects' speech were: the amount of formal pronunciation training received by a given speaker, the speaker's degree of extroversion and the speaker's gender. The data obtained in the 1976 study were later re-examined using more advanced statistical techniques by Purcell and Suter (1980). In the second study, the researchers arrived at somewhat different conclusions and reported that the variables that accounted for the variability in the subjects' pronunciation ratings were: L1 background, the innate aptitude for oral mimicry and the number of years of residence in the US combined with the number of months of cohabitation with native speakers.

A range of different factors that could potentially influence L2 pronunciation were also analysed in a study by Tahta, Wood and Loewenthal (1981), who focused on the following variables: age of English acquisition, age at the time of the study, gender, language(s) spoken at home, length of residence in the L2 country (LOR), musical ability, pronunciation models (i.e. different L2 pronunciation models a given speaker had access to) and the number of languages spoken. The subjects were all immigrants to the UK who had lived in the country for a minimum of two years and learned English as a second language. The participants were recorded while reading a text passage in English; the recordings were rated for degree of foreign accent by three native speakers of English. The results of the study showed that the age at which acquisition commenced had the greatest effect on subjects' accentedness scores. The only other variable that was found to have a significant bearing on pronunciation accuracy was the amount of English use at home, i.e. the more the participants used their L2 at home, the better pronunciation they had. Interestingly, the importance of this variable could not be ascribed to increased amount of English practice since many of the subjects lived with an older generation of immigrants, who could be expected to use foreign-accented English (which meant that the effect of practice should be diminished by the type of input the

participants received at home). Tahta et al. (1981: 271) proposed that the use of English at home might have been "symptomatic of a shift of identification from the LI culture to the British" and that attaining native-like pronunciation could be viewed as "an expression of a desire to sound and to be English".

Two slightly more recent studies that share a similar focus with the aforementioned body of research (i.e. are interested in the variables affecting L2 pronunciation accuracy) were conducted by Thompson (1991) and Elliot (1995). Thompson (1991) investigated the English pronunciation of native speakers of Russian, who were required to perform three tasks: reading specially constructed English sentences, reading a prose passage, and talking spontaneously about their daily routine. The speech samples were rated by two groups of native speakers of English, language experts and laymen. The findings indicated that the best predictors of pronunciation accuracy were age of arrival in the target language country, gender, aptitude for oral mimicry and speaking proficiency in English. It was also found that the assessment and perception of a foreign accent depended on the characteristics of the language samples and the linguistic experience of the raters. More specifically, the samples containing read sentences were judged to be more accented than spontaneous speech and experienced raters were found to be more reliable and lenient in their assessments than inexperienced raters. Additionally, Thompson's (ibid.) findings suggested that regardless of the age of learning, native-like pronunciation may be difficult or impossible to achieve if the learners maintain strong connections to the L1 community and use the L1 extensively. The amount of continued L1 use was also found to affect the degree of foreign accent by Piske, MacKay and Flege (2001).

Elliot (1995) concentrated on three variables that could potentially influence phonetic performance: individual concern for pronunciation, subject's degree of field independence and subject's degree of right hemispheric specialization (the two latter terms refer to individual cognitive styles and are associated with different types of personalities and learning styles). The participants were over sixty native speakers of English learning Spanish (intermediate students studying Spanish at Indiana University). Similarly as in the studies discussed above, the subjects' phonetic performance in Spanish was evaluated according to the degree of foreign accent by three judges. Statistical analysis of the results revealed that the three investigated variables had an effect on participants' pronunciation accuracy.

Other work concerned with the factors affecting pronunciation accuracy includes the research conducted by Alexander Guiora and colleagues. The focal point of their studies was

the role of identity and other social-psychological variables in the acquisition of L2 pronunciation. In Guiora's (1972: 146) own words:

> [...] I would say that second language acquisition in all of its dimensions demands that the individual, to a certain extent, incorporate a new identity. The first step in the completion of this process is pronunciation. Since pronunciation appears to be the feature of language behaviour most resistant to change it can be assumed also to be the most critical to the individual's identity.

Guiora (ibid.) proposed that pronunciation is inextricably linked with one's "language ego" and that greater "ego permeability" (i.e. the extent to which the ego can be flexible and adapt) results in more native-like pronunciation in a foreign language. Interestingly, the hypothesis was also used to explain the discrepancy between children and adults' ability to acquire L2 pronunciation. According to Guiora (ibid.), ego permeability is greater in children since their personalities are still in the early stages of formation. However, once the development of personality is concluded, attaining native-like pronunciation becomes almost impossible.

The relationship between ego peremeability and phonetic performance was tested in two studies in which Guiora and colleagues attempted to relax participants using alcohol (Guiora et al., 1972) and valium (Guiora et al., 1980). In both studies the participants were native English learners of Thai, who were administered varying doses of alcohol or valium and then asked to take a test in Thai pronunciation. The results of the first study revealed that the participants who ingested 1.5. ounces of alcohol performed significantly better than subjects who consumed no alcohol or more than this amount. The finding was interpreted to mean that the consumption of a small dose of alcohol relaxed the speakers' inhibitions and increased their ego permeability, thus resulting in more-native like pronunciation. The results of the second study showed no direct correlation between phonetic performance and the administered dose of valium.

Another key concept in Guiora's work was empathy, which was argued to be crucial for successful L2 acquisition (e.g. Guiora, 1972). The effect of empathic capacity and other personality characteristics on the "authenticity" of pronunciation was examined by Taylor et al. (1971). The participants, native speakers of American English, attended four one-hour sessions in which they were taught basic dialogues and sentence patterns in Japanese by a native speaker of the language. The subjects' phonetic performance was rated according to the degree of foreign accent by native Japanese judges and compared with their scores on a number of psychological tests. The results of the study suggested that pronunciation accuracy

was related to psychological variables such as tolerance to anxiety, intelligence, involvement in emotional experiences and perception of emotional expression.

Related to Guiora's (1972) work are two theories that are not concerned exclusively with L2 pronunciation but deal with the more general subject of L2 acquisition and the factors affecting successful attainment of a foreign language. The first theory is Gardner and Lambert's socioeducational model of second language acquisition (e.g. Gardner and Lambert, 1972; Gardner, 1985), which underscores the role of motivation in foreign-language learning and proposes that successful acquisition is more likely for learners who are motivated to learn a second language, are open to other cultural communities, have favourable attitudes towards the learning situation and possess low levels of language anxiety. The second theory is Schumman's Acculturation Model (Schumann, 1978, 1986), which has its roots in Guiora's (1972) hypothesis about ego permeability and Gardner and Lambert's (1972) model of L2 acquisition. Schumann's (ibid.) framework posits that the key element in the acquisition of a foreign language is acculturation, a macro variable that encompasses a variety of social and psychological factors. The social factors he mentions are concerned with the characteristics of the target language group and the relationship between the target language group and the L2 learning group; they include social dominance patterns, the size of the target language population and the amount of congruence between the learning group and the target language group. The psychological (affective) factors relate to individual characteristics of the learner and include ego permeability, language and culture shock and motivation. The model predicts that learners will acquire the L2 to the extent they acculturate to the target language group, i.e. to the degree they integrate socially and psychologically with the target culture. Some support for the view that language and group identity are interrelated was provided in sociolinguistic studies on L2 pronunciation that are discussed in Section 2.4. (e.g. Gatbonton, 1975; Zuengler, 1982; Gatbonton and Trofimovich, 2008; Gatbonton, Trofimovich, and Segalovitz, 2011).

2.3. The notion of interlanguage

Interlanguage (IL), one of the key concepts in Second Language Acquisition (SLA), has had a major impact on the evolution of L2 phonetics as a scientific discipline. Before the advent of IL, attempts were made at predicting the behaviour of second language learners by focusing on the similarities and differences between a speaker's native language (NL) and the target

language (TL) they were learning. This line of thinking lead to the development of the Contrastive Analysis Hypothesis (CAH) (Fries, 1945; Lado, 1957). One of the central assumptions of CAH was that the knowledge of how a learner's L2 differs from their L1 would enable one to predict all language errors. It was expected that features that were distinct in the L1 and the L2 would be difficult for the L2 learners to acquire, while similarities between the L1 and the L2 would aid the acquisition of certain features. In other words, it was believed that the errors in a speaker's L2 performance could be accounted for and predicted solely by transfer from NL to TL. This kind of outlook on second language performance is visible in many of the early studies on the L2 sound system (see previous section). Admittedly, the early studies focused on a variety of factors that could potentially affect a speaker's pronunciation in a TL (AOL, LOR, L1 background, aptitude for oral mimicry, etc), however, these variables were only viewed as elements that would likely reduce or enhance the main effect of L1 transfer.

Lack of empirical support for the claims of the CAH and the fact that the framework came under some criticism prompted linguists to explore different approaches to the issue of second language acquisition (Major, 2008). The term interlanguage (IL) was introduced by Selinker (1972) and later elaborated on in a detailed account of the historical development of research on learner language (Selinker, 1992). Selinker (1972) defined IL as a separate language system that generates linguistic output when learners attempt to produce utterances in the language they are learning. According to Selinker (ibid.), the IL is distinct from the learner's NL and the TL. At the same time, the three linguistic systems are united psychologically by interlingual identifications made by the learners. The hypothesis claims that IL is shaped by five processes: language transfer, transfer-of-training, strategies of second-language learning, strategies of second-language communication and overgeneralisation of TL linguistic material. Language transfer refers to the interference from the learner's L1. Transfer-of-training takes place when L2 users apply, sometimes erroneously, rules they learnt from textbooks or teachers (Selinker, ibid.; Tarone, 1994). Strategies of second-language learning relate to the methods the learners use to master a second language and the influence of these methods on L2 production (Tarone, 1994). Communication strategies refer to the way learners resolve communication problems and the impact of the selected communication strategies on the L2 utterances they produce (Tarone, 1994). Finally, an overgeneralisation of TL linguistic material can be defined as using an existing TL rule to produce a TL element to which the rule does not apply, e.g. using the *–ed* ending to form the past participle of an irregular English verb (Selinker, 1972). One of the outcomes of these five processes and their

combinations may be the occurrence of fossilized elements in a learner's IL. As explained by Selinker (ibid.: 215),

> Fossilizable linguistic phenomena are linguistic items, rules, and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL, no matter what the age of the learners or amount of explanation and instruction he receives in the TL. [...] Many of these phenomena reappear in IL performance when the learner's attention is focused upon new and difficult intellectual subject matter or when he is in a state of anxiety or other excitement, and strangely enough, sometimes when he is in a state of extreme relaxation.

Importantly, what follows from the tenets of Selinker's (ibid.) framework is that the productions of a given L2 speaker are the outcome of a number of processes, L1 transfer being only one of them. The hypothesis that a learner's phonetic performance in an L2 should be treated as the output of an IL rather than the product of L1 interference was tested by Flege (1980), who examined the production of voiced and voiceless stops by Saudi Arabian learners of English. This particular variable was selected for investigation because voicing contrast in stops is realised differently in Arabic and in English. Acoustic analysis of three correlates of stop voicing (vowel duration, stop closure duration and voice-onset time) in the subjects' productions revealed that, although the subjects generally realised voicing contrasts differently than a reference group of native-speakers, the values they produced were also atypical of their L1. In fact, the result indicated that the subjects' values for vowel duration, stop closure duration and voice-onset time were intermediate to the values found in English and Arabic. Additionally, it was found that some of the Saudi learners did realise certain voicing contrasts in a native-like manner. Flege (ibid.: 117) summarised his findings with the following statement, "[t]he Saudis' English speech, although not typically Arabic or English in phonetic terms, seems to be the product of a fairly stable interlanguage phonetic system which admits the possibility of phonetic strategies by individual speakers". Similar results were obtained by, among others, Caramazza (1973), Flege (1991), Flege and Eefting (1987) and Schmidt and Flege (1996), providing further support for the claim that the phonetic shape of L2 utterances is a product of a separate linguistic system, which differs systematically from both the NL and the TL.

Another claim that can be inferred from Selinker's (1972) seminal paper is that IL should be treated and examined differently than other language systems. Unlike other languages, it tends to fossilize, it is subject to transfer from other systems and it is generated by a different acquisition device ("latent psychological structure", Selinker, ibid.: 211). The view that IL is

somehow different from other linguistic systems was contested by Adjemian (1976), who argued that interlanguage should be regarded as a natural language, i.e. "[...] any human language shared by a community of speakers and developed over time by a general process of evolution" (ibid.: 298). Adjemian claimed that, just as in the case of other natural languages, ILs contain a system of rules and show internal consistency. Also, similarly to other linguistic systems, ILs may be affected by socio-linguistic factors such as style-shifting. An important property of IL is its stability, i.e. the tendency on the part of the learner to repeatedly use a given linguistic feature or rule, whether it be correct or incorrect as compared with the TL norm. In Adjemian's (ibid.) view, the notion of stability fits nicely with the concept of fossilization. He argues that the appearance of fossilized linguistic items may be viewed as a reflection of the IL's stability. What sets IL apart from other natural languages is that its stability can be disrupted by the existence of another IL characteristic, i.e. permeability. This property of learner language manifests itself when "either the IL system is penetrated by rules or forms of the NL not usually evidenced in its speech forms, or an internalized TL rule or form is improperly generalized or distorted in some way" (ibid.: 308). In other words, although ILs can be considered natural languages, they are unique in the sense that they are permeable to invasion from other language systems.

Some support for Adjemian's (ibid.) claims can be found in the data obtained by Dickerson (1975). The researcher investigated variability in the English pronunciation of native Japanese speakers. The participants were recorded three times over a nine month period; each time, they were required to produce English utterances in three different speech styles: free speech, dialogue reading and word-list reading. The phonetic variable under investigation was English /z/, whose production in Japanese learners' speech can show considerable variability (it can be realised as /s/ or /dz/ or omitted altogether). The results demonstrated that there was indeed considerable variability in terms of how the participants realised English /z/. However, Dickerson (ibid.) discovered that there were certain underlying regularities in the distribution of the variants used by the subjects. First, phonetic environment was found to have an influence on the distribution of /z/ in the learners' speech, i.e. different proportions of a given /z/ variant were used according to the sound that followed it. Also, Dickerson (ibid.) observed that the number of TL forms in the subjects' speech increased over time, i.e. more native-like realisations of /z/ were used in the third recording than in the second and the first recording. Finally, it was found that speech style had a bearing on the distribution of the independent variable, i.e. the proportion of TL forms increased in the word-list-reading task as compared with the dialogue-reading task and the free-speech task. Taken together, Dickerson's (ibid.)

findings corroborate Adjemian's (1976) claim that interlanguage is a rule-governed system that shows internal consistency and can be conditioned by socio-linguistic factors such as style-shifting. In Dickerson's words, "The learner [...] is generating utterances which are rule-governed according to his interlanguage system of variable rules" (Dickerson, 1975: 406). Also important in Dickerson's (ibid.) paper is the noticeable shift towards a sociolinguistic approach in the study of L2 phonetic performance, brought on by the conceptualisation of IL as a natural language. More socioliguistically-based studies on L2 pronunciation will be discussed in Section 2.4.

The structure and development of IL were further expanded on by Major, in his Ontogeny Model (Major, 1987), which was later revised to Ontogeny Phylogeny Model (OPM, Major, 2001; Major, 2008). As stated by its author, the purpose of introducing OPM was to "[...] offer an integrated view of the way the L2 learner's phonology changes over time, [...] rather than merely attempting to explain the possible sources of error of an L2 learner at one particular stage" (Major, 1987: 102). According to OPM, interlanguage comprises elements of L1, L2 and language universals (which include markedness considerations, learning principles, Universal Grammar, rules, processes, constraints and stylistic universals). An important claim made in OPM is that there exist and interaction between the IL components and that it develops chronologically. It is hypothesised that at the outset of acquisition, the learner's interlanguage consists solely of L1. Over time, the impact of the L2 and language universals begins to gradually increase. Later the amount of L2 grows further, while the amount of language universals diminishes. Finally, in an idealised learner, IL equals L2. To verify the claims made in OPM, Major (1987) carried out a pilot study that focused on the English pronunciation of Brazilian Portuguese learners. The subjects were teachers and students at a university in São Paulo, who were divided into two groups according to the degree of experience with the L2. The phonetic variables analysed by Major were the production of /r/ and final consonant clusters, devoicing of word-final obstruents and paragogue. The data were collected by asking the participants to read a word list, a sentence list and a short text. Overall, the results indicated that L1 interference and substitutions related to language universals were more common in the pronunciation of the less experienced group of learners. Also, depending on the particular phonetic variable and stage of development, the pronunciation of beginner learners showed either a greater or lesser occurrence of language universals when compared to advanced learners (which confirms the models predictions, as language universals are expected to first increase and then decrease). The claims of OPM

were also confirmed in Flege's (1980) study on VOT in the speech of Saudi Arabians. The findings indicated that more experienced learners produced more native-like values than less experienced speakers, pointing to a gradual approximation of the L2 norm. The existence of different stages in the acquisition of pronunciation were also observed by Wode (1981), who analysed the acquisition of the English sound system by his German-speaking children. Wode reported that English /ʌ/ and /æ/ were first substituted with [a] and [ɛ] respectively (showing L1 transfer from German) and then gradually approximated TL realisations. It was also found that the acquisition of /r/ progressed from [R] to [w] to [ɹ] and finally to a target-like retroflex.

Another important claim made in the OPM is that the rate of development of the IL will vary from individual to individual. Very good learners are expected to progress rapidly; it is predicted that transfer from the L1 will be less pronounced in their IL. In the case of poorer learners, the progress is predicted to be much slower and L1 transfer is expected to predominate in the IL. What is also interesting about Major's (1987, 2001, 2008) model is that it introduces language universals as one of the key components of interlanguage. Indeed, studies have shown that universal principles may play an important part in L2 speech production. For instance, Waniek-Klimczak (2002) examined stress placement by Polish learners of English and found that the participants assigned word stress using complex strategies that could not be simply ascribed to L1 transfer. It transpired that the learners tended to stress long vowels and diphthongs, which is a universal tendency in languages (Major, 2008). Focusing on VOT productions of Polish learners, Waniek-Klimczak (2005) found that voicing lag values were generally longer for /k/ than for /p/ and longer preceding a high vowel than preceding a low vowel. Similar observations were made by Piotrowski (2013), who also examined VOT values in the speech of Polish learners of English. His data show that /k/ was generally realised with longer voicing lag values than /t/, while /t/ was generally produced with longer voicing lag values than /p/. The findings by Waniek-Klimczak (2005) and Piotrowski (2013) agree with two universal tendencies: 1. velar stops generally have longer VOT values than alveolar and bilabial stops (e.g. Lisker and Abramson, 1964; Cho and Ladefoged, 1999), 2. VOT values tend to be longer before high vowels than before mid and low vowels (e.g. Klatt, 1975; Summerfield, 1975).

2.4. Sociolinguistic studies on L2 pronunciation

Following Adjemian's (1976) claim about the nature of interlanguage and inspired by Labovs' (1969) classic work on L1 variability, Tarone (1979) argued that IL can and should be treated as a natural language and that as a natural language, it is variable and undergoes style-shifting just as L1 does. This type of thinking gave rise to a new approach to the study of the L2 sound system and lead to the development of sociolinguistic research on non-native pronunciation. One of the first sociophonetic studies on L2 speech was Dickerson's (1975) work on variability in the pronunciation of Japanese learners of English (discussed in the previous section). Dickerson's findings were expanded on in a subsequent study (Dickerson and Dickerson, 1977), where the researchers focused on the realisation of English /r/ by Japanese speakers of English. Similarly as in the 1975 investigation, it was found that the distribution of TL forms was conditioned by the style of speech; the subjects produced /r/ almost 100% correctly in a word-list reading task but only 50% correctly in free speech. These and similar findings (e.g. Gatbonton, 1975, 1978) prompted Tarone (1979, 1982) to propose that a learner's interlanguage is a continuum of styles, ranging from the superordinate on one end and the vernacular on the other. Tarone (ibid.) argued that the superordinate is the style that is used in formal situations, when the learner pays the greatest amount of attention to language form. The vernacular, on the other hand, was regarded as the style used in informal situations, when the learner pays least attention to language form. According to Tarone (ibid.), the crucial difference between the two styles lies in the fact that interlanguage superordinate is more permeable to invasion from the rule system of the TL. It follows from this assumption that more formal elicitation tasks such as word-list reading (in which the learner is expected to pay increased attention to the language form) should result in the production of more TL forms than less formal tasks such as free speech (in which the learner is expected to be less focused on the language form).

Tarone's (ibid.) claims about the effect of attention to speech were only partly corroborated by empirical data. The results obtained by Dickerson (1975) and Dickerson and Dickerson's (1977) corresponded with the assumption that a more formal style will induce learners to use more TL forms. However, the findings of several other studies suggested that the relationship between task formality/attention to language form and L2 pronunciation is quite complex. For instance, Beebe (1980), who examined the production of /r/ by Thai learners, observed that while the participants used more TL variants of final /r/ in a more formal listing task than in a

conversation task, they also used more L1 variants of initial /r/ in the careful style. Beebe (ibid.) attributed the higher rate of L1 realisations of initial /r/ in the formal task to the fact that they function as a socially prestigious variants in the learner's native language. Schmidt (1987) found that Arabic learners of English used a greater proportion of interdental fricatives when reading minimal pairs as compared with reading a text passage. At first glance, the results seem to support Tarone's hypothesis (1979, 1982) about the effect of attention to language form on L2 pronunciation. However, what needs to be taken into consideration is that the linguistic behaviour of the participants in Schmidt's (1987) study reflected the social stratification of interdental fricatives in the subjects' L1. Overall, it seems that the amount of attention a learners pays to speech form does contribute to some variability in L2 phonetic performance. Nonetheless, it appears to be interconnected with factors such as phonetic environment and prestige variants in the speaker's L1. Other studies that point to a complex relationship between task type/attention to speech form and L2 pronunciation include Sato (1985) and Zampini (1994).

Apart from variation along the speech style continuum, another major area of interest within the sociolinguistic strand of research on non-native speech was the effect of social group affiliation and identity on L2 pronunciation. Anisman (1975) studied the realisation of selected phonetic variables by Puerto Rican speakers of English living in New York. The participants were male adolescents who were grouped according to the amount of contact they had with Black peer groups. Anisman (ibid.) collected data by means of structured interviews and focused his analysis on the subjects' pronunciation of the function word *to*, the /aɪ/ diphthong and the voiced dental fricative, as the realisation of these elements varied between New York Standard English, Black English (the two varieties of L2 available to the participants) and Puerto Rican Spanish (the subjects' L1). The results showed that those informants who interacted frequently with the Black peer group used more Black English variants than Standard English or Spanish variants, the subjects who targeted mainstream values and norms produced more Standard English variants over Black English and Spanish variants, whereas the speakers who interacted predominantly with the Puerto Rican peer group used the greatest proportion of Spanish variants. Anisman's (ibid.) findings highlight the importance of social group membership and input in the acquisition of L2 pronunciation.

In a similar investigation, Thompson (1975) examined the realisation of English word-final /z/ and /aɪ/ by Mexican-Americans in order to determine whether the informants would exhibit variants characteristic of standard English, regional English or Spanish-accented English. It

was found that the production of the two phonetic variables was related to the subjects' socioeconomic status, level of education, type of occupation and attitude towards pronunciation (i.e. whether or not they viewed accent as important for socioeconomic advancement). More specifically, Thompson (ibid.) observed that the speakers who used regional variants attended high school, had higher socioeconomic status, had occupations that dealt with the public and felt that accent was important for socioeconomic advancement. Speakers who exhibited variants typical of standard English had high socioeconomic status but did not attend high school, did not deal with the public in their jobs and did not feel pronunciation to be important for socioeconomic advancement. Finally, speakers who used Spanish-accented variants had lower socioeconomic status, had dropped out of school before high school, usually spoke Spanish at work, had limited contact with native English and appeared unaware of accent or dialect differences in English.

The influence of socioeconomic status on L2 pronunciation was also referred to by Schmidt (1987), who reported that the participants (Egyptian learners of English) of lower socioeconomic background produced significantly less target-like realisations of English interdental fricatives than subjects with higher socioeconomic status. As far as Schmidt's (ibid.) findings are concerned, however, it should be mentioned that the linguistic behaviour of the Egyptian learners' could be attributed to the fact that the distribution of interdental fricatives is socially stratified in the subjects' L1 (Egyptian Arabic). The finding points to an interaction between sociolinguistic factors and L1 transfer in the production of L2 speech.

Gatbonton (1975) examined the relationship between ethnic group affiliation and the use of selected L2 phonetic variables. The subjects were French-Canadian learners of English, who were divided into two groups on the basis of their self-identification as nationalistic (strong pro-French attitudes) or non-nationalistic (strong pro-English attitudes). The results of the study showed that speakers who exhibited a pro-English bias produced more TL forms than the nationalistic learners. Another aim of Gatbonton's (ibid.) study was to gauge the attitudes of French-Canadian learners of English towards the pronunciation of their peers. Interestingly, it was found that the degree of foreign-accent in the peers' L2 was treated as indicative of their ethnic identity, i.e. heavily accented speech was interpreted as an indicator of strong pro-French attitudes. Gatbonton (ibid.) also observed that phonetic environment had some bearing on participants' pronunciation accuracy in the production stage of the experiment, i.e. more TL variants of /h/ and /ʔ/ were found in a postvocalic environment than in a postconsonantal environment. More recently, Gatbonton, Trofimovich and Segalowitz (2011) found further empirical evidence for the influence of ethnic group affiliation on the pronunciation of L2

learners. Examining voiced interdental fricatives in the speech of Canadian Francophone learners, the researchers found that the strength of ethnic group affiliation was negatively correlated with pronunciation accuracy.

Zuengler (1982) sought to determine whether asking the informants an ethnolinguistically-threatening question would affect their L2 phonetic performance. The subjects were native Spanish and Greek learners of English who were asked to respond to remarks made by a native speaker of English, one of the which was designed to threaten the participants' ethnic identity. The learners' responses were analysed with respect to their production of three phonetic variables: pre-vocalic /r/, /ɪ/, and word-final /z/. It was found that, depending on the nature of their reply to the ethnolinguistically threatening question, the participants either increased or decreased the proportion of TL forms in their pronunciation. It transpired that the subjects who provided personal and ethnically related replies produced less native-like variants than informants who objectified their response and made no direct references to their ethnicity. Zuengler (ibid.) suggested that the learners who gave more personal responses may have indentified strongly with their ethnic group and increased the degree of foreign-accent in their speech by way of defending their ethnic solidarity. A somewhat similar statement was made by Thompson (1991), who investigated the pronunciation of Russian born immigrants to the Unites States and argued that participants' strong sense of national identity contributed to their low pronunciation accuracy ratings.

Adamson and Regan (1991) focused on the phonetic realisation of the English (ing) suffix by native Vietnamese and Cambodian speakers of English as compared with the pronunciation of a native-speaker control group. The researchers wished to determine which of the two phonetic variants of (ing), [iŋ] or [ɪn], would be more common in the learners' productions. The results of the study indicated that the phonetic realisation of (ing) was conditioned by the gender of the participants (the effect was observed for both the native and the non-native subjects). It was found that women produced a greater proportion of the prestige variant [iŋ], whereas men used [ɪn] more frequently. A similar tendency was also observed by Hartford (1978), who studied the pronunciation of female and male Mexican-Americans. Adamson and Regan (1991) hypothesised that the difference in L2 phonetic performance occurred because non-native speakers were accommodating to gender-specific norms, i.e. women targeted the variant used by female native speakers, while men preferred the variant employed by male native speakers. Importantly, it was found that non-native

speakers' pronunciation of (ing) was also affected by phonetic environment and grammatical category.

Another issue that should be mentioned in relation to sociophonetic studies on L2 performance is the concept of social marking. Speech (social) markers can be described as phonetic features that serve important social functions in communication. They are generally believed to have two main purposes. On the one hand, they "[...] permit speakers to reveal their association with broadly defended biological, social and psychological states [...]" and "[...] serve the general function of maintaining the social system by indentifying and recognizing members who occupy various roles and hierarchical positions within it" (Giles, Scherer and Taylor, 1979: 343). Additionally, "[...] speech markers permit interlocutors indirectly to communicate important attitudes, beliefs, values and intentions about their own social states [...] (ibid.: 344)". As argued by Dowd, Zuengler and Berkowitz (1990), the concept of social marking can be used to interpret sociolinguistic conditioning in L2 speech. For instance, as far as Anisman's (1975) findings are concerned, the Black English, Standard English and Spanish English variants used by the subjects could be seen as speech markers that served to signal association with a given social group. The findings obtained by Thompson (1975) and Adamson and Regan (1991) could be interpreted similarly. More evidence for the claim that social marking operates in L2 speech can be found in the study by Gatbonton (1975). The researcher found that learners who exhibited strong pro-English attitudes produced more dental fricatives than participants who identified strongly with the French community. The results could be interpreted to mean that dental fricatives functioned as social markers in the subjects' productions and that the learners accentuated/attenuated the amount of TL variants in order to indirectly express their attitudes towards the English and French communities. In a similar vein, it could be assumed that the foreign-accented variants in Zuengler's (1982) study served as social markers that permitted the subjects to defend their ethnic identity.

2.5. The link between perception and production in L2 pronunciation

Another important area of research within the field of L2 phonetics concentrates on the process of L2 speech perception and its effect on the production of L2 sounds. The studies reviewed in this section adopt a psycholinguistic approach towards non-native pronunciation; their interest lies in how the brain processes the sounds of a second language. One of the first

researchers to call attention to the fact that foreign accent stems from inaccurate perceptual patterns rather than from production difficulties was James Flege, a prominent and prolific figure in the field L2 pronunciation (some of his studies have been mentioned earlier in this chapter). Flege is well-known for his work on the effect of perceptual similarity between NL and TL sounds on the production of second-language speech. The issue appeared regularly in his early work (e.g. Flege, 1981; Flege, 1987; Flege and Eefting, 1987; Flege and Eefting, 1988; Bohn and Flege, 1992; Flege, Murray and Skelton, 1992) and gradually developed into the Speech Learning Model (SLM, Flege, 1995). The main focus of the model is on the perception and production of experienced, adult L2 learners. SLM is based on the assumption that L1 and L2 sound categories exist in the same phonological space and are related perceptually at a position-sensitive allophonic level. According to the model, if the learners perceives a difference between an L2 sound and the closest L1 sound, they will establish a new category for the L2 phone. Generally, the greater the perceived dissimilarity between a certain L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be noticed by the learner, which, in turn, should result in accurate production of the said L2 sound. Conversely, if the learner does not perceive a difference between a given sound in their NL and their TL, both sounds will be processed as belonging to the same phonetic category and will resemble each other in production (a mechanism termed equivalence classification). SLM places much emphasis on the age of learning and predicts that the earlier L2 acquisition commences, the easier it is to establish new categories for L2 sounds; increased language experience is also considered to have a facilitative effect on the formation of separate L2 phonetic categories.

Support for the claims of SLM can be found in Flege's early studies. For instance, Flege (1987) examined the production of French /t/ and /y/ by native speakers of English with different levels of experience in French. As argued by Flege, French /t/ has a similar counterpart in the English sound system, while /y/ does not. If one treats TL /t/ as similar to NL /t/, English learners of French can be expected to struggle with a native-like realisation of this sound, as it will be perceptually assimilated to an L1 phonetic category. French /y/, on the other hand, has no close equivalent in the English sound system (ibid.). Consequently, according to SLM, a new category should be established for French /y/, which, in turn, should eventually lead to accurate realisations. Indeed, the results of the study revealed that experienced learners realised /y/ with formant values similar to that of native speakers of French. On the other hand, both experienced and inexperienced learners produced /t/s that differed significantly from the native French speakers' productions, thus corroborating

Flege's (1995) postulate about the relevance of cross-language phonetic similarity in the acquisition of L2 sounds.

Flege and Eefting (1987) focused on the pronunciation of Dutch learners of English who varied in terms of level of proficiency and L2 exposure. The phonetic feature under investigation was VOT duration in the subjects' productions of English /t/. It was assumed that the English alveolar stop may be difficult for Dutch learners to master since a similar stop consonant is also used in their L1, the fundamental difference between the two being that the Dutch counterpart is unaspirated. Acoustical analysis of the data demonstrated that nearly every participant produced a longer mean VOT in English than in Dutch and that the magnitude of the production shift was significantly greater for proficient than non-proficient learners. Also, some of the proficient informants realised /t/ with VOT values that closely resembled those exhibited by the control group of native speakers. These results were taken to mean that proficient subjects formed a new category for English /t/ as a result of extended exposure to the L2, which enabled them to produce more native-like VOT values than less proficient participants (who, presumably, did not receive sufficient L2 input to approximate the native realisation of English /t/). It is also worth mentioning that many of the subjects in Flege and Eefting's (ibid.) study produced mean VOT values that were intermediate to the phonetic norm for VOT in L1 and L2, which was also the case in Flege's (1980) examination of stop realisation by Arabic learners of English and the investigation of French /t/ by native English speakers (Flege, 1987). A similar effect in relation to VOT was also observed in several other studies, e.g. Caramazza et al. (1973), Flege (1991), Schmidt and Flege (1996).

Flege, Murray and Skelton (1992) aimed to test the claims of SLM with respect to the realisation of the English word-final /t/-/d/ contrast in the speech of Spanish and Mandarin speakers of English, all of whom were late learners with differing levels of language experience. Native speakers of Spanish and Mandarin were selected since it could be argued that the languages do not to have a word-final voicing contrast between alveolar plosives (final /t d/ are rare in Spanish; in Mandarin, no obstruents are permitted in word-final position). Following the SLM principle that "new" sounds are less likely to undergo equivalence classification, it was hypothesised that experienced learners should be able to produce native-like distinctions between voiced and voiceless stops. Acoustic analysis of the data revealed that the participants distinguished the two stops by means of vowel duration, F1 offset frequency, stop closure duration and the presence/absence of closure voicing, which concurred with the results obtained for the native-speaker reference group. Although there were some instances in which experienced learners resembled native speakers more closely

than inexperienced learners, on the whole, no significant difference between the two groups of non-native speakers was found. Thus, the results of the study did not substantiate SLM's claims, for which Flege et al. (ibid.) provided several different explanations. One of the possibilities considered by the authors was that the subjects (even those with a higher level of language experience) had not received sufficient phonetic input to fully master the voicing contrast under investigation. It was also argued that "[...] the SLM may be wrong in claiming that all adult learners who have received sufficient L2 phonetic input will master new consonants in an L2. Perhaps new consonants can be mastered by only a *small proportion* of adult L2 learners (ibid.: 140)". Another important finding was that although the subjects did distinguish word-final /t/ and /d/, the differences produced by the non-native speakers were generally smaller in magnitude than the differences produced by the native-speaker reference group. Indeed, the findings of a number of other studies have also shown that late learners of English tend to produce smaller voicing effects than native speakers of the language (e.g. Elsendoorn,1985; Flege and Port, 1981; Mack, 1982).

In a more recent study, Flege, Schirru and MacKay (2003) examined the production of English /e/ by Italian immigrants to Canada. The subjects differed in terms of the amount of L1 use and age of arrival in the L2-speaking country (AOA). The /e/ vowel was selected for analysis since earlier research suggested that native speakers of Italian may identify English /e/ tokens as instances of the Italian /e/ category, even though the English vowel is produced with more tongue movement than the Italian counterpart. The results revealed that both AOA and L1 use were significantly correlated with the participants' production of English /e/. It was found that early learners realised the vowel more accurately than late learners and that low-L1-use subjects tended to produce /e/ more accurately than high-L1-use subjects. The findings were interpreted within the framework of the SLM; the authors argued that the late, high-L1-use learners were unable to realise English /e/ in a native-like manner because they had not established a separate phonetic category for the L2 sound (i.e. English /e/ and Italian /e/ had undergone equivalence classification). Interestingly, it was also observed that the early, low-L1-use learners realised a large number of /e/ tokens with exaggerated movement. Flege et al. (ibid.) suggested that these "hypercorrect" realisations might have stemmed from the fact that the early, low-L1-use learners succeeded in establishing a separate phonetic category for English /e/ and, at the same time, were more likely to identify strongly with Canadian culture and/or the English language (i.e. wished to sound "more Canadian'' for affective or socio-cultural reasons). Overall, the results of the study seem to provide further evidence for the significance of cross-phonetic similarity in the acquisition of L2 sounds.

Furthermore, the findings lend support to SLM's claim that phonetic category assimilation is related to the amount of language experience and input (it seems reasonable to assume that the less the subjects used their L1, the more opportunities they had to use and listen to the L2). Language experience and input were also found to have a considerable bearing on L2 pronunciation in a study by Flege and Liu (2001), who investigated the identification of word-final English consonants by Chinese immigrants to the United States. The researchers observed that participants with a relatively long length of residence obtained higher scores than subjects with a relatively short LOR. Importanly, the difference was only significant for those learners who received substantial native-English input.

Apart from Flege's (1995) Speech Learning Model, the process of L2 speech perception was also examined in Kuhl and Iverson's (1995) Native Language Magnet (NLM) theory and Best's (1995) Perceptual Assimilation Model (PAM). Although all three frameworks deal with the perception of foreign speech sounds, an important difference between SLM and the remaining two theories lies in their focus of interest. SLM concentrates on speech perception in order to predict and explain production difficulties, i.e. it is the production of L2 speech sounds that is of paramount importance in the model. The focal point of NLM and PAM, on the other hand, is the effect of L1 interference on the perception of L2 sounds.

The Native Language Magnet theory is concerned with the way L1 experience modifies and reshapes the underlying representations of phonological categories. According to Kuhl and Iverson (1995), humans are born with the ability to discriminate between the sounds of all languages, but, as they grow older, their perception of speech sounds becomes increasingly language-specific and homed in on the phonological categories typical of the L1. This process is claimed to be a result of the "perceptual magnet effect", which "warps the acoustic space underlying phonetic perception" (Kuhl and Iverson, 1995: 121-122). The magnet effect plays a vital role in second-language learning as it makes it more difficult to recognise and discriminate between the speech sounds of a foreign language, which, in turn, renders it difficult to *produce* the speech sounds of the said foreign language. The findings by Rochet (1995) provide some empirical evidence for Kuhl and Iverson's (ibid.) proposal. The researcher analysed perception and production of French /y/ by native speakers of two languages: Portuguese and English. It was found that the participants perceptually assimilated the French vowel to L1 sounds, i.e. the native Portuguese subjects tended to misidentify it as /i/, while the native English subjects tended to misidentify it as /u/. The results of a subsequent production task echoed the findings of the perception task: native speakers of Portuguese realised French /y/ as /i/; native speakers of English substituted French /y/ with

/u/. Empirical evidence for the tenets of NLM was also found in a study by Iverson et al. (2001), who examined the perception of English /r/ and /l/ by native speakers of Japanese and observed that it was modified by the participants' L1 phonetic system.

Best's (1995) Perceptual Assimilation Model attempts to explain the exact manner in which L1 experience affects the perception of foreign speech sounds. The model is predominantly interested in the gestural similarities and dissimilarities between native and non-native segments. According to PAM, non-native sounds will be perceptually assimilated to L1 phonological categories that they are similar to in terms of the articulatory gestures involved in their production. If there exist large discrepancies between the native and non-native segments, PAM predicts that the foreign speech sounds will not assimilate strongly to any native category. Although the early version of PAM focuses solely on speech perception by naive listeners with no experience with an L2, the framework was later modified to predict patterns of non-native segmental perception by L2 learners. PAM-L2 (Best and Tyler, 2007) assumes that L2 sounds can be assimilated to L1 phonological categories on the basis of similarity in terms of their articulatory-phonetic realisation and phonological functions (as opposed to perceptual similarity as proposed in SLM). Support for the claims put forward in PAM-L2 can be found in some of the studies conducted by Flege. For instance, the results of Flege's (1987) study on the production of French /t/ and /y/ by native speakers of English validate the hypothesis that cross-language phonetic similarity may affect the formation of L2 sound categories in the acquisition of L2 sounds

2.6. Instructed learning of L2 pronunciation

Since the current investigation concentrates on the pronunciation of learners enrolled in an English phonetics and phonology course (see Section 3.6.3.), it is relevant to discuss previous research on the relationship between formal pronunciation instruction and the acquisition of a foreign language sound system. One of the first researchers to investigate this issue was Dziubalska-Kołaczyk, who conducted a series of experiments on the influence of metalinguistic knowledge on the application of phonostylistic processes by L2 learners of English (see Dziubalska-Kołaczyk, 1990). The results of the experiments showed that subjects who were taught English in a formal setting (i.e. received formal instruction and training) used casual speech processes in a greater number of contexts than subjects who learned in a natural setting. As opposed to learners who received formal training, the natural setting

learners tended to produce assimilated forms only in lexicalised versions of utterances (e.g. *Tell me what you want)*. Phonostylistic processes in L2 speech were also investigated by Zborowska (1997), who focused on the pronunciation of Polish learners of English and observed significant improvement in perception and production of casual speech processes in an experimental group (the members of which received explicit training in phonostylistics) as compared with the control group.

Another early study that concentrated on the effect of pronunciation instruction on the acquisition of the L2 sound system was carried out by Derwing, Munro and Wiebe (1998). The researchers asked native speakers of English to evaluate the accentedness, fluency and comprehensibility of speech samples that were collected from three groups of English as a second language (ESL) learners. Each of the three groups received different pronunciation instruction: one group was taught with a focus on segmental aspects of pronunciation (they performed tasks designed to improve their productions of individual sounds), a second group was taught with a global focus (the teacher concentrated on features such as speaking rate, intonation, rhythm), and a third group received no specific pronunciation instruction (these students attended "regular" ESL classes designed to improve their reading, writing and listening skills). The speech samples were collected prior to and following the completion of the course and included short read utterances and extemporaneous narratives. The results revealed that none of the groups showed any noticeable improvement in accentedness scores in the extemporaneous narratives. However, in the case of read utterances, native English listeners judged the segmental group to be significantly less accented at Time 2 than the remaining groups, which suggested that the application and type of pronunciation instruction had some impact on learners' phonetic performance over time.

Wrembel (2004) focused on the pronunciation of Polish learners of English, who were first year students at a teacher training college in Poznań. All subjects participated in a 50-hour practical pronunciation course; the experimental group received explicit theoretical training that was designed to develop their phonological metacompetence and included articulatory descriptions, elements of contrastive analysis, theoretical readings and discussions. The control group received practical training only (listening to tape recordings and imitation). The results revealed that the experimental treatment lead to significant improvement in pronunciation performance. It was also found that the experimental group outperformed the control group in word-list and dialogue reading. In spontaneous speech, however, none of the two groups was significantly more target-like than the other in terms of pronunciation.

The relationship between formal instruction and the acquisition of L2 pronunciation has also been addressed from a theoretical standpoint. Formulating her claim within the framework of Natural Linguistics, Dziubalska-Kołaczyk (2002) argued that conscious knowledge of phonetics and phonology facilitates the acquisition of L2 pronunciation. In the same vein, Wrembel (2005) proposed a metacompetence-oriented model of phonological acquisition that underscores the cognitive aspect of pronunciation learning. The model posits that the acquisition of the L2 sound system can be facilitated by conscious knowledge of phonetics and phonology and a developed metalinguistic awareness. Metaphonological competence is a key concept in the model; it is defined by Wrembel (ibid.: 2) as "conscious knowledge of and about the grammar of the language [...] which may be developed by making the learner metalinguistically aware of L2 phonetics and phonology". It is postulated that phonological metacompetence can act as a facilitating device in different stages of acquisition and can perform the following functions:

1) facilitator of intake - operating at the level of perception and helping input to become conscious intake through formal explicit instruction and guided ear training,
2) acquisition facilitator - forming adequate representations and preventing the mapping into L1 system owing to the conscious analysis of the underlying process,
3) monitoring device - providing reflective feedback on the production by equipping L2 learners with necessary tools for self-monitoring and self-correction as well as promoting conscious awareness of the influencing potential of socio-and psychological factors.

Indeed, the findings of previous work by Wrembel (2004) and Dziubalska-Kołaczyk (the studies on phonostylistic processes in L2 of English; Dziubalska-Kołaczyk, 1990) provide some empirical evidence for the claim that metaphonological awareness can aid the acquisition of L2 phonetic features.

The influence of metalinguistic awareness on L2 phonetic performance was also examined by Ramírez Verdugo (2006). The researcher focused on the effect of adopting a multisensory teaching approach on the prosodic performance of Spanish learners of English. Metalinguistic information, oral models and phonetics software with visual display were used to raise the participants awareness of different intonation contours in English. Two groups of Spanish learners, the experimental group who received the treatment and a control group, participated in pre- and post-test recordings, which were acoustically analysed and assessed by native speakers of English. The results demonstrated a significant improvement in the prosodic

performance of the experimental group, providing further support for the claim that formal instruction can facilitate the acquisition of L2 pronunciation.

Nowacka (2008) conducted a detailed longitudinal study aimed at examining the acquisition of different aspects of English pronunciation as a result of formal phonetic instruction. The subjects were Polish learners of English studying at the Teacher Training College in Rzeszów, who underwent 180 hours of formal phonetic training (both theory and practice) over the period of three years. Data were collected in four recording sessions during which the participants produced free and read speech. The phonetic variables under investigation were 41 English pronunciation features, based on the list of the most frequent Polish pronunciation problems by Szpyra-Kozłowska, Frankiewicz and Gonet (2002). The results revealed an overall significant improvement in the subjects' pronunciation over the three-year period. It was found that half of the analysed pronunciation features (22 out of 41) became significantly more target-like and that subjects' production of consonants improved more markedly than their production of vowels and suprasegmental features. Although the findings seem to provide further support for the claim that formal instruction may facilitate the acquisition of L2 pronunciation, they also indicate that phonetic training does not enhance the acquisition of different pronunciation features to the same extent.

Kennedy and Trofimovich (2010) investigated the relationship between the acquisition of the L2 sound system and pronunciation awareness. The subjects were learners of English as a foreign language who underwent a 13-week university-level pronunciation course that focused on the suprasegmental aspects of English pronunciation. Subjects' phonetic performance was assessed through listener-based ratings of accentedness, comprehensibility and fluency. Language awareness was measured though dialogue journal entries. Subjects' pronunciation was evaluated during week 1 and week 11 of the pronunciation course; the journal entries were written on a weekly basis. The comments were analysed for quantitative aspects (language learning as assimilating a set of discrete items) and qualitative aspects (language learning as a meaningful context in which learning occurs). The results of the study showed a relationship between the subjects' pronunciation ratings and the number of qualitative language awareness comments, i.e. participants who made more qualitative language awareness comments were rated higher in terms of phonetic performance. Moreover, the findings highlight the importance/role of input in the acquisition of L2 pronunciation. It was found that the students who produced the most qualitative language awareness comments were those who reported the largest amount of L2 listening done outside of class.

The relationship between the acquisition of the L2 sound system and pronunciation awareness was also analysed in a subsequent study by Kennedy, Blanchet and Trofimovich (2014). The subjects were adult learners of French as a second language, who underwent a 15-week listening/speaking course. The course was focused on segmental and suprasegmental features of pronunciation; it included connected speech processes, intonation, and fluency. Subjects' phonetic performance was evaluated at the beginning and end of the course. In order to measure pronunciation awareness, the informants were required to write weekly journal entries. The results revealed a significant improvement in learners' segmental production and some aspects of connected speech, intonation and fluency. Several variables were associated with pronunciation awareness measures.

Couper (2011) examined the influence of two specific aspects of formal pronunciation teaching: critical listening and socially constructed metalanguage (which is described as metalanguage developed by students working together with the teacher using already understood first language concepts to help in the formation of target language phonological concepts). The subjects were groups of high-intermediate level adult students, each of which received 45-50 minutes of instruction on pronouncing syllable codas. The instructions were different for each group (instructions that involved both critical listening and socially constructed metalanguage, instructions that included one of these types of pronunciation instruction or instructions that involved none of the two). The results showed that the usage of socially constructed metalagnauge significantly improved the subjects' production, while the usage of critical listening had a significant effect on perception.

Porzuczek (2012) analysed speech timing in the pronunciation of Polish learners of English before and after two semesters of pronunciation training. The subjects were first-year students of the English at a teacher training college in Poland, who were required to read a passage in English. Passage reading was recorded at two points in time, i.e. at the beginning and at the end of the first year of phonetics training. The data collected from the Polish learners were compared to speech samples provided by a native-speaker reference group. The results showed improvement in speech timing following practical phonetics training, thus providing further support for the claim that there may exist a relationship between formal pronunciation instruction and L2 phonetic performance.

Recently, Gralińska-Brawata (2013) examined timing organisation in the speech of Polish learners of English by focusing on durational characteristics of vocalic and consonantal intervals. The participants were students of English studies recruited from the University of Lodz. The subjects were asked to read a text passage in English, which was recorded prior to

and following several months of formal pronunciation training. In the interval between the two recordings, the participants attended numerous courses in English as part of the programme of their studies. Gralińska-Brawata (ibid.) hypothesised that increased language experience and phonetic training will lead to significant improvement in timing organisation (as compared with speech samples provided by a native-speaker reference group). Six different rhythm measures were applied to investigate timing organisation. Although the results of the study showed no global progress in the realisation of English rhythm, some instances of improvement were observed in the case of individual speakers and rhythm measures. The findings provide partial support for the assumption that formal pronunciation instruction can aid the acquisition of L2 phonetic features and point to a more complex relationship between phonetic training and non-native phonetic performance.

2.7. Attitudinal factors in the acquisition of L2 pronunciation

The body of research described in this section focuses on learner attitudes towards different varieties of English (native vs. non-native). Although none of the studies examine actual phonetic performance, they are discussed in this chapter for two reasons. Firstly, the results of sociolinguistic studies on non-native pronunciation point to the importance of attitudinal factors in the development of the L2 sound system (e.g. Gatbonton, 1975; Zuengler, 1982; Gatbonton et al., 2011; see Section 2.4.). Secondly, learner attitudes towards native and accented English are relevant to the study since it concentrates on speech convergence towards two pronunciation varieties: Standard Southern British English and Polish-accented English (see Chapter Three). Also, research on L1 speech convergence suggests that attitudinal factors may affect the magnitude of pronunciation shifts (see Chapter One).

One of the first studies on learner attitudes towards different English varieties was conducted by Matsuura, Chiba and Yamamoto (1994), who asked Japanese college students to evaluate the pronunciation of six Asian learners English and a native speaker of American English. Statistical analysis of the data showed the subjects viewed American English more positively than the remaining varieties. The findings were expounded on in a subsequent investigation (Chiba, Matsuura and Yamamoto, 1995). The subjects were over 150 Japanese university students majoring in English or international business. The participants listened to a short English passage read by nine different speakers: three native speakers of English and six non-native speakers (from Japan, Sri Lanka, Hong Kong and Malaysia). The subjects' task

was to indicate their impression of each speaker among ten sets of adjectives arrayed in bi-polar rating scales. Overall, the results revealed that the Japanese university students favoured native over non-native accents. It was also found that the informants who considered English to be important primarily for communication showed relatively fewer negative attitudes towards non-native pronunciation. Additionally, the results showed that learners that exhibited supportive reactions to American or British culture and language were less approving of non-native accents.

Dalton-Puffer, Kaltenboeck and Smit (1997) examined attitudes towards native and non-native varieties of English in Austria. The informants were 132 university students of English (most of them native speakers of German). The subjects were required to evaluate the pronunciation of three native speakers (speakers of Received Pronunciation, near-RP and General American) and two Austrian learners of English (the Austrian speakers were rated to have weak but recognisable foreign accents). The five speakers read a short text on the topic of bilingualism. The participants were told that the purpose of the experiment was to find speakers for the publishing of an audio-book on child language development. The study used a questionnaire that contained a list of 12 attributes which reflected status and solidarity values; the subjects were instructed to indicate to what degree an attribute applied to a given speaker. The list of adjectives was followed by statements: *I think this person would be a good radio presenter*, and *I would like this person as a friend*. The results revealed an overall preference for the three native accents (the RP speaker received the highest ratings). Examination of subjects' responses to the statement *I would like this person as a friend* showed no clear differences between native and non-native accents. However, in the case of *I think this person would be a good radio presenter,* RP was evaluated more favourably than the remaining native and non-native accents. These findings indicate that although native and non-native pronunciation was not clearly distinguished on a solidarity level, clear distinctions were made between the accents with respect to their social status. Additionally, Dalton-Puffer et al. (ibid.) observed that the respondents gave higher ratings to accents with which they have become familiar at school and/or during stays in English-speaking countries.

A preference for native-like pronunciation was also reported by Scales et al. (2006), who asked 37 learners of English (most of them from Asia) to evaluate the pronunciation of four speakers with different accents: General American, British English, Chinese English, and Mexican English. Interestingly, although the majority of the participants stated that their goal was to sound like a native speaker of English, only 1/3 were able to correctly identify the American accent. Generally, the results indicated that the informants preferred accents that

they found easy to understand. Similar results were obtained by Kawanami S. and Kawanami K. (2010). The study investigated the attitudes of 22 Japanese learners of English (university students living in Hawai‘i), who were divided into two groups according to their level of L2 proficiency. The subjects were required to evaluate the pronunciation of six speakers (two native speakers and four non-native speakers), who read a short passage in English. After listening to all six speakers, the subjects were asked to choose which accent they liked the most and which they found easiest to understand. The findings revealed a preference for native accents regardless of proficiency levels. However, it was also found that more proficient learners were more tolerant towards non-native varieties. Similarly as in the study by Scales et al. (2006), Kawanami S. and Kawanami K. (2010) found a correlation between accents that were preferred and those that were considered easier to understand.

Rather than ask the informants to evaluate different accents, a number of studies have investigated learners' opinions on English pronunciation by means of attitude surveys. For instance, Timmins (2002) collected 400 questionnaire responses from learners of English with diverse L1 backgrounds. The informants were presented with the following two statements (ibid. 242):

> Student A: 'I can pronounce English just like a native speaker now. Sometimes people think I am a native speaker.'
> Student B: 'I can pronounce English clearly now. Native speakers and non-native speakers understand me wherever I go, but I still have the accent of my country.'

The participants' task was to decide whether they would prefer to be like Student A or Student B. The results showed that almost 70% of the subjects selected the former option, which indicates that most of them wished to conform to the L2 norm when speaking English.

Janicka, Kul and Weckwerth (2005) studied attitudes towards different English pronunciation models. The subjects were over 200 Polish students majoring in English who were recruited from the School of English at Adam Mickiewicz University in Poznań. The subjects completed a survey designed to investigate opinions on teaching of different English accents. The survey included questions such as *Do you think there is a need to teach any accent at all?* and *Would you as a teacher teach a specific pronunciation*? The findings indicated that the informants generally preferred native models to be used in pronunciation teaching.

Waniek-Klimczak and Klimczak (2005) compared English pronunciation preferences of two groups of Polish learners: students majoring in English and students majoring in economics and sociology. Due to dissimilarities in expected language use in the future and language experience, the authors expected that the two groups would express different attitudes towards English pronunciation. The survey completed by the participants included questions that pertained to the importance of different aspects of speech (grammar, vocabulary and pronunciation) and preferred pronunciation models. The results showed that although both groups stated they believed pronunciation to be important, English majors attached greater importance to pronunciation than economics and sociology majors. Interestingly, both groups exhibited a strong preference for British English with respect to reference variety.

Waniek-Klimczak, Rojczyk and Porzuczek (2014) conducted a large-scale questionnaire study among Polish learners of English. The participants were over 500 students of English, who were recruited from Polish state universities, teacher training colleges, state schools of professional education and one private college. The subjects were enrolled in English programmes at the BA or MA level. The questionnaire items analysed in the study included the following Likert-scale statements (ibid. 29):

> I care about my pronunciation in English being fully correct.
> I think that my pronunciation in English DOES NOT contain features characteristics for Polish pronunciation.
> I care about my English pronunciation NOT having features characteristic for Polish pronunciation.
> I care about my pronunciation in English signalling that I am from Poland.

Waniek-Klimczak et al. (ibid.) found that the majority of the respondents wished their pronunciation to be correct. Given that over 90% of the subjects declared that they would like to speak English in a native-like manner, the researchers argued that correctness was associated with the L2 pronunciation norm. It was also found that most of the informants preferred Polish features not to be present in their speech and did not want their pronunciation to signal that they come from Poland, which was interpreted to mean that the learners expressed negative attitudes towards an ethnic identity function of their accents. Interestingly, the results revealed that BA level students were significantly more concerned with their pronunciation not having Polish features than MA level students, which suggests that attitudes towards L2 pronunciation may change with language experience.

Taken together, the results of the studies reviewed in this section indicate that L2 learners taught in institutional and classroom settings tend to express negative attitudes towards foreign-accented speech and often show a strong preference for native-like pronunciation.

2.8. Summary

The findings discussed in this chapter, although varied in terms of methododology and approaches, allow one to draw some important conclusions about the formation and development of the L2 sound system. The assumptions of the interlanguage framework imply that speakers' productions in a second language are a product of an independent, self-governed linguistic system, which does not correspond exactly either to the L1 or the L2 of the learner (Selinker, 1976, 1992). Indeed, acoustic measurements have shown that non-native speakers may often exhibit values that match neither the L1 nor the L2 pronunciation norm (e.g. Flege, 1980, 1987, 1991; Mack, 1982; Schmidt and Flege, 1996). Major (1987, 2001, 2008) argues that IL is an amalgam of L1, L2 and language universals, the magnitude of their influence depending on the stage of IL's development. Language universals have been found to affect learner's phonetic performance by, for instance, Waniek-Klimczak (2002, 2005) and Piotrowski (2013). The effect of L1, on the other hand, is clearly visible in the relationship between L2 speech perception and production. The findings concerned with the perception-production link suggests that the L1 sound system is of vital importance in the formation of L2 phonetic categories. It has been claimed that with age, perception of speech sounds becomes increasingly language-specific and homed in on the categories typical of the speaker's L1 (Kuhl and Iverson, 1995). This process may cause difficulties with the recognition and discrimination of foreign speech sounds, which, in turn, may reduce the ability to produce the sounds of a foreign language in a native-like manner (e.g. Flege, 1987; Flege and Eefting, 1987; Rochet, 1995; Flege et al., 2003).

Another important characteristic of IL is the fact that it is not fixed and unchanging, but restructures itself as the learner gains more L2 knowledge, gradually approximating the TL norm (Major, 1987, 2001, 2008). This feature of the IL system has been verified in some L2 pronunciation studies (e.g. Dickerson, 1975; Flege, 1980; Wode, 1981; Major; 1987). It has also been suggested that IL shares many characteristics with natural languages and that it can be treated and examined similarly (Adjemian, 1976; Tarone, 1979). Indeed, the results of a number of socio-phonetic studies indicate that L2 pronunciation can be conditioned by the

same social and psychological factors as L1 speech. One of the sociolinguistic variables found to have some (albeit not entirely straightforward) influence on non-native pronunciation, is speech style (alternatively referred to as attention to language form or monitoring; e.g. Dickerson, 1975; Gatbonton, 1975; Dickerson and Dickerson, 1977; Beebe, 1980; Schmidt, 1987). The work by Anisman (1975), Gatbonton (1975), Thompson (1975, 1991), Zuengler (1982), Schmidt (1987), Adamson and Regan (1991) and Gatbonton et al. (2011) implies that L2 phonetic performance can also be modified by the speaker's social group affiliation and sense of identity. In other words, *how* the learner speaks seems to be affected by *who* they identify with. Importantly, the results obtained by Gatbonton (1975), Zuengler (1982) and Gatbonton et al. (2011) suggest that attitudinal factors may be responsible for some of the variability that is present in L2 pronunciation (attitudes towards different language communities were found to affect L2 productions). Based on this observation, it could be hypothesised that attitudes towards native and non-native varieties of English (which constitute an important element of the current investigation, see Chapter Three) can affect L2 speaker's phonetic performance. Studies that deal with learners' opinions on different L2 accents indicate that L2 learners (at least those who were taught in an institutional rather than a naturalistic setting) tend to favour native varieties and have negative attitudes towards foreign-accented speech, including their own non-native variety (e.g. Chiba et al., 1995; Dalton-Puffer et al., 1997; Timmins, 2002; Janicka et al., 2005; Waniek-Klimczak and Klimczak, 2005; Waniek-Klimczak et al., 2014).

The variety of social-psychological, psycholinguistic and linguistic factors that seem to shape the IL and, by extension, the development of the L2 sound system, have been found to create a complex pattern of interaction. For instance, the effect of L2 category perception on non-native pronunciation appears to be related to variables such as cross-language phonetic similarity, L2 language experience and the quality and amount of L2 input. Generally, it seems that L2 sounds which are "similar" (either in terms of perception or articulation) to those found in the L1 may be more difficult to master than L2 sounds that are "new", i.e. do not have close equivalents in the L1 (e.g. Flege, 1987; Flege et al., 2003). Previous research suggests that the acquisition of L2 pronunciation can be facilitated by increased L2 language experience and native-speaker input (Suter, 1976; Flege, 1987; Flege and Liu, 2001; Flege et al., 2003). Another key element in the acquisition of the L2 sound system is the age of learning: it has been observed in numerous studies that late learners are often unable to achieve native-like performance in L2 pronunciation (e.g Asher and García, 1969; Oyama, 1976; Suter, 1976; Tahta et al., 1981; Flege, 1988; Thompson, 1991). Other factors that have

been found to affect L2 phonetic performance include: linguistic environment (Adamson and Regan, 1991;), phonetic talent (Purcell and Suter, 1980), concern for pronunciation (Dickerson, 1975; Gatbonton, 1975; Suter, 1976; Beebe, 1980; Elliot, 1995; Thompson, 1975), the use of formal pronunciation instruction (e.g. Dziubalska-Kołaczyk, 1990; Derwing et al., 1998; Wrembel, 2004; Nowacka, 2008; Kennedy and Trofimovich, 2010) and certain psychological variables such as motivation (Gardner and Lambert, 1972; Gardner, 1985, Schumann, 1978, 1980), ego permeability (Guiora et al., 1972; Schumman, 1978, 1980), tolerance to anxiety and involvement in emotional experiences (Taylor et al., 1971).

Chapter Three: Speech convergence in the pronunciation of Polish learners of English - the study

3.1. Introduction

The chapter describes the study on speech convergence in the pronunciation of Polish learners of English, which is the central point of this dissertation. As referred to in the Introduction, the term speech convergence (or phonetic convergence) refers to the general phenomenon of adapting one's linguistic behaviour depending on who one is talking or listening to. Two types of speech convergence will be mentioned in this chapter: imitation and accommodation. The term imitation denotes speech convergence that is analysed in a non-interactional setting; the term accommodation is applied in relation to speech convergence examined in an interactional setting (see Introduction and Chapter One). Section 3.2. of the chapter provides a review of previous research on speech convergence in non-native productions. Section 3.3. discusses the rationale and aim of the study. Section 3.4. is concerned with the pilot work that provided the basis for the current investigation. Section 3.5. presents the main assumptions of the study and the hypotheses that will be tested in the course of the investigation. Section 3.6. describes the methodology. Research questions that were formulated for the purposes of the study are listed and discussed in Section 3.7. Finally, Section 3.8. presents the results of the study.

3.2. Speech convergence in L2 pronunciation

The findings of previous research on L2 pronunciation (as discussed in Chapter Two) indicate that the IL sound system shares many characteristics with the pronunciation of an L1 and thus can be treated and examined similarly. What follows from this assumption is that speech convergence phenomena that take place in a speaker's native language should also be present, at least to an extent, in the speaker's second language. The first studies on speech convergence in L2 pronunciation were inspired by the findings of Giles and colleagues (Coupland, 1984; Bourhis and Giles, 1977; Giles, 1973; Giles, Taylor and Bourhis, 1973); and examined accent shifts in conversational interactions between native and non-native speakers. For instance, Beebe (1977) studied the Thai pronunciation of Chinese-Thai bilinguals in conversations with Chinese and Thai interlocutors. She observed that the

subjects used significantly more Thai variants when speaking to Thai interviewers; they also chose Chinese variants more frequently in conversations with Chinese interlocutors. Similar results were obtained for Chinese-Thai children (Beebe and Zuengler, 1985). The findings imply that in both studies, the participants were accommodating their pronunciation towards their interviewers. Interestingly, accommodation took place even when the interviewers did not use the investigated phonetic variants in their speech. As argued by Giles et al. (1987), speakers may accommodate towards the communicative behaviour they *believe* others to exhibit or converge towards what they believe others *expect* them to produce. Perceived ethnicity was also found to affect L2 pronunciation in a study by Sawyer (1973). The subjects were Mexican-American speakers of English who were found to converge towards Anglo and Hispanic interlocutors. More specifically, whenever the informants needed to use a Spanish word, they anglicised its realisation when interacting with an English interlocutor. Conversely, they maintained the native Spanish pronunciation of the words when in conversation with a Hispanic interlocutor.

Zuengler (1985, 1989) focused on the effect of status imbalance in conversational interactions between native and non-native speakers of English. The subjects were female Spanish learners of English, grouped into dyads with female native speakers of English. Status was operationalised as relative expertise in aesthetic perception (which was based on scores from an aesthetic judgement test that th subjects were required to take in the first phase of the experiment). The Spanish-English dyads were assigned to one of two experimental conditions or to the control group. In the first condition, the dyads were told that the native speaker scored low on the aesthetic perception test and that the non-native talker had gotten a high score. The situation was reversed in the second condition; it was the native English speakers who were assigned the role of the "experts". In the control group, dyads were not told how they performed in the test. In the main part of the experiment, each dyad had two 10-minute conversations about the pictures in the aesthetic perception test, one before and one after being told how they performed on the test. The analysed phonetic variables included: /ð/, /r/, word-final consonant clusters, the STRUT vowel and the TRAP vowel. The results of the study showed that the group of non-native "nonexperts" decreased in their TL realisations of /r/ in the second conversation as compared with the first conversation. Zuengler (1991: 232) accounted for these findings by arguing that "[...] if NNS tend to be treated as subordinate in general, being assigned nonexpert status could cause them to simply give up aiming for correctness, thus displaying a kind of passive divergence". Other phonetic variables that were investigated in the study did not undergo significant shifts. Also, no significant differences

were found in the pronunciation of the non-native "experts". These findings suggest that the magnitude of convergence in L2 speech may be related to social status and the phonetic variable under investigation. It should also be mentioned that divergence from the TL norm was found in Zuengler's (1982) earlier study on the effect of ethnolinguistically-threatening questions (see Section 2.4.). Some of the subjects produced less native-like phonetic variants when replying to a remark that threatened their ethnic identity, which Zuengler (ibid.) attributed to a desire to defend their ethnic solidarity.

Young (1988) analysed the production of regular plural marking by Chinese speakers of English residing in the United States. The participants were divided into two groups based on their proficiency in English and were interviewed by two interlocutors, one a native speaker of English and one a Chinese speaker of English. Although the study was concerned with morphology rather than phonetics, the CAT framework was be used to explain some of the observed variability in L2 performance (which makes the study relevant to the present investigation). An important variable in Young's (ibid.) study was the degree of social convergence between the participants and their interlocutors, which was measured by calculating the number of attributes (ethnicity, gender, occupation, educational level, place of origin, age) the interviewers and the interviewees shared. Young (ibid.) hypothesised that the degree of social convergence with the interlocutor will be correlated with the magnitude of accommodation. The results of the study indicated that only the group of high proficiency subjects who talked to the native English speaker behaved as predicted (i.e. converged towards the native interlocutor by increasing the proportion of TL variants in their speech). Young's (ibid.) findings suggest that the degree of convergence in non-native speech may be modified by whether or not the learner identifies with their interlocutor as well as the learner's level of proficiency in the L2.

Berkowitz (1986) investigated the pronunciation of Dominican learners of English in a conversational interaction with a native speaker and observed that the interlocutor's perceived cultural empathy influenced the subjects' production of several phonetic variables. The findings showed that the more the participants perceived cultural empathy on the part of their interviewer, the less TL variants of final consonant clusters they produced. Interestingly, the opposite pattern was observed for two other phonetic variables: the realisation of /r/ and /s/-initial consonant clusters. These results are consistent with Zuengler's (1985, 1989) findings that the magnitude of convergence may differ as a function of phonetic variable.

Recently, L2 phonetic convergence has been examined within a social-psychological framework by Lewandowski (2012). The researcher focused on the issue of phonetic ability

and its effect on speech accommodation. The subjects were advanced German learners of English who were categorized into three groups on the basis of their level of phonetic talent (the participants were selected from the subject pool of a preceding project during which they had been extensively tested on their phonetic abilities). The German learners participated in two consecutive spoken interactions with a native speaker of American English and a native speaker of Standard Southern British English. Acoustic measurements were based on the extraction of amplitude envelopes from the speech signal at different points within conversational interactions. As explained by Lewandowski (ibid.), amplitude envelopes reflect the amount of energy present in the separate frequency bands of the acoustic signal. The results revealed that the level of phonetic ability had a significant effect on the amount of convergence in the subjects' pronunciation. More talented learners were found to accommodate towards native speakers to a greater extent than less talented learners.

Another recent study dealing with accommodation in native-non-native interactions was carried out by Kim et al. (2011). The study is discussed in more detail in Chapter One (see Section 1.3.), as it concentrated on the broader issue of the effect of language distance on phonetic convergence. The analysed conversational interactions took place between subjects who had either the same or different regional dialects, and between native and non-native speakers of English. The results indicated that it was only the pairs of speakers with the same L1 who accommodated towards each other; no significant effect of convergence was found for the native-non-native dyads. Failure to accommodate on the part of the non-native speakers was attributed to high attentional demands and processing load involved in native-non-native communication. It is also possible that no convergence was observed due to the nature of the experimental procedure and data analysis. The magnitude of accommodation was measured with the use of an AXB perceptual similarity task (see Section 1.3.), where the A and B stimuli were speech samples of a given member of a dyad taken from early and late portions of the recorded conversations and X was the interlocutor's speech sample from either early or late stage of the interaction. This means that the stimuli that were contrasted were comprised of different strings of words. Therefore, it is conceivable that no convergence effect was found for the non-native speakers because the procedure did not allow for a precise enough examination of the data.

There have also been a number of recent studies on L2 speech accommodation that investigate the phenomenon in classroom setting. Waniek-Klimczak (2009a) examined the English pronunciation of three teacher trainees (native speakers of Polish teaching English). Phonetic performance was analysed under two conditions: in interaction with learners during

lessons and in interaction with a native speaker of English in a constrained, formal context. The phonetic variables under investigation were: consonant voicing, aspiration, vowel length, rhythmicity and intonation. Auditory analysis of the data suggested that the participants converged towards the native speaker by producing more target-like speech and converged towards the Polish students by using more heavily accented speech. Some degree of divergence was also observed in the teacher-student interactions; it was attributed to hypercorrection resulting from the characteristics of classroom communication. Importantly, the findings of the study call attention to an interesting issue that was not explored in previous studies on L2 speech convergence, i.e. the analysis of convergence towards native speakers of the target language as compared with convergence towards other learners of the TL.

Trofimovich and Kennedy (2014) focused on speech accommodation (which they refer to as alignment) in classroom interactions between speakers communicating in a shared L2. The participants were 30 learners of English with different L1 backgrounds, who engaged in two interactive speaking tasks. Excerpts from the beginning and end of the tasks were presented to a group of native English listeners, who were instructed to rate whether a given pair of participants sounded similar in a particular excerpt. Statistical analysis of the data indicated that pairs of participants were rated as more similar-sounding at the end than at the beginning of the conversation, which indicated that the participants converged towards one another.

Trofimovich, McDonough and Foote (2014) focused on the accommodation of English stress patterns in a classroom setting. The subjects (students enrolled in a university-level English-for-academic-purposes class) participated in four interactive speaking tasks (information exchange quizzes), which involved using multisyllabic words with stress on the second syllable (e.g. *consistent, intelligent*). Convergence (referred to as alignment) was operationalised as higher accuracy rates in discourse contexts where a conversational partner previously produced an accurate target stress. The results of the study suggested that the subjects converged towards each other on stress placement in all four tasks. Taken together, the data obtained by Trofimovich and colleagues imply that phonetic convergence can take place also when L2 speakers are communicating with other learners.

Several studies on phonetic imitation in non-native speech have also been conducted recently. For instance, Rojczyk (2013) examined the production of the English TRAP vowel by Polish learners. The vowel was selected for analysis because it tends to be assimilated to Polish /a/ or /e/ by Polish native speakers. The participants were students recruited from the University of Silesia, whose proficiency in English ranged from intermediate to upper-intermediate. The stimuli used in the experiment were monosyllabic English words containing

the vowel /æ/, which were recorded by a male speaker of Southern British English. The experimental procedure consisted of two phases. First, the subjects were required to read the monosyllabic English words from a computer screen. In the imitation task, they were instructed to immediately repeat the stimuli provided by the native speaker. A comparison of participants' productions from the reading task (baseline) with the realisations from the imitation task revealed significant convergence with the native English model talker. Although the effect of gender was also examined in the study, Rojczyk (ibid.) found no significant effect of this variable on the magnitude of convergence (however, the author admits that the results might have been biased by the fact that male participants were underrepresented in the study).

The aim of another imitation study by Rojczyk (2012) was to test Strange's (2011) Automatic Selective Perception (ASP) model. The phonetic variable under investigation was VOT duration in /p/-, /t/- and /k/-initial English words. The subjects were students of English studies recruited from the University of Silesia, who participated in three experimental tasks: reading a word list (baseline), immediate shadowing after the native English model talker and distracted shadowing after the native English model talker. In the final task, the participants were instructed to listen to a word, read a number presented on the computer screen, and then imitate the word. The results demonstrated that the learners significantly increased VOT values in both the immediate and the distracted imitation task as compared with the baseline (however, the effect was smaller in the distracted imitation task). These findings were interpreted to indicate that "[...] immediate imitation may bypass the influence of native articulatory habits [...]" and that " [...] distraction in imitation results in partial recovery of native phonetic patterns" (Rojczyk, Porzuczek and Bergier, 2013: 5). It was also found that the place of articulation of the voiceless stop influenced the magnitude of convergence. Although gender was observed to have some effect on imitation, once again, the male and female populations were unbalanced.

Rojczyk, Porzuczek and Bergier (2013) focused on Polish learners' immediate and distracted imitation of release burst in English stop sequences. Unlike in English, stops are usually released in Polish, which is an L1 habit that tends to be transferred to L2 pronunciation. The experimental procedure in Rojczyk et al.'s (ibid.) study was much the same as in the two previous investigations. The subjects were native Polish students recruited from the Institute of English at the University of Silesia. First, they were required to read a list of phrases containing English stop sequences that were presented on a computer screen. Next, the subjects were instructed to repeat the phrases after a native English model talker. Finally,

they were instructed to listen to the model talker's voice, read a number from the computer screen, and then imitate the phrase produced by the model talker. The findings of the study were broadly consistent with the results obtained by Rojczyk (2012). The mean durations of release bursts in the subjects' productions were significantly reduced in the immediate imitation task as compared with the baseline, which implies that Polish learners converged towards the native English speaker. The release bursts produced in the distracted imitation task were generally lower than in the baseline but the effect was not significant, which suggests that distracting the participants may somewhat reduce imitative tendencies. It was also found that imitation was selective with respect to phonetic environment, i.e. release burst in homorganic clusters was imitated to a greater extent than in heterogenic clusters, which was attributed to the fact that stops in homorganic clusters can be optionally released in the subjects' L1.

Taken together, the results of the studies reviewed in this section indicate that speech convergence phenomena that have been found to take place in a speaker's L1 are also present in L2 speech. Not unlike the corresponding process in L1, L2 speech convergence appears to be selective from both a linguistic and a social-psychological standpoint. It has been found that the magnitude of convergence in non-native speech may be conditioned by the speaker's social status in an interaction (Zuengler, 1985, 1989), the strength of ethnic affiliation (Zuengler, 1982), the level of social convergence between the interacting partners (Young, 1988), phonetic ability (Lewandowski, 2012), level of language proficiency (Young, 1988) and phonetic context (Zuengler, 1985, 1989; Berkowitz, 1986; Rojczyk, 2012; Rojczyk et al., 2013).

3.3. Aim of the study

L2 phonetic convergence appears to be a fruitful area of research that has not yet been thoroughly investigated. One of the elements that seems to be missing from previous work on L2 speech adjustments is a more careful examination of pronunciation shifts upon exposure to the speech of native speakers of the TL as compared with pronunciation shifts upon exposure to the speech of other learners. Thus, the aim of the study is to investigate and compare L2 convergence strategies upon exposure to native and non-native pronunciation. In this dissertation, the term convergence strategies encompasses three types of linguistic behaviour: convergence, divergence and maintenance (see Introduction and Section 3.5.). The study

concentrates on the phonetic performance of advanced Polish learners of English, who are exposed to two pronunciation varieties: Polish-accented English and native English.

3.4. Pilot studies

L2 convergence strategies upon exposure to native and non-native pronunciation were examined in two pilot studies (Zając, 2013; Zając and Rojczyk, 2014). Both studies used a non-interactional, laboratory-based design and investigated speech convergence patterns in the pronunciation of Polish learners of English. In Zając (2013), the native model talker was a speaker of Standard Southern British English (SSBE), while the non-native model talker was a Polish learner of English who spoke with a noticeable foreign accent (a judgement based on auditory analysis by the author of the study). The phonetic variable under investigation was vowel duration as a cue for consonant voicing. The words selected as stimuli included the following minimal pairs: *bad-bat, bed-bet, bead-beat, bid-bit.* The native model talker used noticeably longer vowels in the voiced context in each word pair, whereas the non-native model talker used shorter vowels in the voiced environment in the *bead-beat* and the *bid-bit* pairs. The subjects were first-year students of English Studies recruited from the University of Lodz, Poland. The experimental procedure was divided into two phases: an auditory naming task (the baseline), and the imitation task. In the first phase, a set of photos was presented sequentially on the computer screen. The participants' task was to identify what they see in the photos. In the imitation phase of the experiment, the same photos were presented on the computer screen, this time accompanied by either a native or a non-native model talker's voice. The subjects were instructed to first listen to the voice and then name what they see in the photos (they were never explicitly instructed to imitate the speech they hear). The results revealed that the participants increased duration contrasts in the *bit-bid* pair in the imitation task as compared with the baseline. This was interpreted to indicate that the subjects converged towards the native English speaker and diverged from the native Polish speaker. Convergence strategies of the participants were ascribed to a desire to sound native-like, which was likely related to the fact that the experiment took place in an institutional setting. Firstly, all the participants were students at the Institute of English Studies and were expected one day to become expert language users. Secondly, they attended an English phonetics and phonology course where SSBE was presented as the preferred pronunciation model. In addition, the study was conducted at the building of the Institute of English Studies and the

subjects' performance was monitored by the author of the study (whom the informants knew as an academic teacher). Thus, it would appear that the situational context of the experiment may have had a bearing on the learners' convergence strategies. Furthermore, since the subjects were neither explicitly instructed whether or not to imitate the speech they heard nor were they asked to comment on their speech behaviour after the experiment had ended, it remains unclear whether they converged towards the native speaker and diverged from the non-native speakers as a result of a bias against foreign-accented speech or, given the institutional setting, simply because they believed this type of speech behaviour was expected of them.

Zając and Rojczyk (2014) replicated the previous study (Zając, 2013) in order to expand on its findings and determine whether the provision of explicit instructions to imitate may affect convergence strategies. The stimuli used in the experiment were monosyllabic English words with the front vowels /æ e ɪ iː/ flanked by word-initial /b/, /m/ or /s/ and word-final /t/ or /d/. They were recorded for the imitation task by a native speaker of Standard Southern British English and a native speaker of Polish (a qualified phonetician imitating Polish accent). The native English model talker provided vowels that were considerably longer in the voiced context, whereas the Polish model talker produced similar durations of vowels before voiceless and voiced stops. As in the previous study, the experiment took place in an institutional setting: the participants were first-year students at the Institute of English, University of Silesia, who were enrolled in an pronunciation course taught by the second author of the study. The experimental procedure consisted of two phases: a reading task designed to establish participants' baseline productions and an imitation task where the subjects repeated the analysed words after the model talkers. Half of the participants took part in the first experimental session in which the words were presented without specific instructions to imitate: the participants were only instructed to wait until the recorded voice stopped producing the word and then read it from the computer screen. The remaining half of the subjects took part in the second session in which they were instructed to imitate the words they heard as faithfully as they could. The results revealed that the informants produced significantly longer vowels before voiced than voiceless stops when imitating both the native and the non-native model talker, which was interpreted as convergence towards the native English speaker and divergence from the native Polish speaker. Interestingly, the type of instructions received in the imitation task was found not to have a significant influence on the magnitude of phonetic convergence, which signifies that the subjects diverged from the non-native model talker even when explicitly instructed to imitate his speech. Similarly as in the

previous study, it seems likely that the convergence strategies evidenced by the participants stemmed from a desire to sound native-like and were related to the fact that the experiment was carried out in an institutional setting. Also, the finding that the learners diverged from the non-native model talker even when explicitly instructed to imitate his speech seems to suggests that it was the situational context of the experiment that had the greatest influence on the subjects' convergence strategies.

The findings of the two pilot studies point to certain methodological issues that are addressed in the present study. Firstly, it was found that the participants diverged from non-native speech, which suggests they may have been biased against foreign-accented pronunciation. The assumption seems plausible given the institutional setting of the experiments and the fact that students of English have been found to favour native over non-native accents in previous studies (e.g. Dalton-Puffer et al., 1997; Janicka et al., 2005; Waniek-Klimczak and Klimczak, 2005; Waniek-Klimczak et al., 2014; see Section 2.7.). However, the hypothesis should be verified by gauging the subjects' attitudes towards native and foreign-accented pronunciation. Secondly, although a tendency to diverge from the non-native model talker was observed in both pilot studies, it seems that it may have resulted from the type of experimental design that was used in the two investigations. It is possible that the informants diverged from L1-accented speech because they wished to make a favourable impression on the experimenters and/or saw the experimental tasks as an extension of the English pronunciation course that was part of their curriculum. The interpretation seems probable in view of the fact that in both pilot studies, the experimenters were known to the subjects as academic teachers of English phonetics and phonology. It is conceivable that the participants would have used different convergence strategies had the experimental design included a non-native interlocutor that would listen to and evaluate their productions, e.g. another Polish learner of English. A possible solution would be to record conversational interactions of Polish learners of English with other non-native and native speakers. Nonetheless, using conversational data may be problematic in the sense that it makes it difficult to control for phonetic context and the number of investigated tokens. Additionally, finding interlocutors willing to devote much of their time and energy to converse with the participants may sometimes prove challenging.

In the present investigation, an attempt is made to extend and improve the controlled experimental design of the pilot studies by: 1. applying a questionnaire to measure the subjects' attitudes towards English pronunciation and ascertain whether they favour native-like realisations, 2. modifying the experimental procedure so that the model talkers act as

interlocutors, thus providing the stimulus for convergence towards both native and foreign-accented English.

3.5. Assumptions and hypotheses

The study is grounded in two large bodies of research: L2 phonetics and work on speech convergence. As discussed in Chapter One, previous findings indicate that speakers tend to adjust their speech when listening to or interacting with others. The process seems to have its origin in a natural human predisposition to imitate actions performed by another (in this case, the sounds produced by another person). The inherent tendency to imitate ambient speech can be sustained (convergence), blocked (maintenance) or reversed (divergence) depending on a variety of social-psychological and linguistic factors. In this dissertation, the three types of linguistic behaviour are referred to as convergence strategies (see Introduction); they can be described in the following manner: convergence is the process of making one's speech more similar to that of another person; divergence is the process of moving away from the speech of another person; maintenance is the process of maintaining one's default linguistic behaviour in spite of exposure to the speech of another person. Convergence strategies may serve as a tool for mediating social distance and/or facilitating communication in an interaction. They may also be affected by attitudinal factors, e.g. bias towards or against a particular social group on the part of the speaker. Additionally, convergence strategies have been found to be sensitive to language structure and to be conditioned by individual speaker differences (see Chapter One).

As discussed in Chapter Two, previous research on L2 phonetics (and second language acquisition) shows that speaker's productions in a foreign language are generated by an independent linguistic system (interlanguage), which contains elements of the learner's L1 and L2 but does not correspond exactly to either NL or TL. Interlanguage is a dynamic system that restructures itself as the learner gains more experience with the L2. The development of IL, and by extension the development of the L2 sound system, has been found to be shaped by a variety of social-psychological, psycholinguistic and linguistic factors.

The main focus of the study are convergence strategies in L2 speech. As discussed in Section 3.2., previous studies have shown that L2 learners may also tend to adjust their speech when listening to or interacting with others. Importantly, the process seems to operate in a similar manner as in L1 speech. Convergence, divergence and maintenance have all been

found to take place when learners are using their L2. The usage of a given strategy and the magnitude of the process appear to be mediated by attitudinal and linguistic factors. Also, previous studies on speech convergence in L2 speech suggest that the process may function as a tool for negotiating social distance and facilitating communication in an interaction.

Based on the aforementioned considerations, the main assumptions in the study are the following:

1. The phenomenon of speech convergence takes place in non-native pronunciation.

2. L2 speech convergence it is motivated and modified by social-psychological and linguistic factors.

3. L2 speech convergence can be analysed and interpreted similarly as convergence in L1 speech.

The study aims to examine L2 convergence strategies following exposure to native and non-native pronunciation. With this objective in mind (and given the aforementioned assumptions), three hypotheses have been formulated to be tested in the study. Hypothesis 1 is the most general of the three and pertains to the overall effect of exposure to two English varieties (native vs. non-native) on subjects' convergence strategies; it assumes that speech behaviour following exposure to native speech will be different than speech behaviour following exposure to non-native speech. Hypothesis 2 and Hypothesis 3 are concerned with specific variables that may influence convergence strategies. Hypothesis 2 pertains to a selected attitudinal factor; it assumes that speech behaviour following exposure to native and non-native speech will be conditioned by participants' attitudes towards English pronunciation. Hypothesis 3 refers to a selected linguistic factor; it assumes that speech behaviour will be affected by phonetic context. The three hypotheses are presented below:

H1: Convergence strategies following exposure to native and non-native English differ as a function of model talker/interlocutor.

H2: Convergence strategies following exposure to native and non-native English are affected by the subjects' attitudes towards native and Polish-accented English.

H3: Convergence strategies following exposure to native and non-native English vary as a function of phonetic context (place of articulation and vowel category).

At this point, it should be noted that the three types of speech behaviour referred to as convergence strategies are operationalised in the following manner:

- convergence: a significant shift towards the values exhibited by a given model talker/interlocutor

- divergence: a significant shift away from the values exhibited by a given model talker/interlocutor
- maintenance: a non-significant difference between the subjects' default realisations and the values exhibited following exposure to the speech of a given model talker/interlocutor

3.6. Method

The section presents the methodology applied in the study of speech convergence in the pronunciation of Polish learners of English. First, the phonetic variables and the stimuli provided by the model talkers/interlocutors are discussed (Sections 3.6.1. and 3.6.2.). Section 3.6.3. provides information about the subjects who participated in the study. Section 3.6.4. discusses the procedure (which uses a new experimental method that merges a controlled experimental setting with an element of social interaction and is based on the experience acquired in the course of pilot work). Section 3.6.5. pertains to acoustic measurements. Section 3.6.6. describes the statistical methods applied in the study.

3.6.1. Phonetic variables

The study examines convergence strategies following exposure to two linguistic varieties: Polish-accented English and native English. The phonetic parameters selected for analysis include: aspiration, pre-voicing in word-initial stops and vowel duration as a cue for consonant voicing. These pronunciation features were selected for analysis since they may be expected to have distinct realisations in the two varieties.

Aspiration and pre-voicing can be described and measured using the Voice Onset Time (VOT) continuum. VOT can be defined as the time interval between the release of the word-initial stop and the onset of vocal fold vibration for the following vowel. The measurement system was introduced by Lisker and Abramson (1964), who were searching for cross-language acoustic features that serve as cues for the voicing of stop consonants in word-initial position. They found that the behaviour of word-initial stops in different languages generally falls into the following phonetic categories:

1. Voiced unaspirated stops, in which voicing begins before the release of the consonant (pre-voicing). Measurements of VOT before the release of the stop are stated as negative numbers and referred to as voicing lead.

2. Voiceless unaspirated stops, in which voicing begins just after the release of the consonant. Here, the VOT measurements are stated as positive numbers and referred to as short voicing lag.

3. Voiceless aspirated stops, in which voice onset lags considerably behind the release. VOT measurements are stated as positive numbers and referred to as long voicing lag.

Both Polish and English have two phonological categories for stop consonants that could be broadly described as voiced and voiceless. However, word-initial stops have considerably different phonetic realisations in each language. English word-initial /p t k/ can be categorised as voiceless aspirated and are generally produced with long voicing lag, while English word-initial /b d g/ can be categorised as voiceless unaspirated and are generally realised with short voicing lag (Lisker and Abramson, 1964; Kopczyński, 1977). In Polish, on the other hand, there is an opposition between a voiceless unaspirated category and a voiced unaspirated category: short voicing lag in word-initial /p t k/ is contrasted with voicing lead in word-initial /b d g/ (Keating, 1980, 1984; Keating et al., 1981). Tables 1-4 show mean VOT values for Polish and English obtained by Lisker and Abramson (1964), Kopczyński (1977) and Keating, Mikoś and Ganong (1981).

stop consonant	VOT value
/p/	+82.5 ms
/t/	+84 ms
/k/	+71 ms
/b/	+18 ms
/d/	+14 ms
/g/	+31 ms

Table 1. Mean VOT values for English word-initial stops (after Kopczyński, 1977: 72).

stop consonant	VOT value
/p/	+58ms
/t/	+70ms
/k/	+80ms
/b/	+1ms
/d/	+5ms
/g/	+21ms

Table 2. Mean VOT values for English word-initial stops (after Lisker and Abramson, 1964: 394).

Given the marked cross-language difference in the realisation of word-initial stops, it is assumed that Polish-accented realisations of English /p t k/ and /b d g/ will either include Polish-like VOT values or values that are intermediate to the VOT values found in English and Polish. Indeed, it was found that Polish speakers of English did not match native-like VOT values by Waniek-Klimczak (2005, 2009b, 2011) and Piotrowski (2013). In a

perception-based study, Rojczyk (2010) found that Polish learners did not match native speakers in categorising VOT values. Also, intermediate VOT values in foreign-accented English were reported, for instance, by Flege (1980) and Flege and Eefting (1987).

stop consonant	VOT value
/p/	+37.5 ms
/t/	+33 ms
/k/	+49 ms
/b/	-78 ms
/d/	-72 ms
/g/	-61 ms

Table 3. Mean VOT values for Polish word-initial stops (after Kopczyński, 1977: 72).

stop consonant	VOT value
/p/	+21.5 ms
/t/	+27.9 ms
/k/	+52.5 ms
/b/	-88.2 ms
/d/	-89.9 ms
/g/	-66.1 ms

Table 4. Mean VOT values for Polish word-initial stops (after Keating et al., 1981: 1262).

It is also expected that VOT values in the productions of Polish learners of English may vary as a function of place of articulation and the quality of the following vowel. It is generally assumed that the further back the closure, the longer the VOT (e.g. Lisker and Abramson, 1964; Cho and Ladefoged, 1999). It has also been observed that VOT may vary depending on the identity of the following vowel: it is generally longer before high vowels than before mid and low vowels (e.g. Klatt, 1975; Summerfield, 1975). The two factors have been found to affect VOT productions of Polish learners of English (e.g. Waniek-Klimczak, 2005).

Another phonetic parameter analysed in the study is vowel duration as a cue for consonant voicing. Although vowels have been generally found to have a tendency to be slightly shorter before voiceless than before voiced obstruents (e.g. Chen, 1970; Lisker, 1974), it has been argued that English exaggerates this universal tendency by rule. Acoustic measurements have shown that English vowels followed by voiced consonants are generally realised as considerably longer than the same vowels followed by voiceless consonants (e.g. Hogan and Rozsypal, 1980) and that the length ratio of one to the other is approximately 3 : 2 (Peterson and Lehiste, 1960). Furthermore, it has been found that vowel duration in English greatly contributes to the perception of voicing of the following final obstruent (e.g. Hogan and Rozsypal, 1980; Raphael, 1972). The following tables show mean vowel durations before voiced and voiceless consonants in English obtained by Chen (1970) and Peterson and Lehiste (1960).

	mean vowel duration		
study	before voiceless consonants	before voiced consonants	mean difference
Chen (1970)	146 ms	238 ms	92 ms
Peterson and Lehiste (1960)	197 ms	297 ms	100 ms

Table 5. Mean vowel durations before voiced and voiceless consonants in English (after Chen, 1970: 138 and Peterson and Lehiste, 1960: 700).

As opposed to English, Polish neutralises the phonological voiced-voiceless contrast between word-final obstruents (Wierzchowska, 1980; Ostaszewska and Tambor, 2000). Thus, vowel length differences before final consonants can be assumed not to be phonologically relevant in Polish. Indeed, Jassem and Richter (1989) found no significant length differences between vowels followed by underlyingly voiced final consonants and vowels followed by underlyingly voiceless final consonants in Polish. Although Slowiaczek and Dinnsen (1985) did report Polish vowels to be longer before underlyingly voiced than underlyingly voiceless obstruents, Table 6. shows that the differences were very slight, especially as compared with the values reported for English. Also, a follow-up perception study revealed that the observed duration contrasts were not functional in perception, i.e. Polish listeners did not use vowel duration in their voicing judgements (Slowiaczek and Szymanska, 1989).

mean vowel duration		
before underlyingly voiceless consonants	before underlyingly voiced consonants	mean difference
118 ms	130 ms	12 ms

Table 6. Mean vowel durations before voiced and voiceless consonants in Polish (after Slowiaczek and Dinnsen, 1985: 333)

Due to the phonological and phonetic differences in the implementation of voicing contrast in word-final obstruents in Polish and English, it is assumed that the English vowel length distinction between following voiced and voiceless obstruents will not be realised in a native-like manner in Polish-accented English. The assumption is supported by some previous findings. Waniek-Klimczak (1999) examined the pronunciation of Polish immigrants to the UK and observed that although the participants did contrast vowel length as a cue for

consonant voicing, they did not implement the durational differences consistently in a native-like way. Similarly, Waniek-Klimczak (2005) found that Polish immigrants to the US exhibited differences in vowel duration as an effect of consonant voicing in their English productions; however, the differences were smaller than for the native speaker control group. In addition, Rojczyk (2008) observed that Polish learners of English resorted to L1 habits and did not match native speakers in their perceptual judgements of vowel duration as a cue for final obstruent voicing.

It is expected that vowel length in the productions of Polish learners of English may be additionally affected by context-independent duration differences: English vowels are traditionally divided into two sets, inherently short and inherently long vowels. For instance, all things being equal, FLEECE is expected to be longer than vowels such as KIT, TRAP or DRESS (e.g. Wells, 1962; Wiik, 1965; see Shockey, 2013 for a discussion of phonemic vowel length in English). Context-independent duration differences in vowels produced by Polish learners of English have been observed, for instance, by Waniek-Klimczak (2009b) and Porzuczek (2012).

3.6.2. Stimulus

The study used the materials recorded for one of the pilot investigations (Zając and Rojczyk, 2014). The stimuli were 48 monosyllabic English words, 24 of which contained the front vowels TRAP, DRESS, KIT and FLEECE flanked by word-initial /b/, /m/ or /s/ and word-final /d/ or /t/. The four vowels were used with a view to conducting a follow-up study on the imitation of vowel quality by Polish learners of English; they were selected on the grounds that assimilating the TRAP vowel with the DRESS vowel and the KIT vowel with the FLEECE vowel are two well-recognised features of a Polish accent in English (e.g. Sobkowiak, 2001; Gonet, Szpyra-Kozłowska and Święciński, 2010; Nowacka, 2010; Weckwerth, 2011).

The tokens containing word-final alveolar stops formed the following voiced-voiceless minimal pairs: *beat-bead, bit-bid, bat-bad, bet-bed, mitt-mid, mat-mad, met-med, meat-mead, seat-seed, sit-Sid, sat-sad, set-said*. The remaining 24 tokens were words with different word-initial stop consonants, arranged into voiced-voiceless minimal pairs followed by the same vowel: *bat-pat, bet-pet, bun-pun, bop-pop, Dutch-touch, dog-tog, dip-tip, Dan-tan, gap-cap, goat-coat, gut-cut, got-cot*.

The target words were recorded by a native speaker of Standard Southern British English (SSBE) and a native speaker of Polish; both speakers were male and of similar age (the native English speaker was in his late twenties, the native Polish speaker in his early thirties). The model talkers/interlocutors were qualified phoneticians; they were given the list of target words and asked to read and record them at their leisure. The author of the study did not interfere in the recording process, thus making the procedure similar to the one used in the experimental phase of the study (the participants were left on their own for the duration of each recording; see Section 3.6.4.). Unlike the subjects, the model talkers/interlocutors were familiar with the purpose of the study and the type of phonetic parameters selected for analysis. The two speakers were told to use natural speaking tempo and falling intonation in each token. The native Polish speaker was additionally asked to imitate a heavy Polish accent for the purposes of the study. Prior to making acoustic measurements, the recordings were examined by the author of the study to ensure high quality.

Aspiration, pre-voicing and vowel duration values in the stimuli provided by the model talkers/interlocutors were measured using Praat speech-analysis software package (Boersma, 2001) by means of waveform and spectrographic display. Aspiration was operationalised as voicing lag values in word-initial /p t k/; it was measured as the temporal span between the first peak of release burst and the onset of the first complete vibration of the vocal folds (see Section 3.6.5.). Pre-voicing was operationalised as voicing lead values in word-initial /b d g/; it was identified as the time interval represented by the voice bar (see Section 3.6.5.). Vowel duration as a cue for consonant voicing was operationalised as the difference in duration between vowels followed by word-final /d/ and the same vowels followed by word-final /t/; vowel length was measured as the temporal span between the onset of periodicity showing clear formant structure and the abrupt diminishment of formant structure preceding a following stop (see Section 3.6.5.).

Table 7 shows the mean VOT values produced by the model talkers in /p t k/-initial tokens. The non-native speaker (NNS) realised the stops as voiceless unaspirated, which is consistent with the realisation of /p t k/ in Polish-accented English. The values are also extremely small, lower even than the mean values observed for Polish by Kopczyński (1977) and Keating et al. (1981). However, similar VOT values for /p t k/ were produced in the baseline task by a few different participants in the current study, e.g. +11 ms in *pun*, +15 ms in *pet*, +14 ms and +19 ms in *touch*, +28 ms in *cot* and +30 ms in *cut*. This observation suggests that even though they are very low, the voicing lag values produced by the Polish model talker in /p t k/ lie within the range of values exhibited by Polish learners of English. The /p t k/ tokens provided by the

native model talker (NS) can be described as voiceless aspirated. Also, the VOT values exhibited by the native speaker are considerably higher than the values reported for English by Kopczyński (1977) and Lisker and Abramson (1964) and may be a result of hyperarticulation. The explanation seems plausible given the fact that the native model talker was aware that aspiration was one of the phonetic variables under investigation. Also, the tokens containing the word-initial stops were presented in the form of voiced-voiceless minimal pairs in the word list that the native model talker was instructed to read. Thus, the native speaker might have "overaspirated" the stops to highlight the contrast between /p t k/ and /b d g/.

	/p/	/t/	/k/	overall
NNS	+13 ms	+14 ms	+27 ms	+18 ms
NS	+112 ms	+158 ms	+133 ms	+134 ms

Table 7. Mean voicing lag values in /p t k/ for the two model talkers; NNS – Polish model talker/interlocutor, NS – English model talker/interlocutor.

Table 8 shows the mean voicing lead values produced by the model talkers in /b d g/-initial tokens. The Polish model talker (NNS) realised word-initial /b d g/ with a considerable amount of prevoicing. The strategy adopted by the non-native speaker appears consistent with heavily accented pronunciation and conforms with the voicing patterns of Polish /b d g/. Nonetheless, the voicing lead values are substantially higher than those reported for Polish by Keating et al. (1981) and Kopczyński (1977). The native model talker (N), on the other hand, devoiced all instances of word-initial /b d g/, as illustrated by the complete lack of prevoicing in his realisations. The values exhibited by the English speaker are in line with the results obtained by Lisker and Abramson (1964) who found that the native English speakers in their study mostly produced word-initial /b d g/ without voicing lead.

	/b/	/d/	/g/	overall
NNS	-161 ms	-169 ms	-149 ms	-160 ms
NS	0 ms	0 ms	0 ms	0 ms

Table 8. Mean voicing lead values in /b d g/ for the two model talkers; NNS – Polish model talker/interlocutor, NS – English model talker/interlocutor.

Another measurement that could be used to characterise voicing patterns in the production of word-initial /b d g/ is voicing lag. However, the variable was not taken into consideration

in the study since preliminary examination revealed that voicing lag values in /b d g/ were similar for the model talkers and the participants. Hence, it was assumed that the subjects would have little room for accommodation and the variable was deemed irrelevant to the study.

Table 9 provides mean vowel length differences produced by the two model talkers. The values were calculated by subtracting vowel durations in the voiceless context from vowel durations in the voiced context for each of the investigated vowels. As referred to in the previous section, Jassem and Richter (1989) reported no significant length differences between vowels followed by voiced consonants and vowels followed by voiceless consonants in Polish. Hence, in order to create an impression of a heavy Polish accent in English, the non-native model talker used a voicing neutralizing pattern when producing the /t d/-final tokens. The PSOLA technique (the time-domain pitch synchronous overlap and add) was used to average any measured length differences to guarantee equal normalized durations of vowels before /t/ and /d/. Consequently, the vowel duration values in the stimuli provided by the Polish model talker are similar before voiceless and voiced stops. Conversely, the English model talker produced consistently longer vowel durations before voiced than voiceless stops, which is analogous to the pattern observed for English by Chen (1970) and Peterson and Lehiste (1960). Also, it can be seen that the native speaker produced the greatest vowel length difference in the minimal pairs containing the inherently long FLEECE vowel, whereas the smallest length difference was produced for the inherently short KIT vowel.

	TRAP	DRESS	KIT	FLEECE	overall
NNS	-3 ms	2 ms	9 ms	9 ms	4 ms
NS	105 ms	90 ms	32 ms	164 ms	98 ms

Table 9. Mean difference in duration between a vowel followed by /d/ and the same vowel followed by /t/ for the two model talkers, NNS – Polish model talker/interlocutor, NS – English model talker/interlocutor.

Although vowel duration and VOT values provided by the Polish model talker could be considered exaggerated and thus induce the participants to diverge, the Polish model talker was intentionally producing realisations that were slightly “over the top” and “Polonised” so that they would not overlap with the participants’ productions. Since the subjects’ English

pronunciation could be generally viewed as mildly accented[1], a similarly slight degree of foreign accent in the Polish model talker's realisations could render it difficult to determine whether the participants were adjusting their pronunciation as a result of exposure to the non-native talker's speech or simply maintaining their default realisation of the investigated phonetic features. For the same reason, the four front vowels in the /t d/-final tokens were also produced in a Polish-like manner: TRAP and DRESS were both realised as Polish /e/, while KIT and FLEECE were both replaced with Polish /i/. As referred to above, assimilating the TRAP vowel with the DRESS vowel and the KIT vowel with the FLEECE can be considered typical features of Polish-accented English.

3.6.3. Participants

Although a total of 44 Polish learners of English were recorded for the purposes of the study, several recordings had to be discarded due to equipment malfunctions. Moreover, two informants failed to follow the experimental procedure (i.e. omitted to repeat the target words in the imitation tasks, see Section 3.6.4.). Ultimately, the group whose pronunciation was analysed in the study consisted of 38 participants, 29 of whom were female and 9 male. The age of the subjects ranged from 20 to 23 (M=20.7). None of them reported any speech or hearing disorders. The participants were all second-year students of English Studies, recruited from the University of Lodz. Their level of English proficiency ranged from upper-intermediate to advanced[2] and they all had long experience with learning English (M=14 years, SD=2.2, Min.=9 years, Max.=15 years). The age at which learning commenced ranged from 4 to 14 (M=7 years old, SD=1.9)[3].

At the time of the experiment, the subjects had completed three semesters of an English phonetics and phonology course taught by the author of the study. Since aspiration and vowel length contrasts as a cue for consonant voicing were discussed at length during the pronunciation course, it is likely that the participants possessed a metaphonological awareness of the two features of English pronunciation. The devoicing of word-initial /b d g/ in English

[1] The judgement is based on the author's experience as an English pronunciation instructor and the subjects' phonetic performance during English phonetics and phonology classes taught by the author.
[2] The judgement was not empirically verified, it is based on the subjects' general language performance during phonetics and phonology classes taught by the author of the study. Also, the fact that the subjects possessed at least an upper intermediate level of English proficiency had been confirmed by the annual practical examinations that the participants took at the end of their first year of study.
[3] Information concerning the subjects' experience in learning English was collected via e-mail, after the experiment had taken place.

was mentioned in the second semester of the course. However, since it was not covered as extensively as the remaining two parameters, it is assumed that the subjects were less familiar with this pronunciation feature (at least on the level of metalinguistic awareness). Questionnaire responses seem to provide some support for this assumption: although several subjects stated that they focused on vowel duration and/or aspiration when producing or recognising the analysed word tokens (see Section 3.8.1.), none of them mentioned voicing in word-initial /b d g/[4].

3.6.4. Experimental procedure

Similarly as in the pilot studies, the participants listened to pre-recorded tokens provided by the two model talkers and produced the tokens under different experimental conditions. However, the procedure was modified so that the model talkers would also act as interlocutors. The modifications involved providing the subjects with false information about the nature and purpose of the experiment and are described in detail below. It is important to note that the experimental conditions which involve a non-interactional design are referred to as imitation tasks, whereas the conditions which involve an interactional design are referred to as accommodation tasks. The model talkers are referred to as such when mentioned in relation with the imitation tasks; they are referred to as interlocutors when mentioned in relation with the accommodation tasks.

Prior to the experiment, the participants were told that the aim of the study was to determine whether it is easier to understand the speech of native English or native Polish users of English. The subjects were informed that two groups of university students had already been recorded for the purposes of the study: a group of Polish learners of English recruited from the University of Silesia and a group of native speakers of English recruited from the University of Reading. The subjects were asked to draw the names of two speakers (one from each group) and told they would be required to listen to the two speakers' pronunciation in English and then provide speech samples for the two speakers to listen to at a different time. Regardless of the names they drew, all the subjects listened to the native Polish and native English model talkers only. The existence of the two groups of students was made up in order to render the pretend purpose of the experiment more credible. Also, it was assumed that

[4] Nonetheless, it is possible that the responses were primed by the fact that aspiration and vowel duration were mentioned in Q8 and Q17 of the questionnaire, whereas devoicing of word-initial stops was not.

presenting the model talkers as the subjects' peers might cause the informants to identify with the two speakers and thus provide a stronger basis for phonetic convergence.

The experiment consisted of six phases: the baseline task, two imitation tasks, two accommodation tasks and questionnaire completion. The whole experimental procedure took approximately 30 minutes for each participant to complete. In the first five phases of the experiment, the 48 tokens selected for analysis were presented on the computer screen in the form of self-running PowerPoint presentations. Four different presentation files were used in the experiment: one file for the baseline task, one file for the two accommodation tasks and two separate files for the two imitation tasks (one contained audio samples provided by the native Polish model talker, the other contained audio samples provided by the native English model talker). The slides changed automatically every 3 seconds in the baseline task, every 6 seconds in the two imitation tasks and every 3 seconds in the two accommodation task. Self-running presentations were used to control for speech tempo; the time intervals were calculated to allow the participants sufficient time to produce the target words with a natural speech tempo. The 48 tokens were presented in the same order in all five phases.

The subjects' productions were recorded using a standalone microphone; the stimuli were presented via headphones. Since the experimental procedure took a considerable amount of time to complete, it was necessary to conduct it in several sessions and in different rooms (all of which were located in the building of the Institute of English Studies, University of Lodz). Special care was taken to select rooms that were quiet and provided similar acoustic environment. At the beginning of each phase of the experiment, the author of the study would explain the task, turn on the microphone and then leave the participant alone in the room. The subjects were left alone so that they would not attempt to converge towards the author of the study and strive for a more native-like pronunciation to create a favourable impression. All of the instructions were provided in Polish in order to reduce the artificiality of the experimental procedure (the author of the study and the participants are all native speakers of Polish).

The purpose of the first phase of the experiment was to elicit subjects' baseline realisations of the 48 tokens. Each slide of the PowerPoint presentation included two English words (minimal pairs) and a picture. An example is shown in Figure 1, all 48 slides are provided in Appendix A. The participants' task was to decide which of the two words is presented in a given picture by reading it out loud. This elicitation method was selected in order to draw the subjects' attention to the semantic content of the analysed words. Presenting the tokens in a meaningful context was considered important since the findings of one of the pilot studies suggested that decontextualising the target words may increase the likelihood of

mispronunciations. More specifically, Zając and Rojczyk (2013) observed that several common and seemingly easy-to-pronounce words were frequently mispronounced in the baseline condition of the study (e.g. *seat* and *set* were often realised with the DRESS and FLEECE vowels respectively) and it was the authors' impression that many of the mispronunciations resulted from presenting the target words with no reference to their semantic value. The second reason for using a forced-choice procedure in the baseline task was to render the experiment more coherent (a similar elicitation method was also employed in the two imitation tasks).

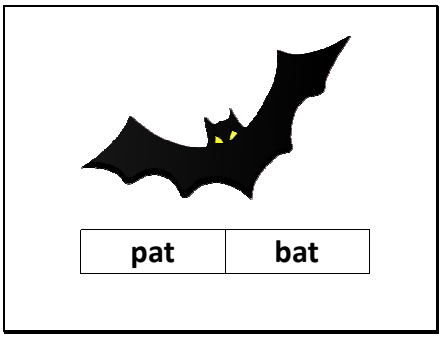


Figure 1. Sample PowerPoint slide from the baseline task.

It is also important to mention that each of the participants performed the baseline task twice: first accompanied by the author of the study (with the microphone off) and then alone (with the microphone on). This was done to ascertain that the participants are familiar with the pronunciation of the target words and are able to recognise which word is presented in a given picture. If an informant mispronounced a target word (e.g. produced *bead* with the DRESS vowel), the author of the study would write the correct pronunciation on a piece of paper using phonetic transcription. If a participant misidentified a word (e.g. produced *bun* instead of *pun*), the author of the study would correct them by saying *it's the first/second one* (the author refrained from using the target words so as not to affect the subjects' baseline productions by providing additional phonetic input).

The second and fourth phases of the experiment were the imitation tasks, referred to as such since they contained no elements of social interaction. In the second phase (1st imitation task), the participants were given an exercise sheet with 48 minimal pairs that contained the analysed tokens (a fragment of the exercise sheet is provided in Figure 2, the whole answer sheet can be found in Appendix B). The subjects were told they were going to listen to the native Polish speaker whose name they drew at the very beginning of the experiment. On the

computer screen, they could see several PowerPoint presentation files with different first names in the title. The presentations were in fact all the same and differed only with respect to the filename; they were created to confirm the assertion that two groups of students had already been recorded for the purposes of the study. The subjects' task was to run the presentation with their assigned speaker's name in the title and listen to his productions. The informants were instructed to identify the words produced by the model talker by reading them out loud and marking them on the provided exercise sheet. The fourth phase of the experiment (2nd imitation task) followed the same procedure as the second phase (1st imitation task), the difference being that the subjects listened to the native model talker's productions.

In the two imitation tasks, the subjects were required to identify the words they heard rather than simply asked to repeat the stimuli so that the instructions would remain consistent with the pretend purpose of the experiment, i.e. determining whether it is easier to understand the speech of native English or native Polish users of English. Also, it was assumed that this type of elicitation procedure might prevent the participants' from concentrating on their pronunciation and thus result in more natural productions. The exercise sheets were included to verify that the informants identified the target words correctly. The forced-choice procedure was used to facilitate the identification of the stimuli. As referred to in the previous section, the Polish model talker provided slightly exaggerated temporal values and used the same vowels in the TRAP-DRESS and FLEECE-KIT minimal pairs - recognising which tokens he produced solely on the basis of the audio stimuli would not be possible.

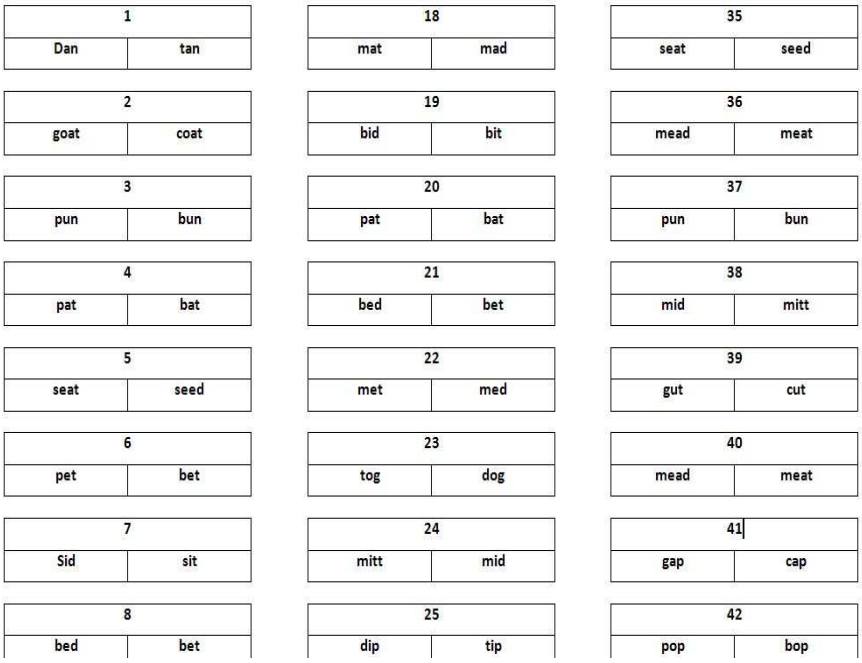


Figure 2. Fragment of the answer sheet used in the second and fourth phases of the experiment (1st and 2nd imitation task).

The third and fifth phases of the experiment were the accommodation tasks, referred to as such since they included a feature of social interaction: the model talkers acting as interlocutors. The same PowerPoint file was used in both phases and contained a list of the 48 target words (each word presented on a separate slide, see Appendix C). In both phases, the participants read the words from the computer screen. Prior to the third phase (1st accommodation task), the subjects were informed that the native Polish model talker they were assigned to would listen to their productions from this task at a later time. The participants were also told that the Polish model talker would be required to identify the stimuli they produced and rate whether their speech was easy to understand. Correspondingly, prior to the fifth phase (2nd accommodation task), the participants were told that the native English model talker they were assigned to would listen to their productions from the task at a later time, then identify the stimuli they produced and assess whether their speech was easy to understand. The subjects were told the two interlocutors would evaluate the intelligibility of their productions to provide incentive for the participants to converge. In other words, it was expected that the subjects may attempt to converge their pronunciation towards that of the model talkers to make it easier for them to understand their productions. For the sake of consistency, the subjects were also instructed to rate how easy or difficult it was to understand the model talkers' speech after completing each imitation task (the question was included in the answer sheet).

Table 10 provides an outline of the first five phases of the experimental procedure. It is important to note that although the experiment is based on a repeated measures design, the measures were not counterbalanced, i.e. all of the participants listened to the two model talkers in the same order; first to the native Polish and then to the native English speaker. Given the institutional setting of the experiment, it was expected that the subjects may view the native model talker as superior in terms of social status and language proficiency. Thus, so as not increase potential bias against the non-native speaker, none of the subjects listened to the native speaker first.

stage	task	design	task instructions
1st phase: baseline task	spoken identification of the target words	non-interactional	*Decide which word is shown in the picture by reading it out loud.*
2nd phase: 1st imitation task	spoken and written identification of the stimuli provided by the NNS model talker	non-interactional	*Listen to the Polish speaker you were assigned to, identify the words he used by saying them out loud and marking them on the exercise sheet, assess whether it was easy or difficult to understand what he was saying.*
3rd phase: 1st accommodation task	reading the target words for the NNS interlocutor	interactional: NNS model talker acting as interlocutor	*Read the words for the Polish speaker to listen to at a later time. The speaker will be asked to identify the words you read and will evaluate their intelligibility.*
4th phase: 2nd imitation task	spoken and written identification of the stimuli provided by the NS model talker	non-interactional	*Listen to the English speaker you were assigned to, identify the words he used by saying them out loud and marking them on the exercise sheet, assess whether it was easy or difficult to understand what he was saying.*
5th phase: 2nd accommodation task	reading the target words for the NS interlocutor	interactional: NS model talker acting as interlocutor	*Read the words for the English speaker to listen to at a later time. The speaker will be asked to identify the words you read and will evaluate their intelligibility*

Table 10. Outline of the first five phases of the experimental procedure.

Following the fifth phase of the experiment (2nd accommodation task), the participants were asked to complete a questionnaire (Appendix D). The questionnaire was written in Polish so as to remain consistent with the previous stages of the experiment (where all the instructions were provided in the subjects' mother tongue). The structure of the questionnaire is presented in Table 11. The purpose of the questionnaire was twofold. Firstly, its aim was to

verify the assumption that the participants favour native pronunciation over Polish-accented English by gauging their attitude towards the pronunciation of the two model talkers' and their attitude towards non-native English pronunciation in general (evaluation component and attitudes component). The second aim of the questionnaire was to provide a fuller account of the informants' convergence strategies by asking them to comment on their speech behaviour during the experimental tasks (self-report component). A few additional questions were also included with a view to facilitating the interpretation of the results (phonetic variables component and pronunciation model component).

component	question no.	question type
evaluation component	1-21	Likert-type scale
attitudes component	22-35	Likert-type scale
self-report component: A. baseline and imitation tasks B. accommodation tasks	19-21 (A) 9, 18 (B)	multiple choice (A) open-ended (B)
phonetic variables component	8, 17	multiple choice
pronunciation model component	22	multiple choice

Table 11. Structure of the questionnaire.

In the evaluation component, the participants were required to state whether they considered each model talker's pronunciation to be correct and pleasant to listen to. They were also asked to assess whether each model talker sounded intelligent, professional, educated, friendly and attractive. The maximum score that the model talkers/interlocutors could receive was 35. Self-report component A was concerned with the subjects' phonetic performance in the baseline and imitation tasks. The participants were required to choose between three options: a) *I concentrated on my pronunciation and tried to sound native-like*, b) *I paid no attention to my pronunciation* or c) *other* (in which case the subjects were asked to specify what they did). In self-report component B, the participants were asked whether they adjusted their pronunciation when reading for the two model talkers (i.e. in the 1st and 2nd accommodation tasks). The phonetic variables component was concerned with whether the subjects noticed and paid attention to given pronunciation feature in the productions of the model talkers/interlocutors (it was assumed that the information may prove useful when interpreting the results of the study). In the pronunciation model component, the subjects stated what they would like to sound like when speaking English; they were required to

choose between the following options: a) *like an American* b) *like a British person* c) *like a Polish person* d) *I don't mind how I sound as long as I'm able to communicate* e) *other*. The attitudes component examined subjects' attitudes towards Polish-accented English, their opinion on the importance of using native-like pronunciation when speaking English and the importance of pronunciation as compared with other language skills. The maximum score in this part of the questionnaire was 65; achieving a score close to this number was expected to signify a strong bias against Polish-accented speech.

3.6.5. Measurements

The phonetic variables examined in the study were aspiration (operationalised as voicing lag values in initial /p t k/), pre-voicing in word-initial stops (operationalised as voicing lead values in initial /b d g/) and vowel duration as a cue for consonant voicing (operationalised as the difference in duration between vowels followed word-final /d/ and the same vowels followed by word-final /t/). All three parameters were measured using Praat speech-analysis software package (Boersma, 2001) by means of waveform and spectrographic display. Voicing lag in /p t k/ tokens was measured as the temporal span between the first peak of release burst and the onset of the first complete vibration of the vocal folds (e.g. Lisker and Abramson, 1964; Cole et al., 2007; Rojczyk, 2010); an example is provided in Figure 3. Voicing lead in /b d g/ tokens was identified as the time interval represented by the voice bar (e.g. Lisker and Abramson, 1964); an example is provided in Figure 4. Vowel duration was measured from the onset of periodicity showing clear formant structure to the abrupt diminishment of formant structure preceding a following stop (e.g. Slowiaczek and Dinnsen, 1985; Rojczyk, 2010); an example is provided in Figure 5. The length difference between the voiced and voiceless context was calculated by subtracting the duration value before word-final /t/ from the duration value before word-final /d/ for each of the investigated vowels.

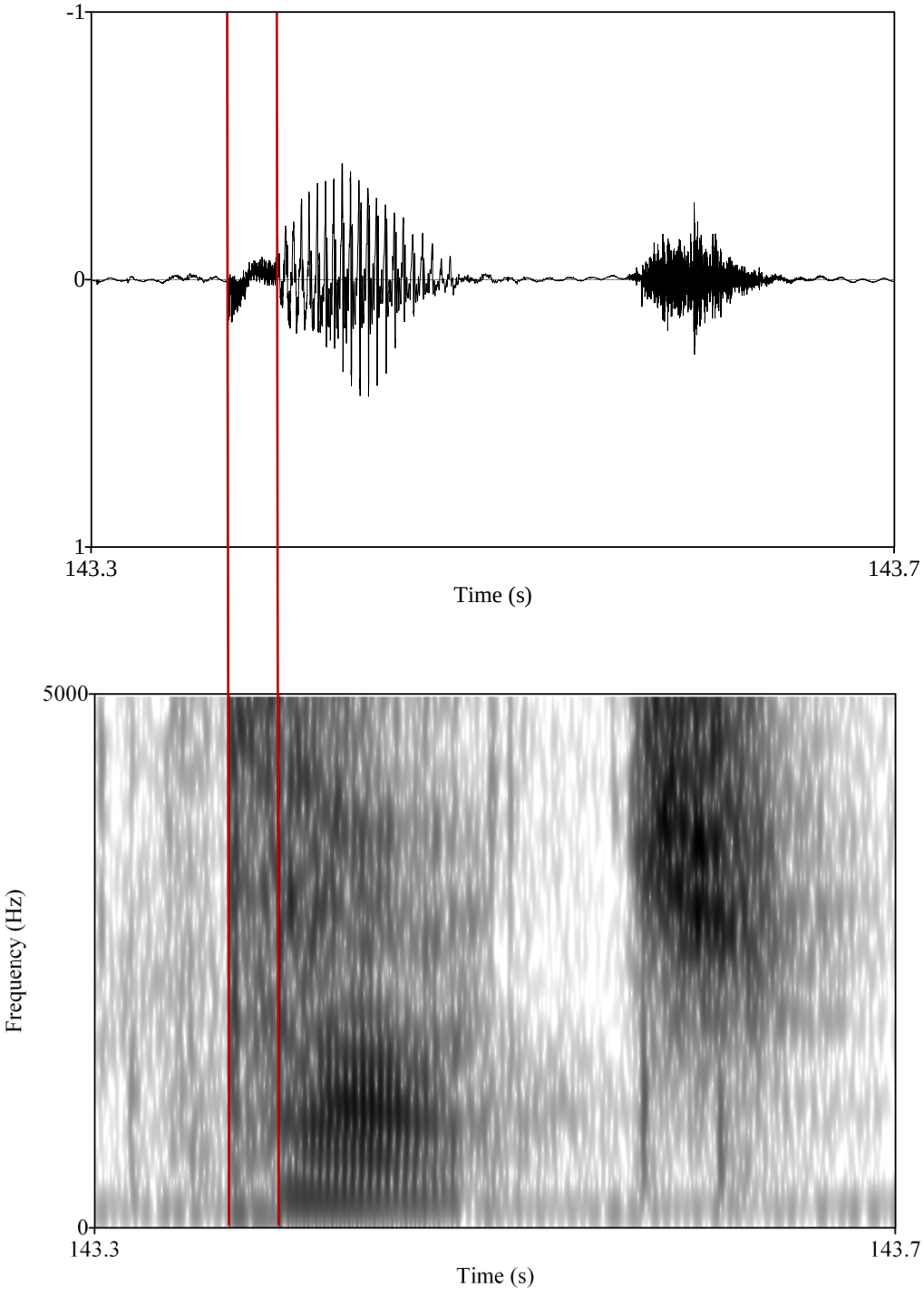


Figure 3. Waveform and spectrogram of *touch* produced by one of the participants; the temporal span that represents voicing lag is marked with red bars.

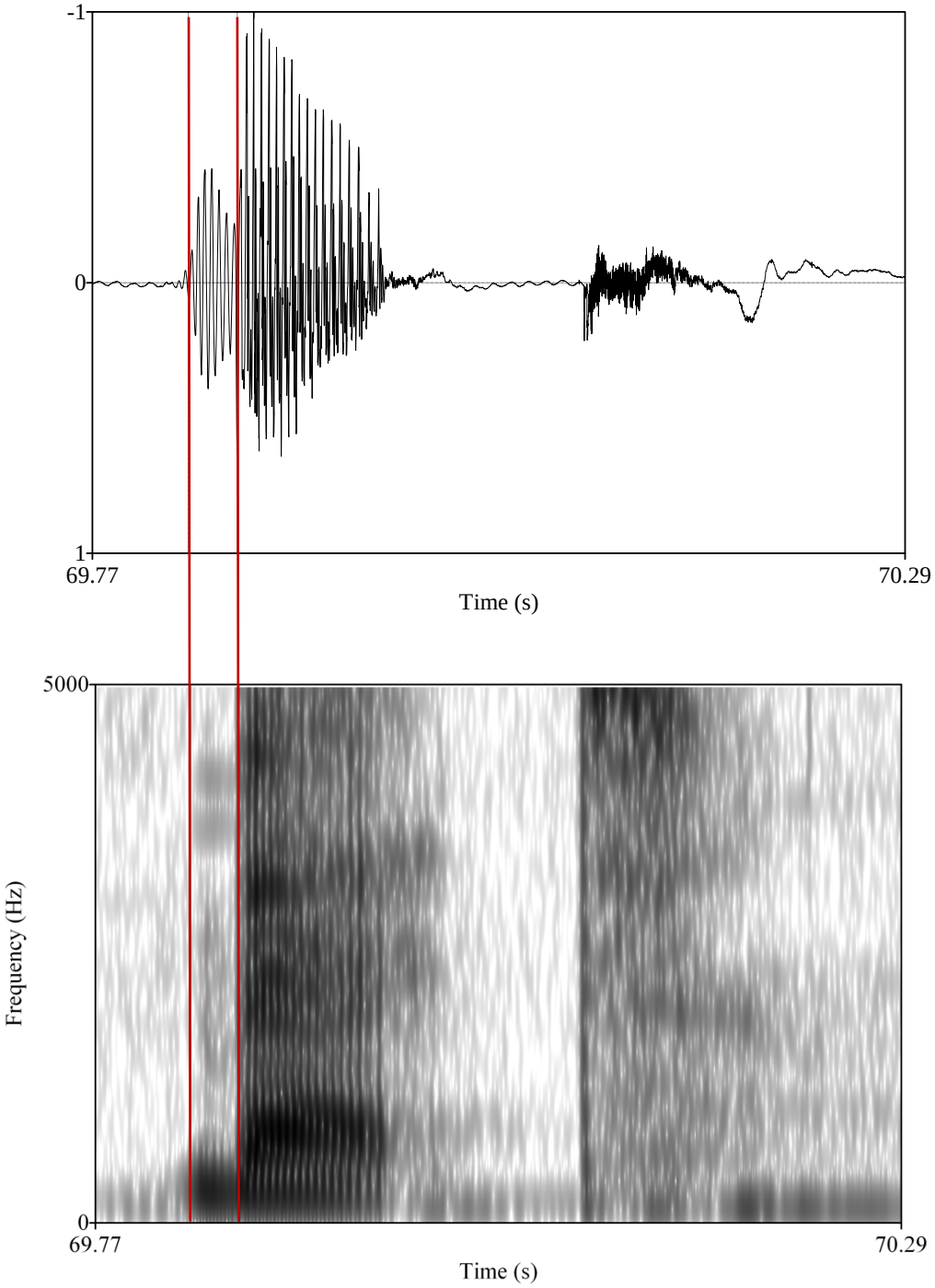


Figure 4. Waveform and spectrogram of *bet* produced by one of the participants; the temporal span that represents voicing lead is marked with red bars.

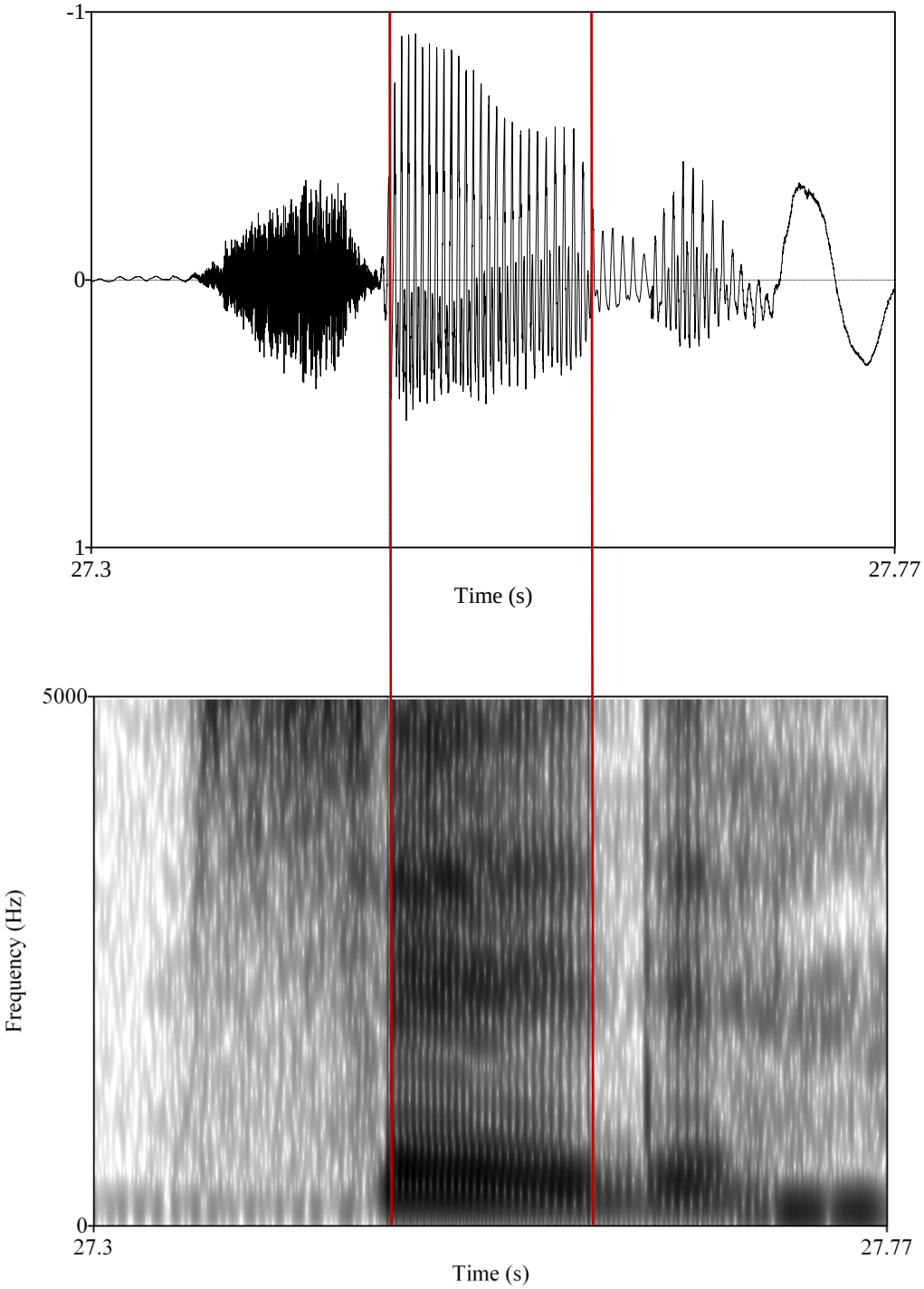


Figure 5. Waveform and spectrogram of *sit* produced by one of the participants; the temporal span that represents vowel duration is marked with red bars.

3.6.6. Statistical analysis

Although a total number of 48 target words served as stimuli in the study, a considerable number had to be excluded from statistical analysis due to frequent misidentifications in the imitation tasks. Examination of the answer sheets used in the two tasks revealed that the target words produced by the model talkers were often confused with their voiced or voiceless counterparts. The /p t k/-initial tokens were frequently mistaken for /b d g/-initial tokens in the 1st imitation task (in which the participants were required to listen to the productions provided by the Polish model talker). For example, the majority of the participants identified *tan* as *Dan* and *coat* as *goat*. The /b d g/-initial tokens, on the other hand, were often confused with their /p t k/-initial counterparts in the 2nd imitation task (in which the subjects were required to listen to the productions provided by the English model talker). For example, several subjects mistook *goat* for *coat* and *bet* for *pet.* As far as vowel duration is concerned, the /t/-final tokens were often confused with /d/-final tokens in both imitation tasks. Ultimately, the minimal pairs that were least frequently misidentified were selected for statistical analysis and include the following word sets: *mat-mad, set-said, mitt-mid, seat-seed, pat-bat, pop-bop, cap-gap, cut-gut*. Tokens containing word-initial /t d/ were misidentified so frequently that they had to be altogether excluded from statistical analysis.

The primary aim of the study was to investigate whether convergence strategies in the speech of Polish learners of English will vary as a function of model talker/interlocutor. With this objective in mind, the relationship between the model talker (native vs. non-native) and the three phonetic parameters (aspiration, pre-voicing and vowel duration as a cue for consonant voicing) was tested by conducting three two-way repeated measures ANOVAs (one for each phonetic parameter). Each ANOVA included two independent variables. The first independent variable had five levels corresponding to the different experimental phases (baseline, 1st imitation, 1st accommodation, 2nd imitation, 2nd accommodation). The second independent variable differed according to the phonetic parameter under investigation. In the case of aspiration and stop voicing, the second independent variable had two levels that corresponded to different places of articulation (bilabial, velar). In the case of vowel duration, the second independent variable had four levels that corresponded to different vowel qualities (TRAP, DRESS, KIT, FLEECE). The dependent variable was a repeated measurement of a particular phonetic parameter. The repeated measurements of aspiration were entered into the statistical model as the mean voicing lag value of /p/ in *pat* and *pop* and the mean voicing lag

value of /k/ in *cap* and *cut*. The repeated measurements of stop voicing were entered into the statistical model as the mean voicing lead value of /b/ in *bat* and *bop* and the mean voicing lead value of /g/ in *gap* and *gut*.

It is also important to mention that the repeated measurements of the three phonetic parameters were not taken from the exact same set of participants. More specifically, the voicing lag and voicing lead measurements were taken from the same set of 25 participants, while the vowel duration measurements were taken from a different set of 33 participants. Both sets were selected out of the group of 38 participants whose productions were analysed in the study; there was a substantial overlap between the two sets (similarity was measured using Jaccard coefficient, Sj=0.66). Using the same set of participants in each case was not possible due to the frequent misidentifications of target words in the imitation tasks. In order to make statistical analysis of the data possible, it was necessary to exclude the productions of those participants who committed the greatest amount of identification errors. Thus, it was not only certain minimal pairs but also the productions of certain participants that had to be excluded from statistical analysis as a result of frequent target word misidentifications in the imitation phases of the experiment.

3.7. Research questions

The research questions are based on the assumptions specified in Section 3.5. and reflect the experimental design of the study. Given the relative complexity of the experimental design, the purpose of formulating research questions was to help structure the analysis and discussion of the results. Thus, the analysis and discussion of the data in Chapter Four will be arranged around the twelve research questions that are presented in this section. The answers to the research questions will be summarised and used to test the hypotheses (see Section 3.5.) in one of the final sections of Chapter Four (Section 4.6.). The research questions are presented below (each is discussed in relation to the hypothesis/-es it will be used to test):

RQ1: How were the three phonetic variables realised in the baseline condition?

RQ1 refers to participants' baseline realisations, which will be used as a point of reference in the analysis of convergence strategies (i.e. convergence, divergence, maintenance; see Section 3.5.). The answer to RQ1 will be essential in testing all three hypotheses formulated for the purposes of the study.

RQ2: What are the attitudes of the participants towards L2 pronunciation in English?

RQ2 pertains to attitudinal factors that could potentially affect participants' convergence strategies (i.e. convergence, divergence, maintenance; see Section 3.5.).; it will be used to test Hypothesis 2.

RQ3: According to the participants, what convergence strategies did they use in the imitation and accommodation tasks?

RQ3 is concerned with self-reported convergence strategies, the knowledge of which is expected to facilitate the interpretation of the results with respect to the effect of attitudinal factors; the answer to RQ3 will be used to test Hypothesis 2.

RQ4: In the case of aspiration, what imitation strategies did the participants use following exposure to native and non-native speech?

RQ5: In the case of pre-voicing, what imitation strategies did the participants use following exposure to native and non-native speech?

RQ6: In the case of vowel duration, what imitation strategies did the participants use following exposure to native and non-native speech?

The term imitation strategies that is used in RQs4-6 denotes convergence strategies (i.e. convergence, divergence, maintenance; see Section 3.5.) that are analysed in a non-interactional context. RQs4-6 apply to speech behaviour in the imitation conditions as compared with the baseline condition; each pertains to a different phonetic variable. The answers to RQs4-6 will be used to test Hypothesis 1 and Hypothesis 2.

RQ7: : In the case of aspiration, what accommodation strategies did the participants use following exposure to native and non-native speech?

RQ8: In the case of pre-voicing, what accommodation strategies did the participants use following exposure to native and non-native speech?

RQ9: In the case of vowel duration, what accommodation strategies did the participants use following exposure to native and non-native speech?

The term accommodation strategies that is used in RQs7-9 denotes convergence strategies (i.e. convergence, divergence, maintenance; see Section 3.5.) that are analysed in an interactional context. RQs7-9 are concerned with speech behaviour in the accommodation conditions as compared with the baseline conditions; each refers to a different phonetic variable. The answers to RQs7-9 will be used to test Hypothesis 1 and Hypothesis 2.

RQ10: In the case of aspiration, what convergence strategies did the participants use with respect to different places of articulation?

RQ11: In the case of pre-voicing, what convergence strategies did the participants use with respect to different places of articulation?

RQ12: In the case of vowel duration, what convergence strategies did the participants use with respect to different vowel categories?

RQs10-12 are concerned with the effect of phonetic context on convergence strategies (i.e. convergence, divergence, maintenance; see Section 3.5.), each pertains to a different pronunciation feature. The answers to RQs10-12 will be used to test Hypothesis 3.

3.8. Results

Questionnaire results are shown in Section 3.8.1. All of the questions and responses mentioned in the text had been translated from Polish into English by the author of the study (the questionnaire was written in Polish, see Section 3.6.4.). The results yielded by the statistical analysis of the data are presented in Section 3.8.2., which is further subdivided according to phonetic variable under investigation. It should be noted that this section of the dissertation concentrates solely on the presentation of the results of the study; they are analysed and discussed in Chapter Four.

3.8.1. Questionnaire results

Table 12 shows the mean scores received by the model talkers/interlocutors in the evaluation component of the questionnaire. The results are presented separately for the two subsets of the subject group (see Section 3.6.6.). The maximum score that the model talkers/interlocutors could receive was 35.

	VOT group (N=25)		V_DUR group (N=33)	
	mean	SD	mean	SD
NNS	20	4.4	21	4.6
NS	30	4.1	30	3.4

Table 12. Mean scores received by the model talkers/interlocutors in the evaluation component; NNS – Polish model talker/interlocutor, NS – English model talker/interlocutor.

Table 13 presents the mean scores of the participants in the attitudes component. The findings are shown separately for the two subsets of the subject group. The maximum score in this part of the questionnaire was 65; achieving a score close to this number was expected to signify a strong bias against Polish-accented speech.

VOT group (N=25)		V_DUR group (N=33)	
mean	SD	mean	SD
50	6.8	49	7.2

Table 13. Mean scores in the attitudes component.

Tables 14, 15 and 16 show the proportion of participants who had selected a given option in the self-report component A of the questionnaire. The number of participants who had selected a given option is given in brackets. The results are presented separately for the two subsets of the subject group (see Section 3.6.6.). Each table represents a different experimental condition (baseline, 1st imitation, 2nd imitation).

	baseline		
	I tried to sound native-like	*I paid no attention to my pronunciation*	*other*
VOT group (N=25)	84% (21)	12% (3)	4% (1)
V_DUR group (N=33)	79% (26)	18% (6)	3% (1)

Table 14. Self-reported speech behaviour in the baseline task - the proportion of participants who selected a given option.

The participant who selected the option *other* in the baseline task stated:

- *Sometimes I thought about the pronunciation of the word after I had produced it. I produced them in a natural way.*

	1st imitation		
	I tried to sound native-like	*I paid no attention to my pronunciation*	*other*
VOT group (N=25)	68% (17)	4% (1)	28% (7)
V_DUR group (N=33)	64% (21)	9% (3)	27% (9)

Table 15. Self-reported speech behaviour in the 1st imitation task - the proportion of participants who selected a given option.

The subjects who selected the option *other* with respect to the 1st imitation task stated:

- *Sometimes I wasn't able to concentrate on my pronunciation, concentrating on what I heard instead*
- *It was sometimes difficult to recognise what he said because of the pronunciation errors. I paid attention to the Polish speaker's pronunciation and I tried to recognise what he had said first. Then I tried to pronounce the word as best as I could.*

- *I imitated some of the sounds produced by the Polish speaker even though I knew he was mispronouncing them*
- *I concentrated on my pronunciation but I pronounced the words similarly to the Polish person I had listened to*
- *I inadvertently imitated his pronunciation*
- *At first I imitated his speech, then I tried to pronounce the words my way.*
- *His pronunciation influenced my pronunciation, I committed the same mistakes he did.*
- *I definitely paid less attention to my pronunciation but I did not completely forget about it.*
- *I paid attention to the Polish speaker's pronunciation and I imitated it.*
- *I concentrated on my pronunciation but I tried to pronounce the words similarly to him.*
- *Sometimes I thought about the pronunciation of the word after I had produced it. I produced them in a natural way.*

	2nd imitation		
	I tried to sound native-like	*I paid no attention to my pronunciation*	*other*
VOT group (N=25)	92% (23)	0	8% (2)
V_DUR group (N=33)	91% (30)	3% (1)	6% (2)

Table 16. Self-reported speech behaviour in the 2nd imitation task - the proportion of participants who selected a given option.

The informants who opted for *other* with respect to their speech behaviour in the 2nd imitation task stated:

- *Sometimes I was difficult to recognise the word if the difference lied in vowel lengthening before a voiced consonant – this is more difficult for me than aspiration. I listened to the British person's speech and tried to imitate it.*
- *I imitated his pronunciation.*
- *Sometimes I thought about the pronunciation of the word after I had produced it. I produced them in a natural way.*

Table 17 shows the number of participants who stated they had converged their pronunciation towards that of the non-native interlocutor (NN) or native interlocutor (N) in the self-report component B of the questionnaire. The results are presented separately for the

two subsets of the subject group (see Section 3.6.6.). Self-reported convergence towards the Polish interlocutor was operationalised as a statement that involves the participant declaring they had imitated (or attempted to) the speech of the non-native speaker and/or had used (or attempted to) more Polish-like realisations. Self-reported convergence towards the English interlocutor was operationalised as a statement that involves the participant declaring they had imitated (or attempted to) the speech of the native speaker and/or had used (or attempted to) more native-like realisations.

	convergence towards NNS (1st accommodation task)	convergence towards NS (2nd accommodation task)
VOT group (N=25)	0	23
V_DUR group (N=33)	4	31

Table 17. Self-reported speech behaviour in the accommodation tasks – the number of participants who stated they converged towards non-native/native interlocutor; NNS – Polish interlocutor, NS – English interlocutor.

Some of the participants who declared they did not converge towards Polish-like values in the 1st accommodation task stated[5]:

- *I paid attention to pronunciation correctness.*
- *I just tried to pronounce the words well.*
- *I tried to use aspiration and to shorten vowels before voiceless consonants, because I want my pronunciation to be correct. Moreover, I want to realise vowels the way a typical native speaker would.*
- *I tried to pronounce the words the way I think they should be pronounced, because I value perfectionism.*
- *I tried to correct the mistakes that I had noticed in his [the Polish speaker's] pronunciation.*
- *No [I didn't adjust], I tried to read the way I normally would, because he wasn't a native speaker and I noticed some mistakes.*
- *I tried to change my pronunciation when I thought a given word had been mispronounced.*
- *I tried to change my pronunciation because I know some of these words were mispronounced. I see no point in repeating somebody's mistakes. Besides, we always want to sound as best as we possibly can.*

[5] The selected answers are the ones that were considered representative (i.e. give an opinion that was expressed by at least several participants) or interesting; they do not include all of the responses given by the participants. Also, the responses are not presented separately for the two subsets of the subject group (see Section 3.6.6.) as it can be seen in Table 17 that the two groups exhibited comparable convergence strategies.

- *I tried to pronounce the words clearly and carefully so that the person that's going to listen to me doesn't have any doubts as to which words I'm pronouncing.*
- *I tried to sound as intelligible as possible so that it would be easier to distinguish different sounds.*
- *I tried to emphasise voiced and voiceless consonants at the end of words so that it would be easier to understand me.*
- *I tried to use my regular pronunciation but I didn't always succeed.*
- *I didn't adjust, I just tried to sound as correct as I can. I sometimes adjust my pronunciation in situations where communication is the goal, in a conversation. Here, I didn't feel that communication was the goal, I just focused on having the best pronunciation possible, I think it should be intelligible to another student of English studies.*

The four participants who stated they converged towards the Polish interlocutor in the 1st accommodation task expressed the following opinions:

- *For some reason, after hearing the words he read, instead of pronouncing them the way I think they should be pronounced, I repeated his incorrect (in my opinion) pronunciation.*
- *I tried to read similarly to that person. You could say that I adopted the way he was speaking to some extent. I tried to sound similar to make sure that he would understand me. I repeated the words he pronounced the way I heard them, even if they were sometimes mispronounced.*
- *Yes, I tried to adjust my pronunciation towards what I had heard earlier, speak more clearly.*
- *Yes [I adjusted my pronunciation], when I was repeating after the recording I pronounced the words automatically. I didn't think about correct pronunciation.*

Some of the participants who attempted to converge towards native-like values in the 2nd accommodation task stated[6]:

- *I just tried to sound as best as I can.*

[6] Similarly as in the case of the 1st accommodation task, the selected responses are the ones that were considered representative (i.e. give an opinion that was expressed by at least several participants); they do not include all of the responses given by the participants. The answers are not presented separately for the two subsets of the subject group (see Section 3.6.6.) as it can be seen in Table 17 that the two groups exhibited comparable convergence strategies.

- *I tried to read the words as correct as I can and to some extent adjust my pronunciation towards what I had heard so that it would sound more natural for the British person.*
- *Yes, I tried to imitate him because he's a native speaker.*
- *I adjusted my pronunciation because I wanted to sound similar to him because in my opinion, British accent is the "correct" one, I also wanted to sound intelligible.*
- *I adjusted because I wanted to sound like a native-speaker, I tried to focus my attention on the sounds /æ/, /ɪ/ and aspiration.*
- *Yes [I adjusted]. I tried to use glottal stops because Dave [the native interlocutor] used them. I concentrated on the quality of /ɪ/, I wanted to sound as natural as I can because a native speaker will spot by mistakes more quickly and easily.*
- *I tried to sound as best as I could. It's more difficult for native speakers to understand accents (I suppose). Besides, I wanted to sound as smart as he did.*
- *I tried to pronounce the words in a similar way so that he wouldn't notice that I'm not a native speaker.*
- *When I was reading the words for this person, I tried to pay attention not only to the correct quality of vowels and consonants but also my accent so as not to sound too "Polish"*
- *I tried to adjust my pronunciation so that I would sound more "English". Apart from that, I tried to sound the same [as him], because I didn't want him to think I can't speak English very well.*
- *I definitely tried to copy that person's pronunciation, because I know he pronounced the words the correct way and I wanted to sound like him.*
- *I adjusted towards his pronunciation – I didn't want it to be so obvious that I'm not a native speaker.*

Three out of the four participants who declared they did not converge towards the native interlocutor in the 2nd accommodation task simply stated *No, I did not adjust*. One participant said she did not adjust because her target accent is General American.

Results obtained in the phonetic variable component of the questionnaire show that in both subsets of the subject group (see Section 3.6.6.), approximately 55% of the participants stated they had noticed aspiration (or lack of it) in the model talkers' pronunciation. Approximately 35% of the subjects declared they had noticed vowel length contrasts (or lack of them) before

voiced/voiceless consonants in the model talkers' pronunciation. Since word-initial devoicing (or pre-voicing) was omitted from the phonetic variable component of the questionnaire, it is not possible to state what proportion of the participants noticed this pronunciation feature in the model talkers' realisations.

Results from the pronunciation model component demonstrated that the vast majority of the participants expressed a preference for British English. Five subjects declared their preferred pronunciation model was American English; none of the informants stated they wished so sound like a Polish person when speaking English.

3.8.2. Results of statistical analysis

This section presents the results yielded by the statistical analysis of the data and is further subdivided into three subsections, each of which is concerned with a different phonetic variable (aspiration, pre-voicing and vowel duration as a cue for consonant voicing).

3.8.2.1. Aspiration

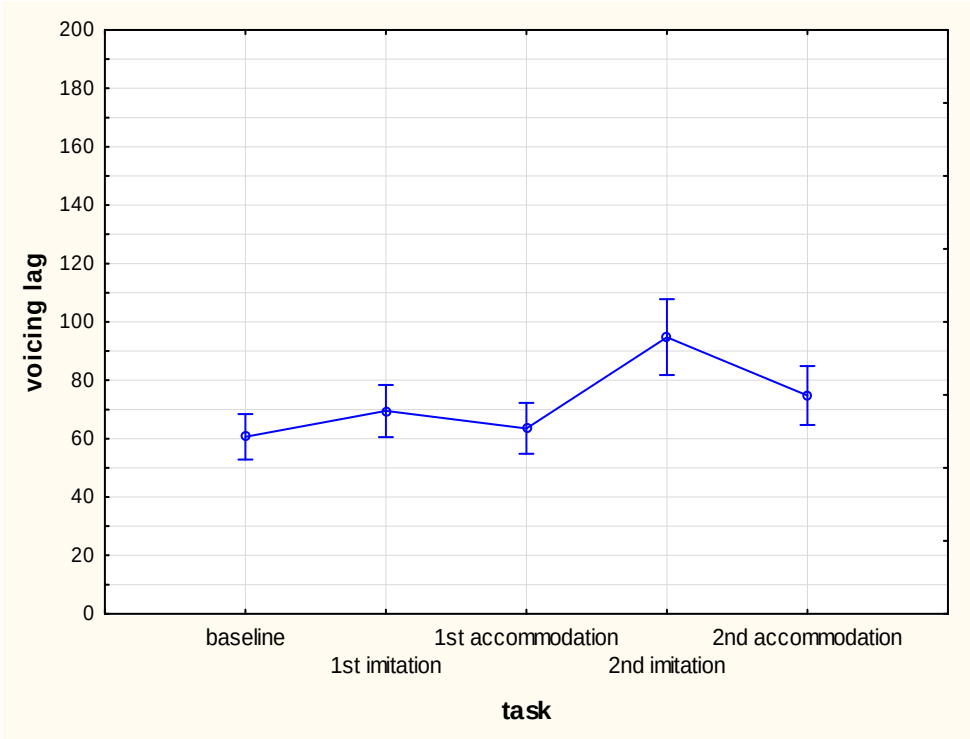


Figure 6. Mean voicing lag values (ms) across different tasks (N=25).

task	mean
baseline	61 ms (3.8)
1st imitation	69 ms (4.3)
1st accommodation	63 ms (4.2)
2nd imitation	95 ms (6.3)
2nd accommodation	75 ms (4.9)

Table 18. Mean voicing lag values across different tasks (N=25), SD given in brackets.

Figure 6 and Table 18 show mean voicing lag values in all five experimental tasks. It can be seen that the mean VOT values consistently exceeded 60 ms. As compared with the baseline, an increase in mean voicing lag values can be observed in each experimental condition. The difference is more marked following exposure to the pronunciation of the native speaker (especially in the imitation condition) than following exposure to the pronunciation of the non-native speaker (especially in the accommodation condition, where the increase is small enough to be considered inconsequential). Statistical analysis of the results revealed that the main effect of task on voicing lag values was highly significant [$F(4, 96)=32.1$, $p<.001$]. Post hoc Bonferroni tests showed that the increase in duration was significant in: the 1st imitation task as compared with the baseline [$p<.01$], the 2nd imitation task as compared with the baseline [$p<.001$], the 2nd accommodation task as compared with the baseline [$p<.001$], the 2nd imitation task as compared with the 2nd accommodation task [$p<.001$].

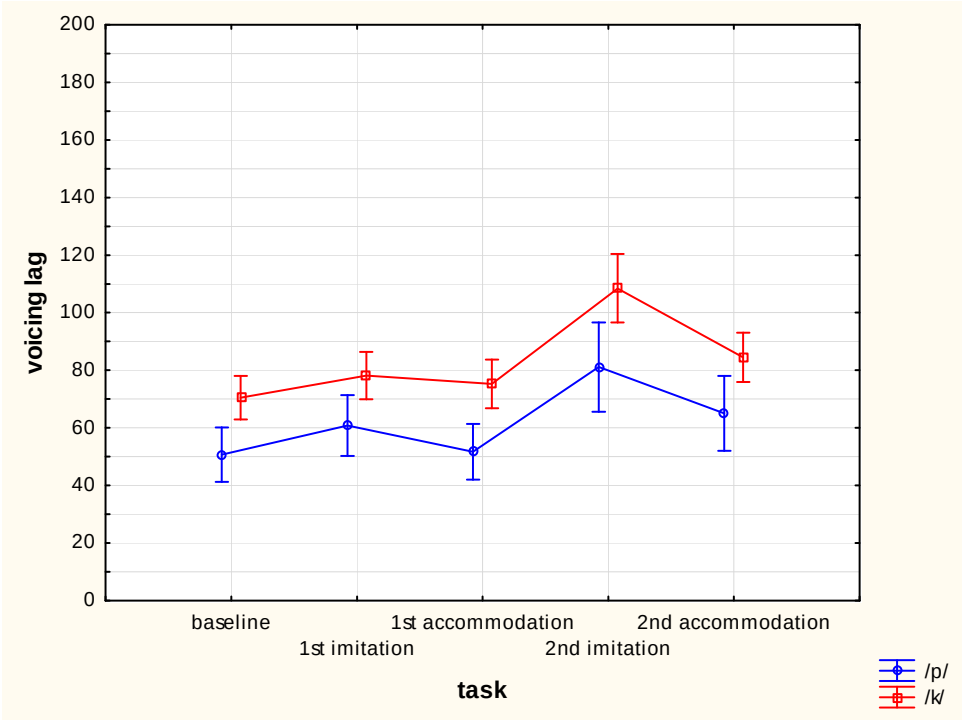


Figure 7. Mean voicing lag values (ms) for /p/ and /k/ across different tasks (N=25).

task	/p/	/k/
baseline	51 ms (4.6)	70 ms (4.6)
1st imitation	61 ms (3.7)	78 ms (4.0)
1st accommodation	52 ms (4.7)	75 ms (4.1)
2nd imitation	81 ms (7.5)	108 ms (5.8)
2nd accommodation	65 ms (6.3)	85 ms (4.2)

Table 19. Mean voicing lag values for /p/ and /k/ across different tasks (N=25), SD given in brackets.

Figure 7 and Table 19 show mean voicing lag values for /p/ and /k/ in all experimental tasks. It can be seen that voicing lag in /k/ was consistently realised as longer than in /p/. For both consonants, an increase in mean voicing lag values can be observed in each experimental condition as compared with the baseline. Also, for both /p/ and /k/, the duration difference between the baseline and the 1st accommodation is very slight, small enough that it could considered immaterial. Statistical results of the analysis revealed that the interaction between task and place of articulation was not statistically significant [F(4, 96)=1.6, p>.05].

3.8.2.2. Pre-voicing

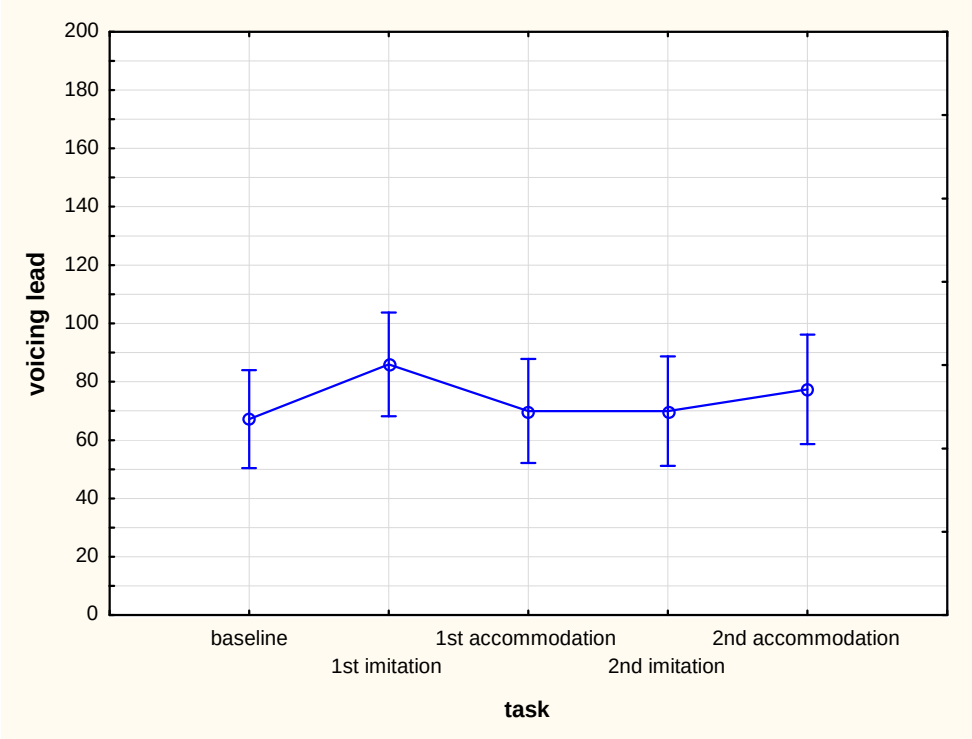


Figure 8. Mean voicing lead values (ms) across different tasks (N=25).

task	mean
baseline	67 ms (8.1)
1st imitation	86 ms (8.6)
1st accommodation	70 ms (8.6)
2nd imitation	70 ms (9.1)
2nd accommodation	77 ms (9.1)

Table 20. Mean voicing lead values across different tasks (N=25), SD given in brackets.

Figure 8 and Table 20 present mean voicing lead values in all five experimental tasks. The values are presented in the form of positive numbers to facilitate the interpretation of the results. It can be seen that the mean values consistently exceeded 60 ms. Interestingly, SD values are considerably higher than those found for aspiration (see Table 18). As compared with the baseline, an increase in mean voicing lead values can be observed in each experimental condition. Nonetheless, the difference appears to be very slight in all but one case, i.e. productions upon exposure to non-native speech in the 1st imitation condition. Statistical analysis of the results revealed that the main effect of task on voicing lead values was significant [$F(4, 96)=2.78$, $p<.05$]. Post hoc Bonferroni tests showed that the increase in duration was significant only for the 1st imitation task as compared with the baseline [$p<.05$].

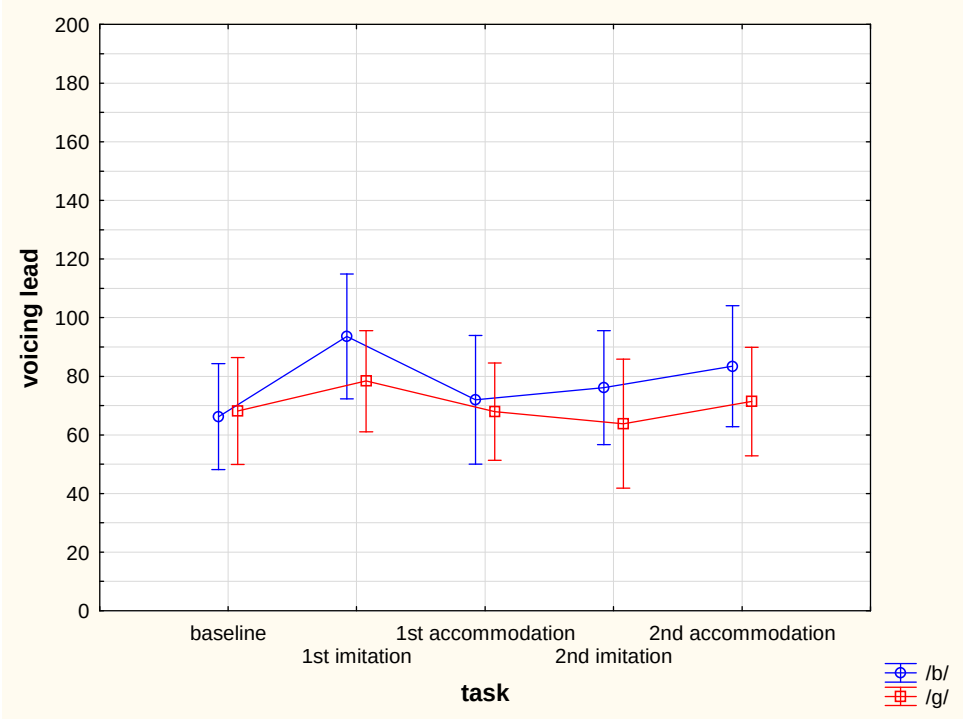


Figure 9. Mean voicing lead values (ms) for /b/ and /g/ across different tasks (N=25).

task	/b/	/g/
baseline	66 ms (8.8)	68 ms (8.9)
1st imitation	94 ms (10.3)	78 ms (8.3)
1st accommodation	72 ms (10.6)	68 ms (8.0)
2nd imitation	76 ms (9.4)	64 ms (10.7)
2nd accommodation	83 ms (10.0)	71 ms (9.0)

Table 21. Mean voicing lead values for /b/ and /g/ across different tasks (N=25), SD given in brackets.

Figure 9 and Table 21 show mean voicing lag values for /b/ and /g/ in all experimental tasks. The values are presented in the form of positive numbers to facilitate the interpretation of the results. It can be seen that that voicing lead in /b/ was slightly longer than in /g/ under all experimental conditions apart from the baseline. For both consonants, an increase in mean voicing lag values can be observed in the 1st imitation task as compared with the baseline. In the 1st accommodation task, mean voicing lead increased for /b/ and remained the same as in the baseline for /g/. In the 2nd imitation task, the mean VOT value increased for /b/ and decreased very slightly for /g/. In the 2nd accommodation task, mean voicing lead increased both for /b/ and /g/ (although the difference is very small for the latter). Statistical analysis indicated that the assumption of sphericity had been violated [$\chi^2(9)=28.7$, $p<.001$], therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon<.75$). The results revealed that there was no significant interaction between task and place of articulation [$F(2.42, 58.2)=.97$, $p>.05$].

3.8.2.3. Vowel duration as a cue for consonant voicing

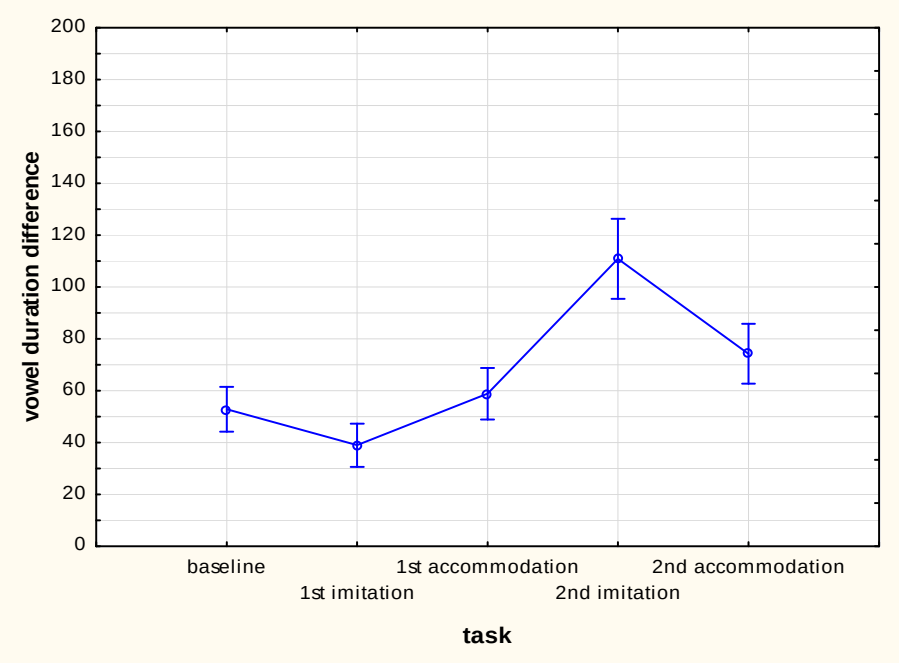


Figure 10. Mean duration differences (ms) between vowels followed by word-final /d/ and vowels followed by word-final /t/ (N=33).

task	mean
baseline	53 ms (4.2)
1st imitation	39 ms (4.1)
1st accommodation	59 ms (4.9)
2nd imitation	111 ms (7.6)
2nd accommodation	74 ms (5.7)

Table 22. Mean duration differences between vowels followed by word-final /d/ and vowels followed by word-final /t/ (N=33), SD given in brackets.

Figure 10 and Table 22 show mean vowel duration differences in all five experimental tasks. It can be seen that the mean values are positive and exceeded 30 ms. A decrease in duration can be observed in the 1st imitation task as compared with the baseline. There is an increase in duration in the 1st accommodation task as compared with the baseline, however, it seems small enough to be considered inconsequential. An increase in mean duration values can also be seen following exposure to the pronunciation of the native speaker (2nd imitation and 2nd accommodation); the difference is more marked in the non-interactional task. Mauchly's test indicated that the assumption of sphericity had been violated [$\chi^2(9)=33.15$, $p<.001$], therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon<.75$). The results showed that there was a highly significant main effect of task on mean differences in vowel duration [$F(2.74, 87.8)=52.5$, $p<.001$]. Post hoc Bonferroni tests revealed that there

was a significant difference in duration values between: the 1st imitation task and the baseline [p<0.01], the 2nd imitation task and the baseline [p<.001], the 2nd accommodation task and the baseline [p<.001], the 2nd imitation task and the 2nd accommodation task [p<.001].

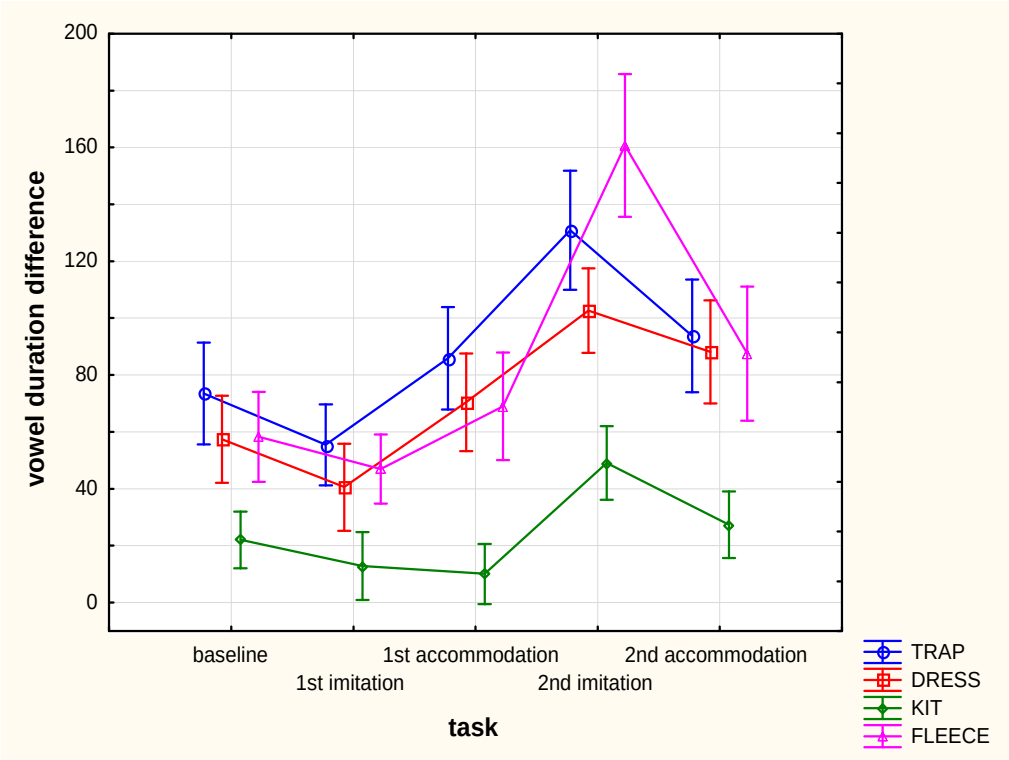


Figure 11. Mean duration difference (ms) for different vowel categories across 5 experimental tasks (N=33).

task	TRAP	DRESS	KIT	FLEECE
baseline	73 ms (8.8)	57 ms (7.5)	22 ms (4.9)	58 ms (7.8)
1st imitation	55 ms (7.0)	41 ms (7.5)	13 ms (5.8)	47 ms (5.9)
1st accommodation	86 ms (8.8)	70 ms (8.4)	11 ms (5.2)	69 ms (9.3)
2nd imitation	131 ms (10.3)	103 ms (7.3)	49 ms (6.4)	161 ms (12.3)
2nd accommodation	94 ms (9.7)	88 ms (8.9)	27 ms (5.7)	88 ms (11.6)

Table 23. Mean duration difference for different vowel categories across 5 experimental tasks (N=33), SD given in brackets.

Figure 11 and Table 23 show mean duration difference for different vowel categories across the five experimental conditions. It can be seen that the duration difference for KIT was consistently realised as the smallest among the three vowels. In the baseline task and following exposure to non-native speech, mean duration difference was greatest for the TRAP vowel; the duration difference for FLEECE and DRESS were comparable. In the 2nd imitation task, the greatest mean duration difference was exhibited in the case of FLEECE; it then gradually decreases from TRAP through DRESS to KIT. In the 2nd accommodation task, the

mean duration difference was comparable for TRAP, FLEECE and DRESS. As far as convergence patterns are concerned, the mean duration difference decreased in the 1st imitation task as compared with the baseline for all of the investigated vowels. In the 1st accommodation task, the vowel duration difference decreased for KIT but increased for the remaining three vowels. Following exposure to native pronunciation, the mean duration difference increased for all four vowels. Statistical analysis showed that the assumption of sphericity had been violated [$\chi^2(77)=108.03$, $p<.05$], therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\boldsymbol{\varepsilon}<.75$). The interaction between task and vowel category was found to be statistically significant [$F(8.1, 259.4)=5.8$, $p<.001$]. However, post hoc Bonferroni tests showed that neither the decrease between the baseline and the 1st accommodation condition observed for KIT nor the increase observed for TRAP, DRESS and FLEECE were statistically significant.

Chapter Four: Speech convergence in the pronunciation of Polish learners of English - analysis and discussion

4.1. Introduction

In this chapter, the results of the study on speech convergence in the pronunciation of Polish learners of English are analysed and discussed. The term convergence strategies will be used extensively throughout the chapter; as referred to previously, convergence strategies are considered to comprise three types of linguistic behaviour: convergence, divergence and maintenance. Convergence is treated as the process of making one's speech more similar to that of another person and was operationalised as a significant shift towards the values exhibited by a given model talker/interlocutor. Divergence, on the other hand, is viewed as the process of moving away from the speech of another person and was operationalised as a significant shift away from the values exhibited by a given model talker/interlocutor. The third strategy, maintenance, refers to the process of maintaining one's default linguistic behaviour in spite of exposure to the speech of another person and was operationalised as a non-significant difference between the subjects' default realisations and the values exhibited following exposure to the speech of a given model talker/interlocutor.

Section 4.2. of the chapter is concerned with the data obtained in the baseline condition; these results are described separately since baseline productions were used as a reference point for the examination of pronunciation shifts. Section 4.3. provides an interpretation of the questionnaire results; it examines attitudes towards English pronunciation and self-reported convergence strategies of the participants. Sections 4.4. and 4.5. are concerned with the results of statistical analysis. The former discusses convergence strategies as a function of model talker/interlocutor; the latter focuses on convergence strategies as a function of phonetic context. Sections 4.2. to 4.5. begin with a relevant research question (see Section 3.7.). A summary of the results is provided in Section 4.6.; in this section, the hypotheses formulated for the purposes of the study are tested. Section 4.7. provides an evaluation of the experimental method. It is included in the chapter since the introduction of a modified experimental procedure was a key component of the current study. The advantages and limitations of the method are discussed together with recommendations for further

modifications and improvement. The final section of the chapter offers suggestions for further research.

4.2. Baseline realisations

RQ1: How were the three phonetic variables realised in the baseline condition?

RQ1 refers to participants' baseline realisations, which will be used as a point of reference in the analysis of convergence strategies. The answer to RQ1 will be summarised and used to test the three hypotheses formulated for the purposes of the study (see Section 3.5.) in Section 4.6.

The results indicate that the participants realised /p/ and /k/ as aspirated in their baseline productions and approximated the voicing lag values reported for native English by Kopczyński (1977) and Lisker and Abramson (1964) (compare Tables 1, 2 and 18). The results also show that the subjects realised word-initial /b g/ with substantial amounts of pre-voicing and produced voicing lead values similar to those reported for native Polish by Kopczyński (1977) and Keating et al. (1981) (compare Tables 3, 4 and 20). Finally, the findings revealed that there was a considerable mean difference in duration between vowels followed by /d/ and vowels followed by /t/. Nonetheless, the mean duration difference produced by the participants was approximately half as long as the mean overall difference provided by the native English model talker (compare Tables 9 and 22) and approximately half as long as the values reported for native English in previous studies (e.g. Chen, 1970; Peterson and Lehiste, 1960; see Table 5).

The data obtained for voicing lag suggest that realising the analysed target words as aspirated did not pose great difficulties for the participants and could be interpreted to mean that the subjects succeeded in establishing new categories for English word-initial /p/ and /k/. Following James Flege and the terminology used in his Speech Learning Model (Flege, 1995), the term „new" is used here to denote sounds that are perceived as separate from their L1 equivalents and are produced native-like as a result. That the subjects had succeeded in establishing new categories for English /p k/ seems likely due to their long language experience and the fact that they had completed three semesters of formal pronunciation training during which aspiration was discussed and practised quite extensively (see Section 3.6.3.). As argued by Dziubalska-Kołaczyk (2002) and Wrembel (2005), conscious

knowledge of the existence and usage of a particular L2 pronunciation feature should facilitate its successful acquisition (see Section 2.6.).

As far as pre-voicing is concerned, the results demonstrate that the subjects generally resorted to L1 habits when producing the analysed /b g/ tokens in the baseline task and that devoicing of English word-initial stops was problematic for them. The findings could also mean that, in spite of extensive experience with learning English and the phonetic training received, the participants did not succeed in forming new categories for English voiced stops. As referred to in Section 3.6.3., devoicing of word-initial /b d g/ in English was mentioned in the theoretical component of the phonetics and phonology course the subjects attended. However, it was not covered as extensively as aspiration and vowel duration and was rarely (if ever) included in the practical component of the course, which, presumably, might have caused some participants to gradually forget about the existence of this feature and thus inhibit its successful acquisition. It should also be mentioned that the /b d g/-initial tokens were frequently confused with their /p t k/-initial counterparts in the 2nd imitation task (see Section 3.6.6.). Difficulties in correctly identifying devoiced instances of /b g/ provide further evidence for the claim that the participants did not succeed in forming new categories for English voiced stops

The results indicate that the subjects used vowel duration as a cue for consonant voicing in their baseline productions of the investigated target words. Since they did exhibit considerable differences in vowel duration but did not match the values reported for native English, the durational differences produced by the informants could be considered intermediate between L1 and L2.. Intermediate values for vowel duration in L2 pronunciation were also reported in previous studies (e.g. Flege, 1980; Waniek-Klimczak, 2005). Importantly, producing intermediate vowel length values could signify that vowel duration as a cue for consonant voicing is in the process of formation in the participants' ILs.

On the whole, the analysed phonetic variables appear to reflect different stages of acquisition of the L2 sound system. According to Major's Ontogeny Phylogeny Model (Major, 1987, 2001, 2008; see Section 2.3.), interlanguage consists of elements of L1, L2 and language universals. It is also postulated that as the learner gains more language experience, the impact of L1 gradually decreases while the effect of L2 and language universals gradually increase. The findings of the current study suggest that the subjects' realisation of initial stop devoicing was still under the influence of their L1, their implementation of aspiration was approximating the L2 norm, whereas their realisation of vowel duration as a cue for consonant voicing showed a target-like tendency but did not match native values. Thus, it could be

hypothesised that initial stop devoicing was in an early, aspiration in a final, and vowel duration in an intermediate stage of acquisition in the subjects' ILs. However, it should be stressed that these claims are based on fairly limited data, extracted from single-word utterances produced in a relatively formal setting. In order to provide a fuller account of the stage of acquisition of a given phonetic feature, it would be advisable to examine their realisation using different types of elicitation tasks. The method of elicitation could be of importance since formal tasks in which learners pay increased attention to language form may result in more native-like productions than less formal tasks such as free speech (Tarone, 1979, 1982; see Section 2.4.). Also, the claims concerning the stage of acquisition of a given phonetic feature are built on a comparison with the values exhibited by one model talker and the measurements from a couple of early studies on English temporal parameters. For a more reliable analysis, baseline productions should be contrasted with measurements taken from a representative, native-speaker reference group and juxtaposed with the subjects' realisations of equivalent sounds in their L1.

4.3. Questionnaire responses

Section 4.3.1. is concerned with attitudes towards English pronunciation; the analysis and discussion are based on the questionnaire responses that pertained to the evaluation of the model talker's phonetic performance and the subjects' opinion on the importance of using native-like pronunciation when speaking English. Section 4.3.2. describes self-reported convergence strategies; the analysis and discussion are based on questionnaire responses to multiple choice questions (concerned with the speech behaviour in the baseline and imitation tasks) and the open-ended questions (concerned with speech behaviour in the accommodation tasks).

4.3.1. Attitudes towards English pronunciation

RQ2: What are the attitudes of the participants towards L2 pronunciation in English?
RQ2 pertains to attitudinal factors that could potentially affect participants' convergence strategies; the answer to RQ2 will be summarised and used to test Hypothesis 2 (see Section 3.5.) in Section 4.6.

Questionnaire results revealed that the native speaker's pronunciation was rated higher than the non-native speaker's pronunciation: the maximum score that the model talkers could achieve was 35; the native English talker received a mean score of 30 from both subsets of the subject group (see Section 3.6.6.), whereas the native Polish talker received a mean score of 20 from the VOT group and a mean score of 21 from the V_DUR group (see Table 12). Also, SD was higher for the non-native speaker than for the native speaker in the case of both subsets of the subject group (see Table 12), which suggests that scores received by the Polish speaker were characterised by more variability than the scores received by the English speaker. The mean score achieved by the subjects in the attitudes component of the questionnaire was 50 for the VOT group and 49 for the V_DUR group (see Table 13; the maximum score in this part of the questionnaire was 65; achieving a score close to this number was expected to signify a strong bias against Polish-accented speech). In the open-ended questions included in the self-report component of the questionnaire (see Section 3.8.1.), many participants pointed to the importance of using "correct", native-like pronunciation. Several subjects suggested that the Polish talker's pronunciation was incorrect and/or that he mispronounced some of the analysed words. Conversely, pronunciation of the native speaker was often referred to as "correct" or "proper".

The findings suggest that the participants viewed the native speaker's pronunciation in a more positive light than the non-native speaker's pronunciation and generally exhibited a preference for native over Polish-accented English. The fact that they underscored the importance of using "correct", target-like realisations points to a prescriptive approach towards English pronunciation and seems to be related to the fact that the subjects were students at the Institute of English Studies, expected one day to become expert language users and training to become English teachers or translators. Also, SSBE (the accent of the native English model talker) was presented as the preferred pronunciation model at the English phonetics and phonology course the participants attended. The findings accord with the results of several previous studies in which a preference for native-like pronunciation was found among advanced Polish learners of English (e.g. Janicka et al., 2005; Waniek-Klimczak and Klimczak, 2005; Waniek-Klimczak et al., 2014; see Section 2.7.).

Interestingly, more variability was observed in the scores received by the non-native speaker than in the scores received by the native speaker. The finding could be connected with the fact that the subjects were required to assess the two model talkers/interlocutors with respect to two disparate variables: friendliness and social status (i.e. level of education, intelligence, professionalism). It is possible that some of the informants gave the Polish

speaker low scores with respect to social status because they did not approve of his heavy accent but, at the same time, gave him higher scores on friendliness because they sympathised with the speaker (e.g. due to shared L1 and nationality or the fact they were told that the non-native speaker was a fellow student of English studies).

4.3.2. Self-reported convergence strategies

RQ3: According to the participants, what convergence strategies did they use in the imitation and accommodation tasks?

RQ3 is concerned with self-reported convergence strategies, the knowledge of which is used to interpret the obtained data with respect to the effect of attitudinal factors; the answer to RQ3 will be summarised and used to test Hypothesis 2 (see Section 3.5.) in Section 4.6.

Data obtained in self-report component A of the questionnaire (see Section 3.8.1.) indicate that the majority of the participants concentrated on their pronunciation and attempted to realise the analysed words in a native-like way in the baseline and two imitation tasks. Also, several participants declared that they converged towards the non-native speaker's pronunciation in the 1st imitation task; most of them stated that they did it inadvertently. Results from self-report component B (see Section 3.8.1.) showed that in the 1st accommodation task, the majority of the subjects did not attempt to converge towards the non-native model talker by making their speech more similar to his pronunciation. Some comments implied that a few of the participants may have converged towards the Polish model talker unwittingly, even though they considered his pronunciation incorrect. Other participants stated that they attempted to adjust their pronunciation by making it more clear and intelligible. As regards convergence towards the native model talker in the accommodation task, the majority of the subjects stated that they attempted to adjust their pronunciation towards that of the English speaker so as to sound more native-like. Also, some of the remarks imply that the subjects wished to make a favourable impression on the English interlocutor.

The findings indicate that the participants generally attempted to sound native-like when producing the target words in the two imitation tasks. The strategy seems to spring from the subjects' preference for native English pronunciation coupled with the fact that the imitation tasks lacked an interlocutor towards whose pronunciation the subjects could potentially accommodate. Also, although the participants were left on their own for the duration of each

experimental task, it has been argued that individuals with whom a speaker is not in direct, face-to-face interaction may still cause them to make pronunciation adjustments (Bell, 1984; see Section 1.2.). Thus, it seems possible that the informants attempted to use more target-like realisations in order to create a favourable impression on the author of the study, whom they knew as their pronunciation instructor and who (as the subjects most probably realised) would at some point listen to their productions from the experimental procedure.

It was also found that the majority of the subjects attempted to adjust their pronunciation and sound more native-like when reading the target words for the English speaker, which indicates that they wished to converge towards the native interlocutor. As in the case of imitation, the finding appears to be linked to the subjects' preference for native English. Also, the strategy to accommodate towards the native interlocutor may have been dictated by a desire to gain his approval. As stated by Giles (1973: 90), "[...] if the sender in a dyadic situation wishes to gain the receiver's social approval then he may adapt his accent patterns towards that of this person [...]". That the informants wished to make a favourable impression on the English interlocutor seems likely also because questionnaire responses suggest that they viewed him as superior in terms of linguistic performance. Indeed, it has been argued that a speaker's status in an interaction may play an important part in speech accommodation (Zuengler, 1985, 1989; Gregory and Webster, 1996; Pardo, 2010; Pardo et al., 2013) and that native speakers tend to be assigned higher status in interactions with non-native language users (Zunegler, 1989). It is also possible that some of the subjects attempted to accommodate towards native-like values because they wished to make their speech more intelligible to the English interlocutor. Some participants did mention intelligibility when stating whether they adjusted their pronunciation towards that of the native speaker and it has been contended that one of the motives for speech accommodation may be communication efficiency (e.g. Gallois et al., 1995; Giles and Ogay, 2007, see Section 1.2.).

The results show that the overwhelming majority of the subjects did not try to adjust their speech in order to sound more Polish-like when reading for the non-native interlocutor (see Table 17), which signals that they did not wish to accommodate towards his pronunciation. If a participant did declare that they modified their pronunciation when reading for the Polish speaker, it was usually stated that they attempted to make their speech clearer and more intelligible rather than more "Polish-sounding". Also, some comments implied that a few of the participants converged towards the Polish model talker unwittingly, even though they considered his pronunciation incorrect. The findings seem interesting since the results of previous studies imply that social convergence may result in more speech adjustments

(Gregory and Hoyt, 1982; Young, 1988) and that a sense of solidarity may play an important role in phonetic accommodation (e.g. Welkowitz and Feldstein, 1969; Welkowitz and Feldstein 1970). Thus, since the participants had a shared L1 and nationality with the native Polish speaker, they could be expected to want to accommodate towards his speech in order to express solidarity. Nonetheless, the hypothesis has not been borne out by the results. It seems that the majority of the informants opted to use a strategy that could be described as maintenance. In Communication Accommodation Theory, the term refers to a situation in which a person continues to use a given speech style or phonetic feature irrespective of the pronunciation of his or her interlocutor (Giles and Ogay, 2007). The unwillingness to accommodate may have been brought about by the preference for native-like speech on the part of the subjects. Also, questionnaire responses suggest that the participants viewed the Polish speaker's pronunciation in a negative light; they rated his phonetic performance lower than the English speaker's productions and often referred his realisations as erroneous. Indeed, the voicing lag values exhibited by the Polish speaker were substantially smaller than the values reported for L1 English and L1 Polish in previous studies, while his voicing lag values were considerably greater than the values reported for L1 English and L1 Polish in previous studies. He also assimilated the TRAP/DRESS contrast (realising both vowels as Polish /e/) and the KIT/FLEECE contrast (realising both sounds as Polish /i/). Additionally, he used spelling pronunciations with respect to word-final /d/, i.e. produced it as phonetically voiced in words such as *bad, mead, bid*, etc., which is inconsistent with both Polish and English articulatory habits (Polish neutralises the phonological voiced-voiceless contrast between word-final obstruents, e.g. Wierzchowska, 1980; Ostaszewska and Tambor, 2000; in English, phonologically voiced stops are also rarely voiced phonetically, e.g. Shockey, 2003). The numerous deviations from both the TL and NL norms may have created an image of the Polish speaker as a low-proficiency learner and thus increased the reluctance to converge towards his pronunciation.

Finally, it should be noted that the reluctance to converge towards Polish-like values may have been to some extent a result of the experimental setting. The experiment took place in a relatively formal context and there was no direct, socially rich interaction between the subjects and the two speakers. If the participants had been in conversation with their interlocutors rather than simply reading target words for them to listen to at a later time, they might have wished to use different accommodation strategies. Indeed, one of the participants implied that she might have converged towards the non-native interlocutor if the task had borne more resemblance to an actual conversational interaction (see Section 3.8.1.).

It is noteworthy that although the majority of the participants declared they did not attempt to make their pronunciation more Polish-like when reading for the non-native interlocutor, one subject stated that she did try to converge towards the Polish speaker's realisations so as to facilitate mutual understanding. Interestingly, she admitted that she had attempted to accommodate even though she had been aware that the Polish interlocutor committed pronunciation errors. Previous findings suggest that speakers may converge to a greater extent if they possess certain personality traits, e.g. openness or need for social approval (e.g. Natale, 1975a; Yu et al., 2013; see Chapter One). Thus, it seems possible that the aforementioned participant possessed some personal characteristics that made her more prone to converge towards her interlocutors. Also, the finding provide some support for the claim that convergence may result from a desire to make the interaction flow more smoothly (e.g. Giles and Ogay, 2007; Gallois et al., 1995; see Chapter One). However, to ascertain whether the self-reported accommodation did indeed take place, convergence strategies of the participant need to be examined separately.

4.4. Convergence strategies as a function of model talker/interlocutor

In this section, pronunciation shifts are discussed with respect to the native/non-native status of the model talkers/interlocutors. The data obtained in the imitation condition and the results from the accommodation tasks are discussed separately. The first three subsections pertain to the non-interactional condition, the last three are concerned with the interactional tasks. The findings are further subdivided according to the phonetic variable under investigation.

4.4.1. Imitation of aspiration

RQ4: In the case of aspiration, what imitation strategies did the participants use following exposure to native and non-native speech?

RQ4 applies to speech behaviour in the imitation conditions as compared with the baseline condition; the answer to RQ4 will be summarised and used to test Hypothesis 1 and Hypothesis 2 (see Section 3.5.) in Section 4.6.

The results show that the informants produced significantly longer voicing lag in both non-interactional tasks as compared with the baseline condition. The finding implies that the participants imitated the pronunciation of the English model talker and diverged from the

pronunciation of the Polish model talker. It was also found that the subjects did not approximate the mean voicing lag values produced by English model talker (61 ms vs. 134ms, see Tables 7 and 18), which implies that they did not imitate the English speaker's productions faithfully.

Both of the observed strategies (convergence towards the native model talker and divergence from the non-native model talker) may have been an effect of bias in favour of native pronunciation and the characteristics of the experimental setting (see Section 4.2.); it seems likely that the observed imitation patterns stemmed from a desire to sound more native-like on the part of the participants. Another explanation for the increase in aspiration in both imitation tasks could be that the subjects had mastered this L2 pronunciation feature (see Section 4.3.). As a consequence, they were able to converge towards higher values when listening to the native model talker and were able to overcome L1 interference when listening to the non-native model talker.

The findings concerning the imitation of voicing lag evoke the concept of social marking (see Section 2.4.), which is closely related to the social-psychological aspect of speech convergence[7]. As argued by Giles, Scherer and Taylor (1979), speech markers can be attenuated or accentuated to indirectly communicate attitudes towards social group membership. In this case, it could be hypothesised that aspiration was the speech marker, L2 learners of English the social group to which the participants belonged, and the belief that one should strive for native-like pronunciation the attitude they wished to express. Thus, the results could be interpreted to mean that the subjects increased the amount of aspiration in their productions in order to indicate their preference for native-like pronunciation. Trudgill (1981) suggested that phonetic features that tend to become social markers are those that are placed relatively high in the speaker's consciousness. It could be argued that aspiration was relatively high in the subjects' consciousness since it was covered fairly extensively in the course of the phonetic training they received. Indeed, some of the participants stated that they attempted to use this feature in their productions (see Section 3.8.4.). Also, tokens that were used to examine the realisation of voicing lag were /p t k/-initial, single-syllable words. This type of stimuli could prime the participants to focus their attention on aspiration. However, since it was in no way empirically measured whether aspiration or any other of the investigated pronunciation features were placed high in the informants' consciousness, the claim remains tentative.

[7] Howard Giles, the founding father of Communication Accommodation Theory, was among the first to discuss speech markers in social interaction (e.g. Giles, Scherer and Taylor, 1979).

It should also be stressed that rather than maintain their default voicing lag values when exposed to the speech of the Polish model talker, the subjects diverged from his pronunciation, i.e. *increased* the amount of aspiration in their productions as compared with the baseline condition. Although the difference in mean voicing lag values between the baseline and the 1st imitation task was relatively small (8 ms, see Table 22), it was still found to be statistically significant. The finding does not lend itself to straightforward interpretation. Perhaps due to their exaggerated nature, the VOT values produced by the Polish model talker sounded overly unnatural and the participants increased voicing lag values in an attempt to make up for the artificiality of the Polish talker's productions. It also seems possible that, as referred to in the previous paragraph, aspiration functioned as a social marker in the subjects' speech and they "emphasised" it by way of indirectly expressing their view on L2 English pronunciation.

Finally, it should be mentioned that although the informants converged towards the native speaker by increasing the amount of aspiration in their realisations, they did not match the mean voicing lag values produced by English model talker. As referred to in Section 3.6.2., VOT values exhibited by the native speaker in word-initial /p t k/ were considerably higher than the values reported for English by Kopczyński (1977) and Lisker and Abramson (1964) and could be considered exaggerated. Conversely, the values produced by the participants, although significantly higher upon exposure to native speech than in the baseline, generally approximated the values reported by Kopczyński (1977) and Lisker and Abramson (1964) (compare Tables 1, 2 and 18). A similar tendency can be observed in the data obtained by Nielsen (2011), who examined the imitation of target words with artificially extended VOT values by native speakers of American English (see Section 1.3.). Nielsen reported that the participants imitated the extended values, but a careful inspection of her data reveals that although there was a significant increase in the subjects' VOTs following exposure to modelled speech, the mean values exhibited by the participants fell nowhere near the artificially extended VOTs. These observations raise the possibility that there is an upper limit to convergence phenomena and that exaggerated values are less susceptible to imitation.

4.4.2. Imitation of pre-voicing

RQ5: In the case of pre-voicing, what imitation strategies did the participants use following exposure to native and non-native speech?

RQ5 applies to speech behaviour in the imitation conditions as compared with the baseline condition; the answer to RQ5 will be summarised and used to test Hypothesis 1 and Hypothesis 2 (see Section 3.5.) in Section 4.6.

The results revealed that the subjects significantly increased the amount of pre-voicing when listening to the non-native model talker and maintained their default realisations of word-initial /b g/ upon exposure to the native talker's pronunciation. These findings imply that the participants imitated the Polish model talker's pronunciation but did not converge towards the English model talker. It was also found that in the 1st imitation task, the mean voicing lead value produced by the informants did not match the mean value of the stimuli provided by the Polish model talker (86 ms vs. 160 ms, see Tables 8 and 20), which implies that convergence was incomplete.

The finding that the participants converged on pre-voicing towards the non-native speaker but not towards the native speaker do not corroborate the results obtained for aspiration. In fact, a pattern opposite to the one found for aspiration can be observed. This seems especially interesting in light of the fact that the subjects exhibited a pro-native-pronunciation bias and mostly stated they wished to sound native-like in the non-interactional tasks (see Section 4.2.). As referred to in Section 4.3., subjects' baseline productions suggest that initial stop devoicing was in an early stage of acquisition in their IL and that it remained under the influence of L1 interference. Hence, it seems possible that it was L1 transfer that prevented convergence on devoicing and facilitated the imitation of extended pre-voicing. These observations raise the possibility that imitation strategies in L2 pronunciation may be mediated by the stage of acquisition of a given pronunciation feature.

It was argued in the previous section that the subjects may have converged on aspiration because the feature was placed relatively high in their consciousness and functioned as a social marker in their speech. Following this line of reasoning, it could be hypothesised that the subjects did not converge towards native-like values when producing the /b g/-initial tokens (even though they expressed a preference for native-like pronunciation) because devoicing of word-initial stops was relatively low in their consciousness and did not serve as a social marker in their speech. As discussed in previous sections, it could be assumed that the subjects were less familiar with this pronunciation feature since it was not covered extensively during the pronunciation training the participants underwent. Nonetheless, whether the feature was indeed low in the subjects' consciousness cannot be readily verified by the data collected in this study.

Also, the fact that the subjects converged towards L1 values even though they mostly stated they wished to sound native-like suggests that imitation of pre-voicing was automatic and unintentional. The observation seems to lend some support for the claim that phonetic imitation is to some extent an automatic reflex of the human brain and that social motivations are not a prerequisite for some degree of imitation to occur (e.g. Goldinger, 1998; Goldinger and Azuma, 2004; Shockley et al., 2004; Delvaux and Soquet, 2007; Kim, 2011; Lewandowski, 2012; see Chapter One).

It should also be noted that although the participants converged towards the non-native speaker by increasing the amount of pre-voicing, they did not approximate the voicing lead values in his realisations. Since the amount of pre-voicing in his speech was substantially higher than the values observed for Polish in previous studies (Kopczyński, 1977; Keating et al., 1981) and could be considered exaggerated, the finding seems to provide further evidence for the claim that atypical values may be less likely to be imitated faithfully (see previous section).

4.4.3. Imitation of vowel duration as a cue for consonant voicing

RQ6: In the case of vowel duration, what imitation strategies did the participants use following exposure to native and non-native speech?

RQ6 applies to speech behaviour in the imitation conditions as compared with the baseline condition; the answer to RQ6 will be summarised and used to test Hypothesis 1 and Hypothesis 2 (see Section 3.5.) in Section 4.6.

The data show that the informants decreased the mean vowel duration difference when listening to the Polish model talker and increased the mean vowel duration difference upon exposure to the English model talker's pronunciation. The findings indicate that the subjects converged towards both model talkers. The results also show that although the participants shifted their realisations towards the non-native talker's pronunciation, the vowel length contrasts between the analysed minimal pairs were not obliterated, which indicates that convergence towards the pronunciation of the Polish speaker was not complete. As regards convergence towards the English model talker, the mean vowel length produced by the subjects upon exposure to his speech was slightly longer than the mean vowel duration in the stimuli that he provided (111 ms vs. 98 ms; see Tables 9 and 22).

In general, the findings suggest that yet another imitation strategy was used in the case of vowel duration, i.e. the pattern was different from the behaviour observed with respect to aspiration and varied from convergence strategies found in the case of pre-voicing. Convergence towards the English model talker could be attributed to a desire to sound native-like, resulting from a bias in favour of target-like pronunciation and the characteristics of the experimental setting (see Section 4.2.). Also, as referred to in Section 4.3., vowel duration as a cue for consonant voicing appears to have been in an intermediate stage of acquisition in the subjects' IL, which could mean that it was mastered well enough for the subjects to be able to increase length contrasts when presented with native productions. Applying the concept of social markers, the finding could be interpreted to mean that the informants converged towards native pronunciation because vowel length as a cue for consonant voicing was relatively high in their consciousness and they accentuated the feature to indirectly communicate their attitude towards L2 English pronunciation. Some support for the claim that the feature was high in the speakers' consciousness could be found in the fact that context-depended length differences in English vowels were frequently practised and discussed during the subjects' phonetic training. Additionally, the target words that were used to examine the realisation of vowel duration were /t/- or /d/-final, single-syllable minimal pairs (see Section 3.6.2.) that may have primed the participants to concentrate on vowel duration. Nevertheless, as in the case of aspiration and pre-voicing, whether a given feature was high in the subjects' consciousness was not empirically tested in this study and so the claims obtaining to the concept of social markers remain to be verified.

Interestingly, the data suggest that upon exposure to native speech, the informants not only matched the mean vowel length difference produced by the native speaker but even slightly exceeded it. Perhaps, as argued in the previous paragraph, vowel duration as a cue for consonant voicing was placed relatively high in the subjects' consciousness and the TL input they received in the 2nd imitation task brought the feature to their attention, giving rise to increased durational contrasts.

Since the obtained data imply that the subjects converged (albeit not completely) towards the Polish model talker, the findings could also mean that imitation of vowel length was to some extent affected by L1 interference (similarly as in the case of pre-voicing). This interpretation of the results accords with the claim that vowel duration as a cue for consonant voicing was in an intermediate stage of acquisition in the participants' ILs. If the attainment was not complete, it would presumably still be permeable to L1 interference.

Moreover, similarly as in the case of pre-voicing, the finding that the participants converged towards the pronunciation of the non-native model talker despite the fact that they mostly declared they wished to sound native-like implies that imitation was unintentional and provides further support for the claim that the process is to some extent an automatic reflex of the brain (see previous section).

4.4.4. Accommodation on aspiration

RQ7: In the case of aspiration, what accommodation strategies did the participants use following exposure to native and non-native speech?

RQ7 applies to speech behaviour in the accommodation conditions as compared with the baseline condition; the answer to RQ7 will be summarised and used to test Hypothesis 1 and Hypothesis 2 (see Section 3.5.) in Section 4.6.

The data indicate that the participants maintained their default aspiration values when reading for the Polish interlocutor and that when reading for the English interlocutor, they produced significantly longer voicing lag than in their baseline realisations. These findings indicate that the informants accommodated towards the native speaker and neither converged nor diverged from the non-native speaker. The results revealed that the mean voicing lag value in the 2nd accommodation task was considerably lower than the mean value produced by the English speaker (75 ms vs. 134 ms, see Tables 7 and 18), which signifies that the participants did not converge completely.

Both convergence towards the native interlocutor and maintenance in the case of the non-native interlocutor seem to be related to bias in favour of native pronunciation and the desire to sound native-like on the part of the participants. The specific reason for accommodating towards the English speaker could be that the informants viewed him as superior in terms of phonetic performance and wished to gain his approval by approximating his pronunciation and/or wanted to make their speech more intelligible to the native interlocutor in order to facilitate communication (see Section 4.2.). The rationale behind maintaining their default aspiration values when reading for the Polish interlocutor could be that the informants perceived his pronunciation in a negative light and did not wish to accommodate towards his speech (see Section 4.2.).

The finding that the subjects accommodated towards the TL norm when reading for both interlocutors could also be interpreted using the concept of social marking, i.e. it could be

considered to mean that long voicing lag was relatively high in the speakers' consciousness and functioned as a social marker in their pronunciation (see Section 4.4.1.). It is also possible that the subjects were able to use long voicing lag in the interactional tasks because they had mastered this feature of English pronunciation (see Section 4.3.). What is more, producing the target words as aspirated (especially in the 2nd accommodation task) may have been facilitated by the effect of practice (e.g. Heiman, 2002). The experiment was based on a repeated measures design and the participants produced the analysed tokens under five different conditions (see Section 3.6.4.). It seems plausible that being made to repeat the exact same words several times may have improved the subjects' phonetic performance, especially in the final experimental condition (i.e. the 2nd accommodation task).

It was also found that the participants did not match the aspiration values provided by the native interlocutor. In fact, a substantial gap can be observed between the mean voicing lag values produced by subjects and the mean values exhibited by the English speaker. This finding could be related to the fact that the participants read the target words several minutes after listening to the native interlocutor's realisations. It seems reasonable to assume that the longer the delay between the perception of the stimulus and its reproduction, the weaker the tendency to imitate. On the other hand, the native speaker produced voicing lag values that were considerably longer than the ones reported for L1 English in previous studies (e.g. Lisker and Abramson, 1964; Kopczyński, 1977, see Section 3.6.1.). Perhaps the informants found them to be somewhat exaggerated and did not fully converge for this reason. A similar tendency was observed in the 2nd imitation task (see Section 4.4.1.), where it was found that the informants converged towards the native speaker by increasing the amount of aspiration in their realisations but did not approximate the mean voicing lag values from the stimuli provided by the native English speaker. To reiterate, these observations could mean that exaggerated values are less likely to undergo speech convergence (see Section 4.4.1. and Section 4.4.2.).

Interestingly, a comparison of pronunciation shifts in the imitation and accommodation condition shows that there are certain similarities between convergence strategies used in the two types of tasks. Firstly, the subjects converged towards the native English speaker in both the non-interactional and the interactional task. Admittedly, the increase in aspiration was considerably more marked in the imitation condition (the difference was found to be highly statistically significant, see Section 3.8.2.). However, the discrepancy can be ascribed to the characteristics of the experimental procedure. In the imitation tasks, the informants produced the target words immediately after hearing them, while in the accommodation tasks, there was

a several-minute delay between exposure and production. Also, in the intervals between the tasks, the subjects received instructions on what to do in the next block of the experiment in Polish. It seems reasonable to assume that the time delay and the distraction lessened imitative tendencies, leading to a smaller degree of convergence in the accommodation task. Somewhat similar results were obtained by Rojczyk (2012) and Rojczyk et al. (2013), who investigated immediate and distracted imitation of native English by Polish learners and observed that distracting the participants (i.e. asking them to read a number prior to the imitation of modelled speech) reduced the tendency to imitate (see Section 3.2.). Generally, convergence strategies observed upon exposure to native speech could be considered instances of one and the same pattern, varying in intensity depending on the type of experimental task. Interestingly, convergence strategies that were used following exposure to non-native speech could also be treated in a similar manner. It was found that the informants realised the analysed words as aspirated both in the 1st imitation and the 1st accommodation task, the chief difference between the two being that there was a slight increase in aspiration in the non-interactional task as compared with the baseline condition. Overall, the observation that the subjects appear to have used comparable convergence strategies regardless of whether or not a given experimental task included an interlocutor seems to lend support to the claim made in Chapter One that the types of speech behaviour examined under the names of imitation and accommodation are in fact instances of one and the same process and can be viewed as complementary. Methodological and terminological differences could lead one to believe that imitation and accommodation are two separate phenomena, the former an automatic reflex of the brain, devoid of social-psychological motivations and the latter associated solely with social interaction. Nonetheless, the findings of the current study suggest that the two terms represent different aspects of a more general process - a natural tendency to converge towards the speech of another person or people that can be impeded or reinforced by social-psychological and linguistic factors.

4.4.5. Accommodation on pre-voicing

RQ8: In the case of pre-voicing, what accommodation strategies did the participants use following exposure to native and non-native speech?

RQ8 applies to speech behaviour in the accommodation conditions as compared with the baseline condition; the answer to RQ8 will be summarised and used to test Hypothesis 1 and Hypothesis 2 (see Section 3.5.) in Section 4.6.

The results show that the participants maintained their default pre-voicing values in both accommodation tasks, which indicates that they did not accommodate either towards the native or the non-native interlocutor as far as the realisation of word-initial /b g/ is concerned.

The lack of accommodation towards the non-native speaker can be explained by the fact that the participants did not wish to converge towards Polish-like values, exhibited a bias in favour of native-like pronunciation and viewed the Polish interlocutor's pronunciation as incorrect (see Section 4.2.). At the same time, although the subjects showed a preference for native English, they still resorted to L1 habits when realising the /b g/-initial tokens, both when reading for the Polish and the English interlocutor. The explanation could be that initial obstruent devoicing was likely in an early stage of acquisition in the informants' ILs (see Section 4.3.) and so they were unable to produce more native-like values. Lack of accommodation towards the TL norm could also be attributed to the fact that the feature was presumably quite low in the subjects' consciousness and did not function as a social marker in their speech (see Section 4.4.3.). Another factor that may have contributed to the occurrence of long voicing lead in the subjects' realisation could be that they wished to sound intelligible (which seems likely in the light of the fact that they were told their interlocutors would later assess how easy or difficult it was to understand their speech). Thus, it seems possible that they maintained voicing in word-initial lenis stops in order to increase and highlight the phonetic contrast between /p/-/b/ and /k/-/g/ minimal pairs.

Another interesting observation is that the subjects appear to have used comparable convergence strategies in the imitation and accommodation conditions. Following exposure to the English speaker's pronunciation, the subjects maintained their baseline realisations of word-initial /b g/. Following exposure to the Polish speaker's pronunciation, the informants first converged towards his pronunciation (imitation condition) and then reverted to their default realisations (accommodation condition), all the while producing the analysed stops with long voicing lead. Thus, convergence strategies upon exposure to non-native speech could be regarded as gradations of one and the same pattern. As referred to in the previous section, the discrepancy in the magnitude of convergence could be attributed to the fact that production was immediate in the imitation task and delayed and distracted in the accommodation task. Another possibility is that the lengthening of voicing lead in the imitation condition was caused by increased cognitive demands. In the accommodation tasks,

the informants were solely required to read the target words, whereas in the imitation tasks, they were instructed to identify the words they heard and produce them. Additionally, high rate of misidentifications in the 1st imitation tasks suggests that recognising the target words produced by the non-native speaker was especially challenging for the participants (see Section 3.6.6.). One participant admitted that identifying the words provided by the Polish model talker was sometimes difficult (see Section 3.8.1.). Another informant stated that she concentrated most of her attention on the Polish talker's speech and became less focused on her own pronunciation as a consequence (see Section 3.8.1.). Increased cognitive load during the 1st imitation task may have reinforced the tendency to converge and made the subjects' ILs more susceptible to L1 interference.

In general, the observation that the subjects seem to have used parallel convergence strategies in the interactional and non-interactional conditions seems to provide further evidence for the claim that the processes examined under the names of imitation and accommodation are instances of one and the same phenomenon (see previous section).

4.4.6. Accommodation on vowel duration as a cue for consonant voicing

RQ9: In the case of vowel duration, what accommodation strategies did the participants use following exposure to native and non-native speech?

RQ9 applies to speech behaviour in the accommodation conditions as compared with the baseline condition; the answer to RQ9 will be summarised and used to test Hypothesis 1 and Hypothesis 2 (see Section 3.5.) in Section 4.6.

The results indicate that the participants used default vowel duration values when reading for the non-native interlocutor and increased vowel length differences when reading for the native interlocutor. These findings imply that the informants accommodated towards the English speaker but not towards the Polish speaker. The mean vowel duration difference produced by the subjects when reading for the native interlocutor was considerably lower than the mean difference in the English speaker's realisations (74 ms vs. 134 ms, see Tables 9 and 22), which can be interpreted to mean that convergence towards the native interlocutor was not complete.

The rationale behind the lack of accommodation towards more "Polonised" values could be that the informants exhibited a bias against foreign-accented pronunciation, were reluctant to converge towards the Polish interlocutor and regarded him as inferior in terms of phonetic

performance (see Section 4.2.). Another explanation could be that vowel length as a cue for consonant voicing was relatively high in the subjects' consciousness and functioned as a social marker in their speech (see Section 4.4.3.), thus preventing them from shifting towards more Polish-like realisations of this feature. The finding that the subjects used vowel length contrasts in all of the analysed minimal pairs may also be connected with the fact that they had already exhibited durational contrasts in their baseline productions, which suggests that they had been fairly successful in the acquisition of this pronunciation feature and so its realisation did not pose great difficulties (see Section 4.3.).

Another interesting possibility is that the participants did not accommodate towards the non-native interlocutor because converging towards values characteristic of the L1 would mean neutralising the contrast between the /t d/-final minimal pairs (Polish is said to neutralise the phonological voiced-voiceless contrast between word-final obstruents; see Section 3.6.1.). Thus, accommodating towards Polish-like values could result in ambiguity, which the informants presumably wished to avoid. Some participants did state in the questionnaire that they attempted to make their pronunciation clearer and more intelligible when reading for the Polish interlocutor (see Section 3.8.1.).

As far as accommodation towards the native interlocutor is concerned, it seems that the informants used this strategy for reasons similar to the ones mentioned in relation to aspiration (see Section 4.4.4.). On the whole, the observed accommodation pattern appears to be connected with the subjects' preference for native-like pronunciation. The subjects may have accommodated because they viewed the English interlocutor as superior in terms of phonetic performance and wished to gain his approval by using more English-like values. It is also possible that the informants converged towards the TL norm in order to make their pronunciation more intelligible to the native interlocutor. Interpreting the results using the concept of social markers (see Section 4.4.3.), it could also be assumed that the participants increased vowel duration differences when reading for the English interlocutor by way of communicating their preference for native-like pronunciation in L2 English. Another explanation could be that vowel duration as a cue for consonant voicing was mastered well enough for the subjects to be able to increase length contrasts following exposure to native input (the subjects did exhibit vowel length differences in their baseline productions, see Section 4.3.). Finally, as referred to with respect to accommodation on aspiration (see Section 4.4.4.), increased vowel length contrasts in the 2nd accommodation task could also be a consequence of the effect of practice (i.e. the more practice the more native-like performance).

It is also noteworthy that accommodation towards the pronunciation of the native interlocutor was not complete, i.e. the participants shifted their realisations towards those of the English speaker but did not match the mean vowel duration difference that he produced. The mean vowel length difference in the subjects' productions was also slightly lower than the values reported for native English in previous studies (e.g. Peterson and Lehiste, 1960; Chen, 1970; see Section 3.6.2.). Incomplete accommodation towards the TL norm is possibly linked to the fact that vowel duration as a cue for consonant voicing had not been fully acquired by the informants (the subjects used vowel length differences in their baseline productions but did not match native-like values, see Section 4.3.), thus preventing them from approximating the values provided by the native interlocutor. The finding could also be related to the time delay between exposure to the stimulus and the production of the target words (see Section 4.4.4.).

Finally, it should be mentioned that convergence patterns exhibited in the imitation and accommodation conditions could be seen as instances of the same strategy that varied in magnitude depending on the characteristics of the experimental procedure. The results revealed that the informants maintained a vowel length contrast in both the 1st imitation and the 1st accommodation condition and that in the former, they converged towards the non-native speaker by slightly decreasing vowel length differences. It was also found that the subjects converged towards the native English speaker in both the interactional and the non-interactional task and that the increase in vowel length contrast was more marked in the imitation condition (the difference in duration between the 2nd imitation task and the 2nd accommodation task was found to be highly statistically significant, See Section 3.8.4.). In both instances, the variation in the magnitude of convergence between the imitation and accommodation condition could be explained by the time delay and distraction in the interval between the two tasks (see Section 4.4.4.). In the case of the non-native speaker, the decrease in vowel length differences in the non-interactional task may also be connected with higher cognitive demands involved it its performance (see Section 4.4.5.). All in all, the observations agree with the statements made in the two previous sections in that they seem to corroborate the assumption that imitation and accommodation can be considered as two facets of one and the same process (see Section 4.4.4.).

4.5. Convergence strategies as a function of phonetic context

In this section, pronunciation shifts are discussed with respect to different phonetic contexts (place of articulation, vowel category). The data obtained in the imitation and accommodation conditions are discussed together. For this reason, the term *convergence* strategies is used (as opposed to the use of *imitation* and *accommodation* strategies in the previous sections). The findings are divided into subsections according to pronunciation feature.

4.5.1. Convergence on aspiration

RQ10: In the case of aspiration, what convergence strategies did the participants use with respect to different places of articulation?
RQ10 refers to the effect of phonetic context on convergence strategies. The answer to RQ10 will be summarised and used to test Hypothesis 3 (see Section 3.5.) in Section 4.6.

The results revealed no significant interaction between task and consonant category, which indicates that convergence strategies in the imitation and accommodation conditions did not vary as a function of place of articulation.

It was found in one of the previous studies on L2 phonetic convergence that place of articulation may have an impact on the magnitude of imitation. Rojczyk et al. (2013) examined Polish learners' realisation of release burst in English stop sequences following exposure to native pronunciation. Release bursts were analysed in two phonetic contexts: in homorganic and heterorganic clusters. The results indicated that the duration of release burst was imitated by the participants in homorganic clusters, but not in heterorganic clusters, which was ascribed it to the fact that unlike stops in heterorganic clusters, plosive consonants in homorganic sequences can be optionally unreleased in the subjects' L1, thus facilitating imitative tendencies in the latter case. In the current study, however, no significant effect of place of articulation on convergence strategies was found. Admittedly, the two phonetic contexts in Rojczyk et al.'s (ibid.) study (homorganic vs. heterorganic clusters) differed between each other in the sense that the pronunciation feature that was investigated (release burst) can have different allophonic realisations in the learner's L1 in these two environments. No such claim can be made about the phonetic contexts analysed with respect to aspiration in the current study. It does not seem likely that one of the analysed sounds could be realised as aspirated and the other as unaspirated as a result of transferring of some L1 allophonic rule. In

fact, although /p t k/ are typically produced with short voicing lag in Polish, *all three* sounds may optionally be realised as aspirated when articulated forcefully (Wierzchowska, 1980). It is conceivable that learners' convergence strategies could vary as a function of place of articulation if, for some reason, the speakers managed to successfully acquire aspiration in one of the investigated phonetic contexts but not the other. Nonetheless, both /p/ and /k/ exhibited relatively long mean voicing lag in the baseline task (see Table 19), which suggests that realisations in the two phonetic environments were broadly similar in terms of stage of acquisition.

It noteworthy than voicing lag in /k/ was consistently realised as longer than in /p/ across all experimental conditions (see Table 19). It has been argued that there is a universal tendency among languages for velar stops to have longer VOT values than alveolar and bilabial stops (e.g. Cho and Ladefoged, 1999). Thus, the results seem to provide further evidence for the claim made in Major's (1987, 2001, 2008) Ontogeny Phylogeny Model that interlanguage consists not only of elements of L1 and L2 but may also comprise language universals. Similar results with respect to aspiration were obtained by, among others, Waniek-Klimczak (2002, 2005) and Piotrowski (2013) (see Section 2.3.).

4.5.2. Convergence on pre-voicing

RQ11: In the case of pre-voicing, what convergence strategies did the participants use with respect to different places of articulation?

RQ11 refers to the effect of phonetic context on convergence strategies. The answer to RQ11 will be summarised and used to test Hypothesis 3 (see Section 3.5.) in Section 4.6.

The results revealed no significant interaction between task and consonant category, which indicates that convergence strategies in the imitation and accommodation conditions did not vary as a function of place of articulation.

The results obtained for pre-voicing agree with the findings concerning aspiration (see previous section). Similarly as in the case of voicing lag, it seems unlikely that one of the analysed sounds (/b/ vs. /g/) could be realised as pre-voiced and the other completely devoiced due to some L1 allophonic rule. What is more, baseline data suggests that irrespective of the place of articulation, both consonants were realised with substantial amounts of pre-voicing, which implies that they represented roughly the same stage of acquisition, i.e. were both affected by L1 transfer. Perhaps realisations in the two phonetic

contexts (bilabial vs. velar) were similar enough (in terms of factors such as stage of acquisition and positive L1 transfer) not to cause the participants to use different convergence strategies.

Interestingly, the results show that voicing lead in /b/ was slightly longer than in /g/ under all experimental conditions except for the baseline (see Table 21). A similar tendency can be seen in the data obtained for native Polish by Kopczyński (1977) and Keating et al. (1981) (see Tables 4 and 5). The observation seems to provide further evidence for the claim that subjects' realisation of word-initial /b g/ may have been strongly affected by L1 interference (see Section 4.3.).

It was also found that standard deviation values for pre-voicing were considerably higher than those observed for aspiration (see Tables 18 and 20), which is likely a result of the fact that voicing lead values exhibited by the participants ranged from 0 up to over 200 ms and varied both within and between speakers. The high degree of variability could also be interpreted to mean that word-initial devoicing was in an early stage of acquisition in the subjects' ILs.

4.5.3. Convergence on vowel duration

RQ12: In the case of vowel duration, what convergence strategies did the participants use with respect to different vowel categories?

RQ12 refers to the effect of phonetic context on convergence strategies. The answer to RQ12 will be summarised and used to test Hypothesis 3 (see Section 3.5.) in Section 4.6.

The results show that vowel duration difference decreased for the KIT vowel and increased for the remaining three vowels in the 1st accommodation task. However, neither the decrease between the baseline and the 1st accommodation condition observed for KIT nor the increase observed for TRAP, DRESS and FLEECE were found to be statistically significant. These findings imply that convergence strategies in the imitation and accommodation conditions did not vary as a function of vowel category.

The results of one of the pilot studies (Zając, 2013) indicated that convergence strategies varied according to vowel category. More specifically, it was found that the subjects converged towards the native speaker on vowel duration in some vowels but not in others. A similar tendency was also observed in the case of convergence towards non-native speech. The results of the current study do not corroborate these findings. Admittedly, convergence

strategies in the pilot study were found to vary with respect to imitation of vowel duration and not vowel duration *contrasts* (as is the case in the present investigation). It is possible that, had one examined vowel length rather than vowel length differences, variable convergence strategies would be also detected in the data obtained in the current study. A plausible explanation for the lack of variable convergence strategies could be that, as in the case of aspiration and pre-voicing, vowel duration as a cue for consonant voicing was in a similar stage of acquisition in each of the analysed phonetic contexts. Baseline data show that although the magnitude of the contrast varied with each category, the subjects used greater vowel length before voiced than voiceless consonants in all of the analysed minimal pairs (see Table 23).

Interestingly, although neither the decrease between the baseline and the 1st accommodation condition observed for KIT nor the increase observed for TRAP, DRESS and FLEECE were found to be statistically significant, a closer examination of the data revealed that while in the case of TRAP, DRESS and FLEECE, the greater part of the subjects (over 50% in the case of FLEECE and over 60% in the case of TRAP and DRESS) increased vowel duration difference when reading for the Polish interlocutor, an opposite tendency could be observed with respect to KIT, i.e. .it was found that 67% of the subjects decreased vowel duration difference when reading words that contained /ɪ/ for the Polish speaker. It is possible that more participants accommodated towards the non-native interlocutor by decreasing the vowel duration difference for KIT because they associated the vowel with reduced length and, as a result, were less concerned with its durational characteristics when reading the target words in the 1st accommodation task. It is assumed that the subjects may have associated the vowel with reduced length since it is intrinsically short in SSBE (see Section 3.6.1.). and was typically referred to as *short 'i'* during the subjects' pronunciation training. All in all, these observations suggest that there may be a weak relationship between convergence strategies and vowel identity and that the effect could potentially be related to a given vowel's inherent durational characteristics in the TL.

It should also be pointed out that some of the results raise the possibility that the *extent* rather than the direction of speech convergence may have been affected by vowel identity. The data presented in Figure 11 suggest that the increase in vowel duration difference for the FLEECE vowel in the 2nd imitation task as compared with the baseline task was considerably greater than the increase in vowel duration difference for TRAP, DRESS and KIT. The observation could imply that the degree of convergence towards the native model talker was

greater for /iː/ than for the remaining three vowels. The FLEECE vowel is inherently longer in SSBE than TRAP, DRESS or KIT (e.g. Wells, 1962; see Section 3.6.2.) and was often referred to as *long 'i'* during the pronunciation course the informants attended. Thus, it seems reasonable to assume that the subjects associated English /iː/ with extended length and focused on its durational characteristics more than in the case of /æ e ɪ/. If so, it is possible that when reading the target words containing FLEECE, the informants increased its duration to a greater extent than in TRAP, DRESS or KIT. This may have resulted in greater context-dependent contrasts in the 2nd imitation task and could explain the apparent difference in the degree of convergence. Nevertheless, further analysis is necessary in order to verify the hypothesis that the degree of convergence differed as a function of vowel category.

It is also worth mentioning that the duration difference for KIT was consistently realised as the smallest among the three vowels (see Table 23). The pattern appears to be related to the fact that, as referred to in one of the previous paragraphs, the vowel is intrinsically short in SSBE (see Section 3.6.1.). It was also mentioned in one of the previous paragraphs in this section that the KIT vowel was typically referred to as *short 'i'* during the subjects' pronunciation training. Thus, it could be assumed that they associated the vowel with reduced length and had received sufficient amount of TL input (the subjects had long experience with learning English) to be able to produce its durational characteristics in a native-like manner. Consequently, they may have realised the KIT vowel as shorter, which, in turn, resulted in smaller durational contrasts between the analysed minimal pairs.

Another interesting observation is that in the first three experimental tasks, the mean duration difference was greatest for the TRAP vowel (see Table 23) even though the /æ/ of SSBE is typically categorised as a short vowel (e.g. Shockey, 2013). The American TRAP vowel, on the other hand, is sometimes classified as intrinsically long (ibid.). It is probable that the subjects received a considerable amount of American English input through media (music, film, tv series, etc.). If so, it may have influenced their realisation of TRAP and resulted in the production of longer durational characteristics (increasing vowel length in isolated words could lead to increased duration differences between the minimal pairs).

4.6. Summary of the results and hypotheses testing

This section of the dissertation summarises the results of the study on speech convergence in the pronunciation of Polish learners of English and addresses the three hypotheses formulated

for the purposes of the investigation. Hypothesis 1 is tested on the basis of the results that were analysed and discussed in Sections 4.2. and 4.4. (RQ1 and RQs4-6); Hypothesis 2 is tested on the basis of the results analysed and discussed in Sections 4.2., 4.3. and 4.4. (RQs1-9); Hypothesis 3 is tested on the basis of the results analysed and discussed in Section 4.5. (RQs10-12).

H1: Convergence strategies following exposure to native and non-native English differs as a function of model talker/interlocutor.

The data obtained for voicing lag indicate that the subjects converged towards native pronunciation and diverged from non-native pronunciation in the imitation tasks. In the accommodation tasks, the participants converged towards native English and neither converged nor diverged from Polish-accented English. The results obtained for pre-voicing imply that in the imitation condition, the informants converged towards the non-native model talker and neither converged nor diverged from the native model talker. In the accommodation condition, no pronunciation shifts were observed either following exposure to Polish-accented or native English. In the case of vowel duration as a cue for consonant voicing, the findings show that the subjects converged both towards native and non-native pronunciation in the imitation tasks. In the accommodation tasks, the participants converged towards native English but did not shift their pronunciation following exposure to Polish-accented English. Taken together, the results suggest that speech behaviour following exposure to native and non-native English varied as a function of model talker/interlocutor in all but two instances (accommodation on pre-voicing and imitation of vowel duration) and provide partial support for Hypothesis 1. The finding suggests that when using a second language, speakers may use different convergence strategies depending on the native/non-native status of the model talker or interlocutor. It is noteworthy that the results of the study agree with some of the previous findings (e.g. Beebe, 1977; Berkowitz, 1986; Lewandowski, 2012; Rojczyk et al., 2013; Trofimovich and Kennedy, 2014; see Section 3.2.), i.e. they seem provide further evidence for the claim that the process of phonetic convergence (whether examined in an interactional or a non-interactional setting) does operate in L2 speech.

H2: Convergence strategies following exposure to native and non-native English are affected by the subjects' attitudes towards native and Polish-accented English.

Questionnaire responses revealed that the informants generally favoured native-like over Polish-accented English and perceived the Polish speaker's pronunciation more negatively

than the English speaker's accent. The majority stated that they attempted to sound native-like in the imitation tasks, tried to accommodate their pronunciation towards the L2 norm when reading for the native interlocutor and were reluctant to converge towards more Polish-like values when reading for the non-native interlocutor. These findings suggest that all instances of convergence towards the native English speaker and divergence from the Polish speaker stemmed from a preference for target-like pronunciation and were dictated by a desire to sound native-like. Similar motives seem to have prompted some of the instances of a *lack* of a pronunciation shift. Namely, maintenance observed with respect to aspiration and vowel duration seems to be related to bias against foreign-accented speech. Generally, the results imply that much of the subjects' linguistic behaviour was affected by their attitude towards English pronunciation and lend support to previous findings that attitudinal factors may influence phonetic convergence even in controlled experimental settings (Babel, 2009; Babel, 2010).

In spite of the subjects' aversion to foreign-accented English, some instances of convergence towards the Polish speaker were also observed (the informants imitated the non-native speaker's realisation of pre-voicing and vowel duration in the non-interactional task). Additionally, the results indicate that the participants failed to converge towards the native interlocutor on voicing lead. These patterns can be explained by the data collected in the baseline condition, which indicate that the phonetic variables under investigation could be arranged according to how closely they matched native-like values. Values obtained for pre-voicing resembled those reported for L1 Polish, vowel duration as a cue for consonant voicing was realised with values intermediate between L1 and L2, while the implementation of aspiration approximated the L2 norm. Importantly, the three phonetic variables could be likewise ordered based on the direction of convergence in the imitation and accommodation tasks. As regards pre-voicing, the subjects either maintained Polish-like realisations or converged towards more "Polonised" productions. Th data obtained for vowel duration indicate that the participants converged towards both the L1 and the L2 norm. In the case of aspiration, the informants either maintained native-like realisations or converged towards increased values. Taken together, the results suggest that the degree and direction of convergence on a given L2 pronunciation feature may be conditioned by its stage of acquisition. The findings imply that phonetic alignment with the native language is more likely if a learner has not fully mastered a given L2 pronunciation feature and that features that are in later stages of acquisition are more permeable to invasion from the target language.

In conclusion, the findings of the study lend partial support for Hypothesis 2. It was found that a preference for target-like pronunciation may prompt learners to converge towards native speech and diverge from foreign-accented speech. However, the factor does not seem to operate if a learner has not succeeded in mastering a given TL pronunciation feature. In other words, the influence of attitudinal factors on the magnitude of convergence in L2 speech appears to be mediated the stage of acquisition of a given L2 pronunciation feature.

H3: Convergence strategies following exposure to native and non-native English vary as a function of phonetic context.

The hypothesis is not supported by the data obtained in the study. The results show that convergence strategies did not vary as a function of place of articulation (bilabial, velar) or vowel category (TRAP, DRESS, KIT, FLEECE). The findings do not corroborate some of the previous observations on the effect of phonetic environment on the magnitude of convergence in L2 speech (e.g. Rojczyk, 2012; Rojczyk et al., 2013; Zając, 2013). However, it is worth mentioning that although statistical analysis of the results revealed no significant effect of vowel category on convergence strategies, a closer examination of the data showed that when reading for the Polish speaker in the 1st accommodation task, the greater part of the subjects increased vowel duration difference in words containing TRAP, DRESS and FLEECE but decreased vowel duration difference in words containing KIT (see Section 4.5.3.). The observation suggests that there may exist (albeit rather weak) relationship between convergence strategies and vowel identity. Additionally, the data raise the possibility that the degree rather than direction of speech convergence may have been to some extent affected by vowel identity (i.e. it was observed that the increase in vowel duration difference for the FLEECE vowel in the 2nd imitation task as compared with the baseline task was considerably greater than the increase in vowel duration difference for TRAP, DRESS and KIT; see Section 4.5.3.). The issue requires further investigation and could be addressed in a follow-up study. It should also be recognised that even though convergence strategies did not differ depending on phonetic context, the findings suggest that they did vary as a function of phonetic variable. As referred to with respect to Hypothesis 2, slightly different convergence strategies were used in the case of each of the investigated pronunciation features and the findings suggest that the magnitude of convergence may be interrelated with the stage of acquisition. At the same time, statistical analysis of the three phonetic parameters was not conducted on the exact same set of participants (see Section 3.6.6.). For this reason, speech behaviour with respect to each of the three features cannot be directly compared. In order to

provide stronger support for the hypothesis that convergence strategies varied depending on phonetic variable, additional statistical analysis should be conducted on the same group of subjects.

4.7. Evaluation of the method

As referred to in Section 3.4., one of the limitations of the pilot studies was that the experimental design lacked interlocutors towards whose pronunciation the participants could converge, thus making it difficult to determine whether the observed pronunciation shifts were generated by the characteristics of the experimental setting or resulted from a bias against foreign-accented speech. For this reason, an attempt was made in the current study to improve the experimental design of the pilot work by making the model talkers act as interlocutors.

Data collected from the questionnaires suggest that the attempt to establish the model talkers as interlocutors was successful. Some of the responses to Qs 9 and 18 (in which the participants were asked whether they adjusted their pronunciation in the accommodation tasks) indicate that the subjects believed the author's assertion that the model talkers would later listen to and evaluate their pronunciation, e.g.

- *I tried to pronounce the words clearly and carefully so that the person that's going to listen to me doesn't have any doubts as to which words I'm pronouncing.*
- *[...]I tried to sound similar to make sure that he would understand me.[...]*
- *[...]I tried to sound the same [as him], because I didn't want him to think I can't speak English very well.*
- *[...]I tried to sound as best as I could. It's more difficult for native speakers to understand accents [...]*

Thus, the modifications introduced into the design of the study seem to provide a fairly effective method of combining a controlled experimental setting with an element of social interaction (the presence of an interlocutor). One of the main advantages of this experimental method is that it makes it possible to control for phonetic context and the number of investigated tokens, rendering the analysis easier and more reliable. Another considerable advantage is that the procedure does not require direct interaction between the participants

and their interlocutors, relieving one of the need to find speakers that would be willing to devote their time and energy to converse with dozens of participants.

Nevertheless, it should be pointed out that the experimental procedure proved more time-consuming than initially expected. Completing all six blocks of the experiment took each participant approximately 30 minutes, which seems an inordinate amount of time compared with the fact that the experimental procedure lasted for 10 to 15 minutes per participant in the pilot studies (Zając, 2013; Zając and Rojczyk, 2014). What is more, 30 minutes would presumably be sufficient to have the participants engage in an actual conversational interaction with an interlocutor.

Another weak point in the methodology were the numerous misidentifications in the imitation tasks, which resulted in a severe reduction in the number of investigated words and necessitated basing the statistical analysis on subsets of the subject group. As referred to in Section 3.6.5., the subjects listened to pre-recorded realisations of the target words and were required to identify them by saying them out loud and marking their answer on the exercise sheet. A forced-choice procedure was used to facilitate the identification of the stimuli; the subjects were required to select one item from minimal pairs that contrasted voiced and voiceless stops (e.g. *Dan-tan, goat-coat, mid-mitt, bed-bet*). The most frequent misidentifications included mistaking /p t k/-initial tokens for /b d g/-initial tokens in the stimuli provided by the Polish speaker and confusing /b d g/-initial tokens with their /p t k/-initial counterparts in the stimuli provided by the English speaker. Hence, it would appear that the minimal pairs selected for forced-choice recognition were often too difficult to the participants to distinguish. Contrasting sounds that would be easier to differentiate for Polish learners may have resulted in fewer identification problems (e.g. *bed-bid, bad-mad, dog-fog,* etc.). On the other hand, using minimal pairs that are too easy to distinguish could have affected the results of the study by arousing suspicions as to the pretend purpose of the experiment (i.e. determining whether it is easier to understand the speech of native English or native Polish users of English, see Section 3.6.5.). Overall, the elicitation method used in the imitation tasks seems to have been effective in the sense that it detracted the subjects' attention from the real purpose of the experiment and helped convince them that the model talkers would listen to and evaluate their productions. At the same time, it can be seen that the procedure was not entirely successful as it brought new methodological problems.

One other area that could be improved upon concerns the repeated measures design of the study. It was argued that since the participants produced the analysed tokens under as many as five different conditions, their phonetic performance might have been affected by the effect of

practice (see Section 4.4.4. and Section 4.4.6.). A possible solution to this problem could be to divide the experimental procedure into two sessions that would be conducted over the period of a few days or weeks. Each session could involve exposing the subjects to different type of stimuli (native vs. non-native speech); the order in which the informants participate in the sessions could be counterbalanced.

It is also important to note that despite the addition of interlocutors, the experimental procedure used in the current study cannot be considered equally socially rich as a typical conversational interaction. As referred to in Section 4.2., it is possible that the reluctance to converge towards L1-accented speech was partly related to the characteristics of the situational context. If the participants had been recorded while in conversation with their interlocutors, perhaps the observed convergence strategies would have been somewhat different. In a direct, face-to-face conversational interaction, the interacting partners can develop a rapport, their relationship can become more personal. This type of situational context could create a stronger desire to accommodate and thus counterbalance bias against foreign-accented speech. All in all, it seems that the procedure used in the current study constitutes a relatively effective method of examining phonetic convergence in a controlled experimental setting. Nevertheless, in order to provide a fuller understanding of pronunciation shifts in L2 speech, the method should be complemented by the analysis of actual conversational interactions.

Another modification introduced to the methodology of the study was the addition of the questionnaire, the purpose of which was to gauge attitudes towards English pronunciation and provide a fuller account of subjects' convergence strategies. The questionnaire included questions based on Likert-type scale that were designed to measure attitudes towards the pronunciation of the model talkers/interlocutors and L2 English pronunciation in general. It also included multiple choice and open-ended questions whose aim was to elicit subjects' assessment of their own speech behaviour. Interestingly, it was the open-ended questions that proved the most insightful, as they provided valuable information about the informants' attitude towards L2 English pronunciation - the remarks about mispronunciations in the Polish speaker's productions and the importance of sounding native-like and "correct" were made in the open-ended part of the questionnaire (see Section 3.8.1.). The observations suggests that open-ended questions may be more suitbale for the examination and interpretation of convergence phenomena. An interesting (though time-consuming) alternative could be to interview the subjects about their speech behaviour following the experimental procedure.

Such an approach could make it possible to provide a very detailed and comprehensive interpretation of the findings.

4.8. Suggestions for further research

Further analysis could involve focusing on individual cases and determining whether speech behaviour of any of the subjects deviated from the general patterns observed in the data, thus providing a more detailed insight into their convergence strategies. Another possibility could be to group the participants according to how closely they matched the TL norm in their baseline productions in order to investigate whether the magnitude of convergence varied as a function of pronunciation accuracy. If the data show that informants who were able to produce a given pronunciation feature in a native-like manner in the baseline condition converged towards the TL norm, while subjects who used L1-accented realisations in their baseline productions did not converge towards native-like values (in spite of a self-reported desire to sound native-like), the findings would provide additional support for the claim that pronunciation shifts may be conditioned by the stage of acquisition of a given L2 pronunciation feature (see Section 4.7.). Additionally, a case study could be conducted using the data obtained from the one participant who stated she attempted to accommodate towards L1-accented speech even though she believed it to be erroneous (Section 4.2.). It might prove interesting to verify whether the self-reported accommodation did indeed take place and which phonetic features it applied to.

The results of the study suggest that the magnitude of imitation may have been greater for some vowels than others (see Section 4.5.3.). Hence, another issue that could be addressed in a follow-up study is the effect of phonetic context on the degree of convergence. Degree of convergence could be operationalised as the difference in mean values between the baseline and each of the imitation/accommodation conditions. Alternatively, it could be operationalised as the number of participants who converged (or diverged) on vowel duration with respect to a given vowel.

Apart from the three temporal parameters analysed in the current study, another phonetic feature that could be examined using the collected data is vowel quality. The stimuli contained the front vowels TRAP, DRESS, KIT and FLEECE, which were selected on the grounds that maintaining TRAP-DRESS and KIT-FLEECE contrasts has often been found to be problematic for Polish learners of English (e.g. Sobkowiak, 2001; Gonet, Szpyra-

Kozłowska and Święciński, 2010; Nowacka, 2010; Weckwerth, 2011; see Section 3.6.2.). The two vowel contrasts were assimilated in the stimuli provided by the Polish speaker (TRAP and DRESS were both realised as Polish /e/, while KIT and FLEECE were both substituted with Polish /i/). Thus, it could prove worthwhile to inspect the subjects' realisation of the four vowel qualities following exposure to native and Polish-accented productions. A re-examination of the data obtained in one of the pilot studies (Zając, forthcoming) indicated that exposure to modelled speech caused some subjects to modify the spectral characteristics of their vowels. The results revealed considerable variability in terms of imitation strategies and suggested that more participants converged towards the native Polish speaker than towards the native English speaker, especially in the case of the KIT/FLEECE contrast.

As far is further research is concerned, it would be interesting to examine pronunciation shifts in the speech of learners *not* majoring in English. The author's former students were selected as participants since they were more easily accessible, however, it should be recognised that due to their educational background, such learners form a rather idiosyncratic subject group. It is possible that participants who had not received phonetic training and are not aiming to become language experts would exhibit less negative attitudes towards foreign-accented speech and use different convergence strategies as a consequence.

It has been repeatedly argued in this chapter that the subjects' speech behaviour may have been to some extent conditioned by whether or not a given pronunciation feature was placed high in their consciousness. In order to verify this claim, one could use self-report measures and ask the participants (in a form of a written questionnaire or a recorded interview) whether they are aware of the existence of a particular pronunciation feature and whether they had adjusted it under any of the experimental conditions. However, to obtain more objective and easily quantifiable data, it might be advisable to use some indirect method of measurement. For instance, one could ask the informants to listen to short extracts of native speech that include the investigated phonetic variables and instruct them to identify the features that are in their view characteristic of native English pronunciation. The phonetic features which are noticed first would presumably represent those that are placed the highest in the learners' consciousness. In a similar vein, one could have the subjects listen to extracts of non-native speech that include L1-accented realisations of the phonetic variables under investigation and ask them to enumerate mispronunciations made by the speaker.

The findings of the study suggest that the stage of acquisition of a given pronunciation feature may affect learners' convergence strategies. To validate this claim, more detailed investigation of the phonetic variables could be performed. As referred to in Section 4.3.,

different types of elicitation tasks could be used (task type and thedegree of formality may affect the proportion of native-like realisations in the learners' speech, see Tarone, 1979, 1982). Productions provided by a representative, native-speaker reference group and the subjects' realisations of equivalent sounds in their L1 could also be included in the analysis. Another solution could be to examine participants' perceptual discrimination of the analysed phonetic variables. According to Flege's Speech Learning Model (Flege, 1995), foreign accent in L2 speech is generated by inaccurate perceptual patterns. More specifically, it is argued that the process of equivalence classification (processing a given L2 sound and its L1 equivalent as belonging to the same phonetic category) may prevent successful acquisition and production of TL sounds (See Section 2.5.). Thus, it could be hypothesised that a given L2 sound has not been fully acquired if a learner is unable to distinguish it perceptually from its L1 counterpart (e.g. aspirated /p t k/ in English and unaspirated /p t k/ in Polish).

The results of the study imply that the participants viewed the Polish speaker's pronunciation in a negative light, which may have stemmed from the marked deviations from L1 and L2 pronunciation norms in his productions (see Section 4.2.). Since the subjects' attitude towards the non-native speaker's pronunciation was likely one of the reasons behind divergence from L1-influenced values, it might prove worthwhile to examine convergence strategies towards a speaker with a milder foreign-accent in follow-up projects. Learners might be more inclined to imitate or accommodate towards non-native speech if the pronunciation of the model talker/interlocutor is not as strongly accented. At the same time, if the values provided by the non-native speaker are comparable to those exhibited by the participants, pronunciation shifts may be difficult to detect. A possible solution would be to manipulate a smaller number of phonetic variables than in the current study, e.g. use stimuli that contain unaspirated stops but are otherwise "mildly-accented".

Another suggestion for further research would be to correlate quantitative attitudinal data with results obtained with the use of acoustic measures. Although questionnaire findings in the current study included some quantitative data, this type of analysis was not undertaken as the purpose of the questionnaire was to verify the assumption that students of English Studies will favour native over foreign-accented English rather than to provide a comprehensive account of the subjects' attitudes. Also, questionnaire results were found to be relatively uniform and correlating them with the acoustic data did not seem necessary. Nonetheless, it could be worthwhile to carry out a more detailed investigation of attitudinal factors (e.g. degree of bias towards/against the target-language and the native-language groups) and their impact on the magnitude of convergence in L2 speech. This type of research was conducted

by Babel (2009, 2010), who found a correlation between implicit attitudes towards a given social group and the extent and direction of convergence towards the pronunciation of the group's representatives. Measuring implicit rather than explicit attitudes could also yield interesting results. As explained by Pantos and Perkins (2013: 5), "*implicit* attitudes [can be defined] as a person's automatic, immediate reactions to an attitude object based on preexisting stereotypes and cognitive connections, and *explicit* attitudes as reactions formed through additional controlled cognitive processing." Explicit attitudes are examined using self-report measures, which have been criticised by social psychologists for their susceptibility to factors such as demand characteristics (which refer to a situation when participants form an interpretation of the purpose of the experiment and unconsciously change their behaviour to fit that interpretation; Orne, 1962) or evaluation apprehension (a desire on the part of the subjects to gain the experimenter's approval; Rosenberg, 1969) (Greenwald et al., 2002). Implicit attitudes, on the other hand, are inspected with the use of indirect measures (which are said to be more objective than self-reflective methods), such as the Implicit Association Task (IAT, Greenwald, McGhee and Schwartz, 1998). The IAT uses reaction times in association of two target concepts (e.g. Poland vs. England) with an evaluation attribute (e.g. pleasant vs. unpleasant words) to gauge implicit biases towards the two target concepts.

Conclusions

The aim of the study was to examine L2 convergence strategies upon exposure to native and non-native pronunciation. The term convergence strategies was used to refer to three types of linguistic behaviour: convergence (making one's speech more similar to that of another person), divergence (moving away from the speech of another person) and maintenance (maintaining one's default linguistic behaviour in spite of exposure to the speech of another person). The study was grounded in two bodies of research: L2 phonetics and work on speech convergence. As discussed in Chapter One, previous findings indicate that speakers tend to adjust their speech when listening to or interacting with others. The process seems to have its origin in a natural human predisposition to imitate actions performed by other individuals. As indicated by the results of previous studies, the natural tendency to imitate ambient speech can be sustained (convergence), blocked (maintenance) or reversed (divergence) depending on a variety of social-psychological and linguistic factors. Previous research on L2 phonetics (as discussed in Chapter Two) shows that speakers' productions in a foreign language are generated by an autonomous linguistic system (interlanguage), which contains elements of the learner's L1 and L2 but does not correspond exactly to either NL or TL. Interlanguage appears to be a dynamic system that restructures itself as the learner gains more experience with the L2. It has been found that the development of IL, and by extension the development of the L2 sound system, is dependent upon a range of different social-psychological, psycholinguistic and linguistic factors. The issue of speech convergence and the examination of L2 phonetics have been merged in a number of previous studies on phonetic convergence in non-native pronunciation; their results suggest that L2 learners may tend to adjust their speech when listening to or interacting with others and that the magnitude of the process may be affected by attitudinal and linguistic factors (as discussed in Section 3.2.).

The study was concerned with the speech behaviour of advanced Polish learners of English, who were exposed to two pronunciation varieties: Polish-accented English and native English. The participants were 38 second-year students of English Studies, recruited from the University of Lodz. A new experimental procedure was developed on the basis of pilot work. Convergence strategies upon exposure to native and non-native English were analysed under two conditions: imitation and accommodation. In the imitation condition, the subjects repeated target words produced by two model talkers immediately after hearing them. In the accommodation condition, the subjects read target words for two interlocutors to listen to at a

later time. Convergence strategies were examined by contrasting productions from the imitation and accommodation conditions with realisations collected before the subjects' were exposed to modelled speech (baseline condition). In the final phase of the experiment, the participants were required to complete a questionnaire that was designed to measure their attitudes towards native and foreign-accented English. The model talkers/interlocutors were a native speaker of Standard Southern British English and a Polish speaker of English who imitated a heavy Polish accent for the purposes of the study. The phonetic variables under investigation were the following: aspiration (operationalised as voicing lag values in initial /p k/), pre-voicing in word-initial stops (operationalised as voicing lead values in initial /b g/) and vowel duration as a cue for consonant voicing (operationalised as the difference in duration between vowels followed word-final /d/ and the same vowels followed by word-final /t/).

Three hypotheses were formulated to be tested in the course of the study; they were based on previous findings that the phenomenon of speech convergence may take place in L2 pronunciation and can be affected by social-psychological and linguistic factors. Hypothesis 1 related to the overall effect of exposure to two English varieties (native vs. non-native) on subjects' convergence strategies; it predicted that convergence strategies following exposure to native and non-native English would vary as a function of model talker/interlocutor. Hypothesis 2 and Hypothesis 3 were concerned with specific social-psychological and linguistic variables that could potentially influence convergence strategies. Hypothesis 2 predicted that convergence strategies following exposure to native and non-native English would be affected by the subjects' attitudes towards native and Polish-accented English. Hypothesis 3 predicted that convergence strategies following exposure to native and non-native English would differ as a function of phonetic context (place of articulation and vowel category).

The findings of the study show that the subjects modified their speech behaviour following exposure to the speech of the model talkers/interlocutors, which implies that L2 convergence phenomena are present in L2 pronunciation. Importantly, all three convergence strategies were observed in the subjects' speech behaviour: convergence (operationalised as a significant shift towards the values exhibited by a given model talker/interlocutor), divergence (operationalised as a significant shift away from the values exhibited by a given model talker/interlocutor) and maintenance (operationalised as a non-significant difference between the subjects' default realisations and the values exhibited following exposure to the speech of a given model talker/interlocutor). Convergence towards native English was found in the case

of aspiration and vowel duration when the participants were repeating the target words in the imitation condition and when they were reading the words in the accommodation condition. Convergence towards Polish-accented English was found with respect to vowel duration when the subjects were producing the target words in the imitation condition. Divergence from native English was not found in the data obtained for the study, however, the results imply that the informants diverged from Polish-accented English in the case of aspiration in the imitation condition. Maintenance with respect to native English was observed in the case of pre-voicing (both in the imitation and in the accommodation condition); maintenance with respect to foreign-accented speech was observed in the accommodation condition in the case of all three phonetic parameters. Overall, the findings indicate that convergence strategies varied as a function of model talker/interlocutor, thus confirming Hypothesis 1. The discrepancy in convergence strategies towards foreign-accented and native English was found to be related to attitudinal factors. The results suggest that a bias in favour of target-like pronunciation may prompt learners to converge towards native speech and diverge from foreign-accented speech. Nonetheless, it should be pointed out that some instances of convergence towards Polish-accented English were observed despite a strong preference for native English on the part of the subjects. This was interpreted to mean that convergence strategies depended on the phonetic parameter under investigation (aspiration, pre-voicing, vowel duration) and, more specifically, on the stage of acquisition of a given pronunciation feature (it was observed that the three phonetic variables reflected different stages of acquisition in the participants' IL: early in the case of pre-voicing, intermediate in the case of vowel length and late in the case of aspiration). Thus, the findings suggest that learners who exhibit a strong preference for native-like pronunciation will tend to converge towards the TL norm on L2 phonetic features that they had successfully acquired. If, on the other hand, a given pronunciation feature is in an early stage of acquisition in the learner's IL, convergence towards L1-accented values appears to be more likely. These findings lend partial support for Hypothesis 2, i.e. the results suggest that attitudinal factors may affect the magnitude of convergence in L2 speech, however, their influence appears to be mediated by the stage of acquisition of a given TL pronunciation feature. Hypothesis 3 was not supported by the data obtained in the study. Although previous research (e.g. Rojczyk et al., 2013; Zając, 2013) suggests that convergence strategies may differ as a function of phonetic context, a similar effect was not found in the current study. Taken together, the findings corroborate the claim that the phenomenon of speech convergence operates in L2 speech and imply that learners'

convergence strategies may be affected by certain social-psychological and psycholinguistic factors.

References

Adamson, H. D. and Regan, V. M. (1991) The acquisition of community speech norms by Asian immigrants learning English as a second language. *Studies in Second Language Acquisition* 13, 1-22.

Adjemian, C. (1976) On the nature of interlanguage systems. *Language Learning* 26, 297-320.

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. and Weinert, R. (1991) The HCRC Map Task corpus. *Language and Speech* 34, 351–366.

Anisman, P. H. (1975) Some aspects of code switching in New York Puerto Rican English. *Bilingual Review* 2, 56-85.

Asher, J. J. and García, R. (1969) The optimal age to learn a second language. *The Modern Language Journal* 53, 334-341.

Babel, M. (2009) Phonetic and social selectivity in speech accommodation. Unpublished PhD dissertation. University of California, Berkeley.

Babel, M. (2010) Dialect divergence and convergence in New Zealand English. *Language in Society* 39, 437–456.

Babel, M. and Bulatov, D. (2012) The role of fundamental frequency in phonetic accommodation. *Language and Speech* 55, 231-248.

Babel, M., McGuire, G., Nicholls, A. and Walters, S. (2012) Variability in imitation based on voice profile. Presentation at the 2nd Annual Workshop on Sound Change. Seeon-Seebruck, Germany.

Beebe, L. M. (1977) The influence of the listener on code-switching. *Language Learning* 27, 331-339.

Beebe, L. M. (1980) Sociolinguistic variation and style shifting in second language acquisition. *Language Learning* 30, 433-448.

Beebe, L. M. and Zuengler, J. (1985) Accommodation theory: An explanation for style shifting in second language dialects. In N. Wolfson and E. Judd (eds.), *Sociolinguistics and language acquisition*. Rowley, MA: Newbury House Publishers, Inc., 195-213.

Bell, A. (1977) The language of radio news in Auckland. A sociolinguistic study of style, audience, and subediting variation. PhD dissertation. University of Auckland.

Bell, A. (1982a) Radio: The style of news language. *Journal of Communication* 32, 150-164.

Bell, A. (1982b) This isn't the BBC: Colonialism in New Zealand English. *Applied Linguistics* 3, 246-258.

Bell, A. (1984) Language style as audience design. *Language in Society* 13, 145-204.

Berkowitz, D. (1986) The effects of perceived cultural empathy on second language performance. Unpublished PhD dissertation. Columbia University.

Best, C. T. (1995) A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press, 171–204.

Best, C. T., and Tyler, M. D. (2007) Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn and M. J. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins, 13–34.

Bilous, F. and Krauss, R. (1988) Dominance and accommodation in the conversational behaviours of same- or mixed-gender dyads. *Language & Communication* 8, 183-194.

Bock, J.K. (1986) Syntactic persistence in sentence production. *Cognitive Psychology* 18, 355-387.

Boersma, P. (2001) Praat, a system for doing phonetics by computer. *Glot International* 10, 341-345.

Bohn, O.-S. and Flege, J. E. (1992) The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14, 131-158.

Bourhis, R. (1984) Cross-cultural communication in Montreal: Two field studies since Bill 101. *International Journal of the Sociology of Language* 46, 33-47.

Bourhis, R. and Giles, H. (1977) The language of intergroup distinctiveness. In H. Giles (ed.), *Language, ethnicity, and intergroup relations*. Waltham: Academic Press, 119–136.

Branigan, H. P., Pickering, M. J. and Cleland, A. A. (2000) Syntactic co-ordination in dialogue. *Cognition* 75, B13-B25.

Brouwer, S., Mitterer, H. and Huettig, F. (2010) Shadowing reduced speech and alignment. *The Journal of the Acoustical Society of America* 128, EL32-EL37.

Byrne, D. (1971) *The attraction paradigm*. New York: Academic Press.

Caramazza, A., Yeni-Komshian, G., Zurif, E., and Carbone, E. (1973) The acquisition of a new phonological contrast: The case of stop consonants. *The Journal of the Acoustical Society of America* 54, 421-428.

Chartrand, T. and Bargh, J. (1999) The Chameleon Effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology* 76, 893-910.

Chen, M. (1970) Vowel length variation as a function of the voicing of consonant environment. *Phonetica* 22, 129-159.

Chiba, R., Matsuura, H. and Yamamoto, A. (1995) Japanese attitudes towards English accents. *World Englishes* 14, 77-86.

Cho, T. and Ladefoged, P. (1999) Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27, 207-229.

Cole, J., Kim, H., Choi, H. and Hasegawa-Johnson, M. (2007) Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics* 35, 180–209.

Couper, G. (2011) What makes pronunciation teaching work? Testing for the effect of two variables: Socially constructed metalanguage and critical listening. *Language Awareness* 20, 159–182.

Coupland, N. (1984) Accommodation at work: Some phonological data and their implications. *International Journal of the Sociology of Language* 46, 49-70.

Dalton-Puffer, C., Kaltenboeck, G. and Smit, U. (1997) Learner attitudes and L2 pronunciation in Austria. *World Englishes* 16, 115-128.

Delvaux, V. and Soquet, A. (2007) The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64, 145-73.

Derwing, T. M., Munro, M. J. and Wiebe, G. (1998) Evidence in favor of a broad framework for pronunciation instruction. *Language Learning* 48, 393–410.

Dickerson, L. (1975) The learner's interlanguage as a system of variable rules. T*ESOL Quarterly* 9, 401-407.

Dickerson, L. and Dickerson, W. (1977) Interlanguage phonology: current research and future directions. In S. P. Corder and E. Roulet (eds.), *The notions of simplification. Interlanguages and pidgins: Actes du Seme Colloque de Linguistique Applique de Neufchatel*, 18-30.

Dowd, J., Zuengler, J. and Berkowitz, D. (1990) L2 social marking: Research issues. *Applied Linguistics* 11, 16-29.

Douglas-Cowie, E. (1978) Linguistic code-switching in a Northern Irish village: Social interaction and social ambition. In P. Trudgill (ed.), *Sociolinguistic patterns in British English*. London: Edward Arnold, 37-51.

Dziubalska-Kołaczyk, K. (1990) *A theory of Second Language Acquisition within the framework of Natural Phonology. A Polish-English contrastive study*. Poznań: AMU Press.

Dziubalska-Kołaczyk, K. (2002) Conscious competence of performance as a key to teaching English. In E. Waniek-Klimczak and P.J. Melia (eds.), *Accents and Speech in English*. Frankfurt am Main: Peter Lang, 97-106.

Elliot, A. R. (1995) Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *The Modern Language Journal* 79, 369-371.

Elsendoorn, B. (1985) Production and perception of Dutch foreign vowel duration in English monosyllabic words. *Language and Speech* 28, 231-254.

Flege, J. E. (1980) Phonetic approximation in second language acquisition. *Language Learning* 30, 117-134.

Flege, J. E. (1981) The phonological basis of foreign accent. *TESOL Quarterly* 15, 443-455.

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

Flege, J. E. (1988) Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America* 84, 70-79.

Flege, J. E. (1991) Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America* 89, 395–411.

Flege, J. E. (1995) Second Language Speech Learning. Theory, findings and problems. In W. Strange (ed), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press, 233-277.

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

Flege, J. E. and Eefting, W. (1988) Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America* 83, 729-740.

Flege, J. E. and Liu, S. (2001) The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23, 527-552.

Flege, J. E., Murray, M. J. and Skelton, L. (1992) Production of the word-final English/t/-/d/contrast by native speakers of English, Mandarin, and Spanish. *The Journal of the Acoustical Society of America* 92, 128-143.

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

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

Fries, C. C. (1945) *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press.

Gallois, C., Giles, H., Jones, E., Cargile, A. C. and Ota, H. (1995) Accommodating intercultural encounters: Elaborations and extensions. In R. Wiseman (ed.), *Intercultural communication theory*. Thousand Oaks: Sage, 115-147.

Gardner, R. C. (1985) *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C. and Lambert, W. E. (1972) *Attitudes and motivation in second-language learning*. Rowley, Mass.: Newbury House Publishers, Inc.

Gatbonton, E. (1975) Systematic variations in second language speech: A sociolinguistic study. Unpublished PhD dissertation. McGill University, Montreal.

Gatbonton, E. (1978) Patterned phonetic variability in second-language speech: A gradual diffusion model. *Canadian Modern Language Review* 34, 335-347.

Gatbonton, E., Trofimovich, P. and Segalowitz, N. (2011) Ethnic group affiliation and patterns of development of a phonological variable. *The Modern Language Journal* 95, 188-204.

Garrett, P. (2007) *Attitudes to language. Key topics in sociolinguistics*. Cambridge: Cambridge University Press.

Giles, H. (1973) Accent mobility: a model and some data. *Anthropological Linguistics* 15, 87-105.

Giles, H., Coupland, N. and Coupland, J. (1991) *Accommodation Theory: Communication, context and consequence*. In J. Coupland, N. Coupland, and H. Giles (eds.), *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press, 1-68.

Giles, H. and Ogay, T. (2007) Communciation Accommodation Theory. In W. Samter and B.B. Whaley (eds.), *Explaining communciation. Contemporary theories and exemplars*. New Jersey: Lawrence Erlbaum Associates, Inc., 325-244.

Giles, H. and Sachdev I. (2008) Bilingual accommodation. In T. K. Bhatia (ed.), *The handbook of bilingualism*. Oxford: Blackwell Publishing, 353-378.

Giles, H. Scherer, K. R. and Taylor, D. M. (1979) Speech markers in social interaction. In K. R. Scherer and H. Giles (eds.), *Social markers in speech*. Cambridge: Cambridge University Press, 343-381.

Giles, H., Taylor, D. M. and Bourhis, R. (1973) Towards a theory of interpersonal accommodation through language: some Canadian data. *Language in Society* 2, 177-192.

Goldinger, S. (1998) Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251-279.

Goldinger, S. and Azuma, T. (2004) Episodic memory in printed word naming. *Psychonomic Bulletin & Review* 11, 716–722.

Gonet, W., Szpyra-Kozłowska, J. and Święciński, R. (2010) Clashes with ashes. In E. Waniek-Klimczak (ed.), *Issues in accents of English 2: Variability and norm*. Newcastle upon Tyne: Cambridge Scholars Publishing, 213-230.

Gralińska-Brawata, A. (2013) The dynamics of linguistic rhythm in Polish learners of English. Unpublished PhD dissertation. University of Lodz.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., and Mellott, D. S. (2002) A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review* 109, 3-25.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998) Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74, 1464-1480.

Gregory, S. and Hoyt, B. (1982) Conversation partner mutual adaptation as demonstrated by fourier series analysis. *Journal of Psycholinguistic Research* 11, 35-46.

Gregory, S. and Webster, S. (1996) A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology* 70, 1231-1240.

Guiora, A. Z. (1972) Construct validity and transpositional research: Toward an empirical study of psychoanalytic concepts. *Comprehensive Psychiatry* 13, 139-150.

Guiora, A. Z., Beit-Hallahmi, B., Brannon, R. C. L., Dull, C. Y. and Scovel, T. (1972) The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry* 13, 421-428.

Guiora, A. Z., Acton, W. R., Erard, R. and Strickland, F. W. (1980) The effects of benzodiazepine (valium) on permeability of language ego boundaries. *Language learning* 30, 351-361.

Hannah, A. and Murachver, T. (1999) Gender and Conversational Style as Predictors of Conversational Behavior. *Journal of Language and Social Psychology* 18, 153-174.

Hartford, B. (1978) Phonological differences in the English of adolescent female and male Mexican-Americans. *International Journal of the Sociology of Language* 17, 55-64.

Heiman, G. W. (2002) *Research methods in psychology.* Boston: Houghton Mifflin Company.

Hogan, J. T. and Rozsypal, A. J. (1980) Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America* 67, 1764-1771.

Honorof, D., Weihing, J. and Fowler, C. (2011) Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39, 18-38.

Iverson, P., Kuhl, P., Yamada, R., Diesch, E., Tohkura, Y., Ketterman, A. and Siebert, C. (2001) A perceptual interference account of acquisition difficulties for non-native phonemes. *Speech, Hearing and Language: Work in Progress* 13, 106-118.

Janicka, K., Kul, M., and Weckwerth, J. (2005) Polish students' attitudes to native English accents. In K. Dziubalska-Kołaczyk, and J. Przedlacka (eds.), *English pronunciation models: A changing scene.* Bern: Peter Lang, 251-292.

Jassem, W. and Richter, L. (1989) Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17, 205-212.

Kawanami, S. and Kawanami, K. (2010) Japanese learners' perception and attitudes towards English accents. *Let Kyushu-Okinawa Bulletin* 10, 1-14.

Keating, P. A. (1980) A phonetic study of voicing contrast in Polish. Unpublished PhD dissertation. Brown University.

Keating, P. A. (1984) Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286-319.

Keating, P. A., Mikoś, M. J., and Ganong III, W. F. (1981) A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America* 70, 1261-1271.

Kennedy, S. and Trofimovich, P. (2010) Language awareness and second language pronunciation: A classroom study. *Language Awareness* 19, 171–185.

Kennedy, S., Blanchet, J. and Trofimovich, P. (2014) Learner pronunciation, awareness, and instruction in French as a second language. *Foreign Language Annals* 47, 79–96.

Kim, M. (2011) Phonetic convergence after perceptual exposure to native and non-native speech: preliminary findings based on fine-grained acoustic-phonetic measurement. *Proceedings of the 17th International Congress of Phonetic Sciences*,1074-1077.

Kim, M., Horton, W. S., Bradlow, A. R. (2011) Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Journal of Laboratory Phonology* 2, 125-156.

Klatt, D. D. (1975) Voice Onset Time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research* 18, 686-706.

Kopczyński, A. (1977) *Polish and American English consonant phonemes: A contrastive study*. Warszawa: Państwowe Wydawnictwo Naukowe.

Kuhl, P. K. and Iverson, P. (1995) Linguistic experience and the "Perceptual Magnet Effect". In W. Strange (ed), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press, 121–154.

Kuhl, P. K. and Meltzoff, A. (1996) Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America* 100, 2425-38.

Labov, W. (1966) *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. (1969) The study of language in its social context. *Studim Generale* 23, 30-87.

Lado, R. (1957) *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Lenneberg, E. H. (1967) *Biological foundations of language*. New York: Wiley.

Lewandowski, N. (2012) Talent in nonnative phonetic convergence. Unpublished PhD dissertation. University of Stuttgart.

Levelt, W. J. M. and Kelter, S. (1982) Surface form and memory in question answering. *Cognitive Psychology* 14, 78-106.

Lisker, L. (1974) On "explaining" vowel duration variation. *Glossa* 8, 233-246.

Lisker, L. and Abramson, A. S. (1964) A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 527-565.

Llamas, C., Watt, D. and Johnson, D. E. (2009) Linguistic accommodation and the salience of national identity markers in a border town. *Journal of Language and Social Psychology* 28, 381-407.

Mack, M. (1982) Voicing dependent vowel duration in English and French: Monolingual and bilingual production. *The Journal of the Acoustical Society of America* 71, 173-178.

Major, R. C. (1987) A model for interlanguage phonology. In G. Ioup and S. Weinberger (eds.), *Interlanguage Phonology*. Rowley: Newbury House Publishers, 101-124.

Major, R. C. (2001) *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Major, R. C. (2008) Transfer in second language phonology. In J.G. Hansen-Edwards and M. L. Zampini (eds), *Phonology and second language acquisition*. Amsterdam: John Benjamins, 63-94.

Marlowe, D. and Crowne, D, (1961) Social desirability and responses to perceived situational demands. *Journal of Consulting and Clinical Psychology* 25, 109-115.

Matsuura, H., Chiba, R. and Yamamoto, A. (1994) Japanese college students' attitudes towards non-native varieties of English. In D. Graddol and J. Swann (eds.), *Evaluating Language*. Clevedon: BAAL/Multilingual Matters, 52-61.

Miller, R. M, Sanchez, K. and Rosenblum, L. D. (2010) Alignment to visual speech information. *Attention, Perception, & Psychophysics* 72, 1614-1625

Mitterer, H. and Ernestus, M. (2008) The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109, 168-173.

Namy, L, Nygaard, L. and Sauerteig, D. (2002) Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21, 422-432.

Natale, M. (1975a) Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology* 32, 790-804.

Natale, M. (1975b) Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills* 40, 827–830.

Neumann, R. and Strack, F. (2000) Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology* 79, 211-223.

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

Nowacka, M. (2008) Phonetic attainment in Polish university and college students of English: a study in the productive and receptive pronunciation skills. Unpublished PhD dissertation. Maria Curie-Skłodowska University, Lublin.

Nowacka, M. (2010) The ultimate attainment of English pronunciation by Polish college Students: a longitudinal study. In E. Waniek-Klimczak (ed.), *Issues in Accents of English 2. Variability and Norm*. Newcastle upon Tyne: Cambridge Scholars Publishing, 233-260.

Olson, L. L. and Samuels, J. S. (1973) The relationship between age and accuracy of foreign language pronunciation. *The Journal of Educational Research* 66, 263-268.

Oostdijk, N. (2000) The spoken dutch corpus project. *The ELRA Newsletter* 5, 4-8.

Orne, M. T. (1962) On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist* 17, 776–783.

Ostaszewska, D. and Tambor, J. (2000) *Fonetyka i fonologia współczesnego języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.

Oyama, S. (1976) A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 3, 261-283.

Pantos, A. J. and Perkins, A. W. (2013) Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology* 32, 3-20.

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

Pardo, J. (2010) Expressing oneself in conversational interaction. In E. Morsella (ed.), *Expressing oneself/expressing one's self. Communication, cognition, language, and identity*. New York: Psychology Press, 183-196.

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

Pardo, J., Cajori Jay, I., Hoshino, R., Hasbun, S., Sowemimo-Coker, C. and Krauss, R. (2013) The influence of role-switching on phonetic convergence in conversation. *Discourse Processes* 50, 276-300.

Peterson, G. E. and Lehiste, I. (1960) Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America* 32, 693-703.

Pickering, M. J. and Garrod, S. (2004) Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27, 169-226.

Piper, T. and Cansin, D. (1988) Factors influencing the foreign accent. The *Canadian Modern Language Review* 44, 334-342.

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

Piotrowski, M. (2013) Voice Onset Time in Polish students of English - A sociolinguistic approach. Unpublished MA dissertation. University of Lodz, Poland.

Porzuczek, A. (2012) *The timing of tone group constituents in the advanced Polish learner's English pronunciation*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Purcell, E. T. and Suter, R. W. (1980) Predictors of pronunciation accuracy: a re-examination. *Language Learning* 30, 271-287.

Ramírez Verdugo, D. (2006) A study of intonation awareness and learning in non-native speakers of English. *Language Awareness* 15, 141–159.

Raphael, L. J. (1972) Preceding vowel duration as a cue to the perception of voicing of word-final consonants in American English. *Journal of the Acoustical Society of America* 51, 1296-1303.

Rochet, B. (1995) Perception and production of L2 speech sounds by adults. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Timonium, MD: York Press, 379-410.

Rojczyk, A. (2008) Cross-linguistic priming on vowel duration and delayed plosion in Polish-English bilinguals. In E. Waniek-Klimczak (ed.), *Issues in Accents of English*. Newcastle: Cambridge Scholars Publishing, 44-63.

Rojczyk, A. (2010) *Temporal and spectral parameters in perception of the voicing contrast in English and Polish*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Rojczyk, A. (2012) Phonetic and phonological mode in second-language speech: VOT imitation. Paper presented at EuroSLA22 - 22nd Annual Conference of the European Second Language Association, Poznań, Poland, 5-8 September.

Rojczyk, A. (2013) Phonetic imitation of L2 vowels in a rapid shadowing task. In J. Levis and K. LeVelle (eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*. Ames, IA: Iowa State University, 66-76.

Rojczyk, A., Porzuczek, A. and Bergier, M. (2013) Immediate and distracted imitation in second-language speech: Unreleased plosives in English. *Research in Language* 11, 3-18.

Rosenberg, M. J. (1969) The conditions and consequences of evaluation apprehension. In R. Rosenthal and R. L. Rosnow (eds.), *Artifact in behavioral research.* New York: Academic Press, 279–349.

Sato, C. (1985) Task variation in interlanguage phonology. In S. M. Gass and C. G. Madden (eds.), *Input in Second Language Acquistion*. Rowley, MA: Newbury House.

Sawyer, J. (1973) Social aspects of bilingualism in San Antonio, Texas. In R. W. Bailey and J. L. Robinson (eds.), *Varieties of present-day English*. New York: Macmillan Publishing Co., Inc.

Scales, J. Wennerstrom, A., Richard, D. and Hui Wu, S. (2006) Language learners' perceptions of accent. *TESOL Quarterly* 40, 715-738.

Schmidt, R. (1987) Sociolinguistic variation and language transfer in phonology. In G. Ioup and S. H. Weinberger (eds.), *Interlanguage phonology: The acquisition of a second language sound system*. New York: Newbury House, 365-377.

Schmidt, A. M. and Flege, J. E. (1996) Speaking rate effects on stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica* 53, 162–179.

Schumann, J. (1978) The acculturation model for second-language acquisition. In R. C. Gingras (ed.), *Second language acquisition and foreign language acquisition*. Washington, DC: Centre for Applied Linguistics, 27-50.

Schumann, J. (1986) Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development* 7, 379-392.

Schweitzer, A. and Lewandowski, N. (2012) Accommodation of backchannels in spontaneous speech. In the booklet of the International Symposium on Imitation and Convergence in Speech.

Scovel, T. (1969) Foreign accents, language acquisition, and cerebral dominance. *Language Learning* 19, 245-253.

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

Selinker, L. (1992) *Rediscovering interlanguage*. London: Longman.

Shockey, L. R. (2013) English vowels: The long and short of it. In E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and researching English accents in native and non-native speakers.* Heidelberg: Springer, 139-159.

Shockley, K., Sabadini, L. and Fowler, C. A. (2004) Imitation in shadowing words. *Perception & Psychophysics* 66, 422-429.

Slowiaczek, L. M. and Dinnsen, D. (1985) On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13, 325-341.

Slowiaczek, L. M. and Szymanska, H. J. (1989) Perception of word-final devoicing in Polish. *Journal of Phonetics* 17, 205-212.

Snow, C. E. and Hoefnagel-Höhle, M. (1977) Age differences in the pronunciation of foreign sounds. *Language and Speech* 20, 357-365.

Sobkowiak, W. (2001) *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Strange, W. (2001) Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics* 39, 456-466.

Summerfield, Q. (1975) How a full account of segmental perception depends on prosody and vice versa. In A. Cohen and S. G. Nooteboom (eds.), *Structure and Process in Speech Perception*. Berlin: Springer-Verlag, 51-68.

Suter, R. W. (1976) Predictors of pronunciation accuracy in second language learning. *Language Learning* 26, 233-253.

Szpyra-Kozłowska, J., Frankiewicz, J. and Gonet, W. (2002) Aspekty fonetyki angielskiej nauczane w polskich szkołach średnich. In W. Sobkowiak and E. Waniek-Klimczak (eds.), *Dydaktyka fonetyki języka obcego I. Konferencja w Soczewce, 11-13.5.2001*. Płock: Wydawnictwo PWSZ, 9-27.

Tahta, S., Wood, M. and Loewenthal, K. (1981) Foreign accents: factors relating to transfer of accent from the first language to a second language. *Language & Speech* 24, 265-272.

Tajfel, H. (1974) Social identity and intergroup behaviour. *Social Science Information* 13, 65-93.

Tajfel, H. (1981) *Human groups and social categories: studies in social psychology*. Cambridge: Cambridge University Press.

Tajfel, H. (1982) Social psychology of intergroup relations. *Annual Review of Psychology* 33, 1-39.

Tajfel, H. and Turner, J. (1979) An integrative theory of intergroup conflict. In Austin, W. and Worchel, S. (eds.) *The social psychology of intergroup relations*. Monterey, CA: Brooks-Cole, 35-53.

Tarone, E. (1979). Interlanguage as chameleon. *Language Learning* 29, 181-191.

Tarone, E. (1982) Systematicity and attention in interlanguage. *Language Learning* 32, 69-84.

Tarone, E. (1994) Interlanguage. In R.E. Asher (ed.), *The encyclopedia of language and linguistics.* Oxford: Pergamon, 1715–1719.

Taylor, L. L., Catford, J.C., Guiora, A. Z. and Lane, H. L. (1971) Psychological variables and ability to pronounce a second language. *Language and Speech* 14, 146-157.

Thompson, R. M. (1975) Mexican-American English: Social correlates of regional pronunciation. *American Speech* 50, 18-24.

Thompson, I. (1991) Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning* 41, 177-204.

Timmins, I. (2002) Native-speaker norms and International English: a classroom view. *ELT Journal* 56, 240-249.

Trofimovich, P. and Kennedy, S. (2014) Interactive alignment between bilingual interlocutors: Evidence from two information-exchange tasks. *Bilingualism: Language and Cognition* 17, 822-836.

Trofimovich, P., McDonough, K., and Foote, J. (2014) Interactive alignment of multisyllabic stress patterns in a second language classroom. *TESOL Quarterly 48*, 815-832.

Trudgill, P. (1981) Linguistic accommodation: Sociolinguistic observations on a sociopsychological theory. In C. Masek, R. Hendrick and M. Miller (eds.), *Papers from the parasession on language and behavior. Chicago Linguistic Society 1981*.Chicago: University of Chicago Press, 218-37.

Turner, J. (1975) Social comparison and social identity: Some prospects for intergroup behaviour. *European Journal of Social Psychology* 5, 5-34.

Van Santen, J. P. H. (1992) Contextual effects on vowel duration. *Speech Communication* 11, 513–546.

Waniek-Klimczak (1999) Sources of variability in L2 phonology: vowel duration in 'Polish English'. In M. C, Gruber, D. Higgins, K. Olson and T. Wysocki (eds.), *CLS 34 - The Panels Vol. 34-2: The Status of Constraints; The Acquisition of Spoken Language; Acquisition and the Lexicon*. Chicago: Chicago Linguistic Society, 387-401.

Waniek-Klimczak (2002) How to predict the unpredictable – English word stress from a Polish perspective. In E. Waniek-Kimczak and P. J. Melia (eds.), *Accents and speech in teaching English phonetics and phonology*. Frankfurt: Peter Lang, 221–241.

Waniek-Klimczak, E. (2005) *Temporal parameters in second language speech: An applied linguistic phonetic approach*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Waniek-Klimczak, E. (2009a) Accommodation in interlanguage speech: a pilot case study. In J. Arabski and A. Wojtaszek (eds.), *Multidisciplinary perspectives on Second Language Acquisition and Foreign Language Learning*. Katowice: Oficyna Wydawnicza, 198-207.

Waniek-Klimczak (2009b) *Socio-psychological conditioning in ESL pronunciation: Consonant voicing in English spoken by Polish immigrants to Britain*. Wydawnictwo Państwowej Wyższej Szkoły Zawodowej we Włocławku: Włocławek.

Waniek-Klimczak, E. (2011) Aspiration and Style: A sociophonetic study of the VOT in Polish learners of English. In K. Dziubalska-Kołaczyk, M. Wrembel and K. Kul (eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010*. Bern: Peter Lang, 301-314.

Waniek-Klimczak, E. and Klimczak, K. (2005) Target in speech development: learners' views. In K. Dziubalska-Kolaczyk and J. Przedlacka (eds.), *English pronunciation models: A changing scene.* Bern: Peter Lang, 229-250.

Waniek-Klimczak, E., Rojczyk, A. and Porzuczek, A. (2014) 'Polglish' in Polish eyes: What English studies majors think about their pronunciation in English. In E. Waniek-Klimczak and M. Pawlak (eds.), *Teaching and researching the pPronunciation of English: Studies in honour of Włodzimierz Sobkowiak*. Heidelberg: Springer, 23-34.

Weckwerth, J. (2011) English TRAP Vowel in Advanced Polish Learners: Variation and System Typology. *City University of Hong Kong, Volume Proceedings of the 17th International Congress of Phonetic Sciences*, 2110-2113.

Welkowitz, J. and Feldstein, S. (1969) Dyadic interaction and induced differences in perceived similarity. *Proceedings of the 77th Annual Convention of the American Psychological Association* 4, 343-344.

Welkowitz, J. and Feldstein, S. (1970) The relation of experimentally manipulated interpersonal perception and psychological differentiation to the temporal patterning of conversation. *Proceedings of the 78th Annual Convention of the American Psychological Association* 5, 387-388.

Welkowitz, J., Finklestein, M., Feldstein, S. and Aylesworth, L. (1972) Changes in vocal intensity as a function of interspeaker influence. *Perceptual and Motor Skills* 35, 715-718.

Wells, J. C. (1962) A study of the formants and pure vowels of British English. Unpublished MA dissertation. University of London.

Wierzchowska, B. (1980) *Fonetyka i fonologia języka polskiego*. Wrocław: Ossolineum.

Wiik, K. (1965) Finnish and English vowels: a comparison with special reference to the learning problems met by native speakers of Finnish learning English. *Annales Universitatis Turkensis, Series B, No 94*. Turku: Turun Yliopisto.

Wrembel, M. (2004) Phonological "know that" or "know how"? - in pursuit of determinants of second language pronunciation attainments. In W. Sobkowiak and E. Waniek-Klimczak (eds.), *Zeszyt Naukowy Instytutu Neofilologii (3), Materiały z konferencji "Dydaktyka fonetyki języka obcego w Polsce" Mikorzyn k. Konina 10-12 maja 2004 r*. Konin: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej w Koninie, 163-170.

Wrembel, M. (2005) Metacompetence-oriented model of phonological acquisition: implications for the teaching and learning of second language pronunciation. *Proceedings of the Phonetics Teaching and Learning Conference PTLC2005*. London: UCL, 1-4.

Wode, H. (1981) *Learning a second language, I: An integrated view of language acquisition*. Tübingen: Gunter Narr Verlag.

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Kim, M. and Bradlow, A. R. (2010) The Wildcat Corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language & Speech* 53, 510 –540.

Young, R. (1988) Variation and the Interlanguage Hypothesis. *Studies in Second Language Acquisition* 10, 281-302.

Yu, A. C., Abrego-Collier, C. and Sonderegger, M. (2013) Phonetic imitation from an individual-difference percpective: Subjective attitude, personality and "autistic" traits. *PLoS ONE* 8, 1-13.

Zając, M. (2013) Phonetic imitation of vowel duration in L2 speech. *Research in Language* 11, 19-29.

Zając, M. (forthcoming) Variability in L2 English pronunciation examined through the prism of phonetic imitation. In E.Waniek-Klimczak and A. Cichosz (eds.), *Variability in English across time and space*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Zając, M. and Rojczyk, A. (2014) Imitation of English vowel duration upon exposure to native and non-native speech. *Poznan Studies in Contemporary Linguistics* 50, 495-514.

Zampini (1994) The role of native language transfer and task formality in the acquisition of Spanish Spirantization. *Hispania* 77, 470-481.

Zborowska, J. (1997) Self-organization in second language acquisition. The case of phonostylistics. Unpublished MA dissertation. Adam Mickiewicz University, Poznań.

Zuengler, J. (1982) Applying Accommodation Theory to variable performance data in L2. *Studies in Second Language Acquisition* 4, 181-192.

Zuengler, J. (1985) The effect of induced and perceived expertise on the language performance of native and non-native speakers. Unpublished PhD dissertation. Teachers College, Columbia University.

Zuengler, J. (1989) Assessing an interaction-based paradigm: How accommodative should we be? In M. R. Eisenstein (ed.), *The dynamic interlanguage: empirical studies in second language variation*. New York: Plenum, 49-67.

Zuengler, J. (1991) Accommodation in native-non-native interactions: Going beyond the "what" to the "why" in second-language research. In Coupland, J., Coupland N. and H. Giles H. (eds.), *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press, 223-244.

Summary

This dissertation examines variability in the phonetic performance of L2 users of English and concentrates on speech convergence as a result of exposure to native and non-native pronunciation. The term speech convergence refers to a process during which speakers adapt their linguistic behaviour according to who they are talking or listening to. Previous studies show that the phenomenon may take place both in a speaker's L1 (e.g. Giles, 1973; Coupland, 1984; Gregory and Webster, 1996; Pardo, 2006; Babel; 2010) and L2 (e.g. Beebe, 1977; Berkowitz, 1986; Lewandowski, 2012; Rojczyk, 2013; Trofimovich and Kennedy, 2014). Speech convergence can be subdivided into three types of linguistic behaviour: convergence (the process of making one's speech more similar to that of another person), divergence (the process of moving away from the speech of another person) and maintenance (the process of maintaining one's default linguistic behaviour in spite of exposure to the speech of another person).

The dissertation consists of four chapters; the first two provide theoretical background, the next two describe the study and its findings. Chapter One is concerned with previous research on speech convergence. The chapter reviews the methodology and approaches used in previous work and discusses the range of factors that may affect convergence strategies. Chapter Two provides an overview of relevant studies in the field of L2 phonetics. It describes the structure and formation of the L2 sound system and the numerous social-psychological, linguistic and psycholinguistic variables that may influence L2 phonetic performance. Chapter Three describes the study on speech convergence in the pronunciation of Polish learners of English, i.e. the aims, hypotheses, methodology and results. In Chapter Four, the results of the study on phonetic convergence in the speech of Polish learners of English are analysed and discussed.

The phenomenon of speech convergence has been explored under different names and with the use of various frameworks and methodological procedures. Some researchers refer to the process as accommodation and investigate it by analysing spontaneous conversational data (e.g. Giles, 1973; Bourhis and Giles, 1977; Coupland, 1984; Gregory and Webster, 1996). Other researches use the term imitation and examine the phenomenon in socially minimal, laboratory-based settings (e.g. Goldinger, 1998; Schokley et al., 2004; Delvaux and Soquet, 2007; Nielsen, 2011). Irrespective of terminological and methodological differences, the results of previous studies on phonetic convergence indicate that the process is conditioned by

a variety of linguistic (e.g. Mitterer and Ernestus, 2008; Babel, 2009; Brouwer et al., 2010; Nielsen, 2011) and social-psychological factors (Giles, 1973; Bilous i Krauss, 1988; Gregory and Webster, 1996; Pardo, 2006; Babel, 2009, Yu et al., 2013)

Research on L2 acquisition and non-native pronunciation shows that the development of the L2 sound system is a complex and dynamic process. It has been argued that the productions of L2 users are generated by interlanguage (IL), an independent linguistic system that encompasses elements of the learner's L1 and L2 but does not correspond exactly to either the NL or the TL (e.g. Selinker, 1972; 1992). Importantly, previous findings indicate that the phonetic performance of non-native speakers is influenced not only by their L1 and L2 sound systems but also by a range of various psycholinguistic (e.g. Flege, 1987; Flege et al., 2003) and social-psychological factors (e.g. Taylor et al., 1971; Zuengler, 1982; Gatbonton et al., 2011).

The process of adapting one's pronunciation as a result of exposure to another person's speech has been detected in the productions of L2 users (e.g. Beebe, 1977; Berkowitz, 1986; Lewandowski, 2012; Rojczyk, 2013; Trofimovich and Kennedy, 2014). Similarly as in the case of L1 speech convergence, previous studies show that the magnitude of L2 speech convergence may depend upon a variety of social-psychological and linguistic variables.

An interesting aspect of L2 phonetic convergence that has not yet been thoroughly explored is the comparison of pronunciation shifts upon exposure to the speech of native speakers of the TL as compared with pronunciation shifts upon exposure to the speech of other learners. The aim of the study was to address this issue by investigating and comparing L2 convergence strategies upon exposure to native and non-native pronunciation. The study concentrated on the phonetic performance of advanced Polish learners of English, who were exposed to two pronunciation varieties: Polish-accented English and native English.

The participants were 38 native speakers of Polish, majoring in English Studies and recruited from the University of Lodz. The subjects listened to pre-recorded productions provided by two model talkers/interlocutors: a native speaker of Standard Southern British English and a native speaker of Polish (a qualified phonetician imitating a heavy Polish accent in English). The phonetic variables under investigation were the following: aspiration in word-initial /p t k/, pre-voicing in word-initial /b d g/, vowel duration as a cue for consonant voicing in English /æ e ɪ iː/. The experimental procedure consisted of several phases. First, the informants were instructed to identify the target words in an auditory naming task (baseline condition). Next, they were asked to listen to pre-recorded English words provided by the two

model talkers/interlocutors and to identify the words by saying them out loud (imitation condition). Finally, the subjects were required to read the target words for the two model talkers/interlocutors to listen to at a later time (accommodation condition). Following the production stage of the experiment, the participants completed a questionnaire whose purpose was to gauge attitudes towards native and foreign-accented English.

Three hypotheses were formulated to be tested in the course of the study. Hypothesis 1 predicted that convergence strategies following exposure to native and non-native English will vary as a function of model talker/interlocutor. Hypothesis 2 predicted that convergence strategies following exposure to native and non-native English will be affected by the subjects' attitudes towards native and Polish-accented English. Hypothesis 3 predicted that convergence strategies following exposure to native and non-native English will differ as a function of phonetic context (place of articulation and vowel category).

Acoustic and statistical analysis of the data revealed that the subjects modified their linguistic behaviour following exposure to the speech of the model talkers/interlocutors, which corroborates the claim that L2 speech convergence phenomena are present in non-native pronunciation. Hypothesis 1 was partially supported by the results of the study. It was found that speech behaviour following exposure to native and non-native English varied as a function of model talker/interlocutor in all but two instances (accommodation on pre-voicing and imitation of vowel duration). The results suggests that when using a second language, speakers may use different convergence strategies depending on the native/non-native status of the model talker or interlocutor. Hypothesis 2 was partially supported by the data. The results indicate that a strong preference for target-like pronunciation may prompt learners to converge towards native speech and diverge from foreign-accented speech. However, the factor does not seem to operate if a learner has not succeeded in mastering a given TL pronunciation feature, i.e. the impact of attitudinal factors on the magnitude of convergence in non-native pronunciation appears to be conditioned by the stage of acquisition of a given TL phonetic feature. Hypothesis 3 was not borne out the results obtained in the study. It was found that convergence strategies following exposure to native and non-native English did not vary depending on phonetic context. Overall, the findings of the study provide support for the claim that the process of speech convergence operates in L2 pronunciation and imply that certain social-psychological and psycholinguistic factors may have an impact on learners' convergence strategies.

Streszczenie

Niniejsza rozprawa jest poświęcona dynamice zmian w wymowie nierodzimych użytkowników języka i skupia się na zjawisku konwergencji językowej w wymowie polskich użytkowników języka angielskiego. Termin konwergencja fonetyczna lub konwergencja językowa (ang. phonetic convergence, speech convergence) odnosi się procesu, podczas którego dana osoba zmienia swój sposób mówienia na skutek zetknięcia się z mową innej osoby. Badania pokazują, że zjawisko konwergencji językowej może mieć miejsce zarówno kiedy dana osoba mówi w swoim języku ojczystym (np. Giles, 1973; Coupland, 1984; Gregory i Webster, 1996; Pardo, 2006; Babel; 2010) jak i w języku obcym (np. Beebe, 1977; Berkowitz, 1986; Lewandowski, 2012; Rojczyk, 2013; Trofimovich i Kennedy, 2014). W literaturze wyróżnia się trzy strategie związane z tym zjawiskiem: konwergencję (ang. convergence; Giles, 1973; Giles i Ogay, 2007), która polega na przybliżaniu swojego sposobu mówienia do sposobu mówienia innej osoby, dywergencję (ang. divergence; Giles, 1973; Giles i Ogay, 2007), która polega na oddalaniu się w swoim sposobie mówienia od mowy innej osoby oraz podtrzymywanie (ang. maintenance; Giles i Ogay, 2007), które polega na zachowaniu swojego zwyczajowego sposobu mówienia pomimo zetknięcia się z mową innej osoby.

Pierwsze dwa rozdziały rozprawy koncentrują się na kwestiach teoretycznych. Rozdział pierwszy przedstawia zagadnienia związane z konwergencją językową i omawia wcześniejsze badania poświęcone temu zjawisku. Rozdział drugi odnosi się do kształtowania się systemu dźwiękowego w języku obcym i opisuje wyniki najważniejszych (z punktu widzenia rozprawy) badań nad wymową w drugim języku. Dwa kolejne rozdziały dotyczą badania empirycznego nad konwergencją językową w wymowie polskich użytkowników języka angielskiego. Rozdział trzeci omawia wcześniejsze badania poświęcone zjawisku konwergencji w wymowie nierodzimych użytkowników języka a także przedstawia cele, hipotezy, metodologię oraz wyniki obecnego badania. Szczegółowa analiza oraz omówienie wyników znajdują się w rozdziale czwartym.

Proces konwergencji językowej był badany przy użyciu różnych rodzajów metodologii, a także pod różnymi nazwami. Część badaczy określa to zjawisko mianem akomodacji (ang. accommodation) i bada je poprzez analizę danych pochodzących z konwersacji między uczestnikami badania a ich interlokutorami. (np. Giles, 1973; Bourhis i Giles, 1977; Coupland, 1984; Gregory i Webster, 1996). Niektórzy nazywają ten proces imitacją (ang.

imitation) i badają go w kontekście laboratoryjnym, instruując uczestników badania, aby powtarzali słowa za głosem z nagrania (np. Goldinger, 1998; Schokley et al., 2004; Delvaux i Soquet, 2007; Nielsen, 2011). Niezależnie od zastosowanej metodologii i terminologii, wyniki wcześniejszych badań nad konwergencją językową pokazują, że zjawisko to jest zależne od rozmaitych czynników, w tym zarówno językowych (np. Mitterer i Ernestus, 2008; Babel, 2009; Brouwer et al., 2010; Nielsen, 2011) jak i socjologicznych czy psychologicznych (np. Giles, 1973; Bilous i Krauss, 1988; Gregory i Webster, 1996; Pardo, 2006; Babel, 2009, Yu et al., 2013).

Wyniki badań nad formowaniem się systemu dźwiękowego w drugim języku wskazują, że proces ten jest niezwykle złożony. Powszechnie uważa się, że wypowiedzi osób uczących się obcego języka są generowane przez niezależny system językowy, nazywany interjęzykiem (ang. interlanguage, np. Selinker, 1972; 1992). System ten zawiera elementy języka ojczystego jak i języka obcego, ale nie odpowiada ściśle żadnemu z nich. Co ważne, wcześniejsze badania pokazują, że poza językiem ojczystym i językiem obcym, na wymowę osoby uczącej się drugiego języka mogą mieć wpływ różnorakie uwarunkowania psycholingwistyczne (np. Flege, 1987; Flege et al., 2003), socjologiczne i psychologiczne (np. Taylor et al., 1971; Zuengler, 1982; Gatbonton et al., 2011).

Zjawisko konwergencji językowej zostało wykryte w wymowie nierodzimych użytkowników języka w kilku poprzednich badaniach (np. Beebe, 1977; Berkowitz, 1986; Lewandowski, 2012; Rojczyk, 2013; Trofimovich i Kennedy, 2014). Podobnie jak w przypadku konwergencji w języku ojczystym, badania te pokazują, że kiedy proces ten zachodzi w drugim języku, mogą na niego oddziaływać czynniki socjologiczne oraz psychologiczne jak i czynniki językowe.

Podstawową motywacją dla przeprowadzenia badania, które stanowi podstawę tej rozprawy był fakt, iż zjawisko konwergencji językowej w wymowie nierodzimych użytkowników nie zostało jeszcze dogłębnie przebadane. Ponadto, bardzo niewiele z dotychczasowych badań skupiało się na porównaniu strategii językowych (tj. konwergencji, dywergencji i podtrzymania) stosowanych po zetknięciu się z mową rodzimych użytkowników danego języka ze strategiami językowymi stosowanymi po zetknięciu się z mową nierodzimych użytkowników tego samego języka. Głównym celem badania empirycznego opisanego w tej rozprawie była analiza wpływu zetknięcia się z wymową rodzimych i nierodzimych użytkowników języka angielskiego na proces konwergencji fonetycznej w mowie Polaków uczących się angielskiego.

Uczestnikami badania było 38 studentów II roku filologii angielskiej, którzy słuchali angielskich słów wypowiadanych przez dwóch mówców: Anglika oraz Polaka mówiącego po angielsku z wyraźnym polskim akcentem. Badanie skupiało się na trzech zmiennych językowych: przydechu (aspiracji) w angielskich spółgłoskach /p t k/, dźwięczności w angielskich spółgłoskach /b d g/, długości angielskich samogłosek /æ e ɪ iː/ przed spółgłoskami dźwięcznymi i bezdźwięcznymi. Zastosowana metodologia opierała się na kilkuczęściowym eksperymencie. W pierwszej fazie eksperymentu uczestnicy badania zostali poproszeni o wypowiedzenie 48 angielskich słów przedstawionych na obrazkach. Ta część eksperymentu została nazwana fazą bazową (ang. baseline), a słowa w niej użyte zawierały wyżej wymienione zmienne językowe. Następnie studenci słuchali tych samych 48 angielskich słów uprzednio nagranych na potrzeby badania przez wyżej wspomnianych mówców (Anglika i Polaka) i powtarzali je zaraz po usłyszeniu. Ta część eksperymentu nosiła nazwę imitacji i została podzielona na dwie części: najpierw studenci słuchali słów wypowiedzianych przez polskiego mówcę, a potem tych samych słów wypowiedzianych przez angielskiego mówcę. W kolejnej części eksperymentu uczestnicy zostali poproszeni o przeczytanie 48 angielskich słów (tych samych, które zostały użyte poprzednio) oraz poinformowani, że Polak i Anglik, których słuchali również zostaną poproszeni o wysłuchanie słów przeczytanych przez uczestników badania w tej części eksperymentu. Ta faza nosiła nazwę akomodacji i została podzielona na dwie części: podczas pierwszej z nich uczestnicy czytali słowa dla Polaka, podczas drugiej z nich czytali je dla Anglika. Po zakończeniu eksperymentu uczestnicy badania zostali poproszeni o wypełnienie ankiety, która miała na celu zbadanie stosunku studentów do wymowy w języku angielskim. Należy również wspomnieć, że zastosowana metodologia została oparta na badaniach pilotażowych (badania pilotażowe są opisane w rozdziale trzecim).

Przeprowadzone badanie miało na celu weryfikację trzech hipotez. Według pierwszej z nich, po zetknięciu się z wymową polskiego mówcy uczestnicy badania zastosują inne strategie językowe (tj. konwergencję, dywergencję, podtrzymanie) niż po zetknięciu się z wymową angielskiego mówcy. Druga hipoteza przewidywałą, że zastosowane strategie językowe będą zależne od stosunku uczestników badania do wymowy w języku angielskim. Hipoteza trzecia zakładałą, iż kontekst fonetyczny (miejsce artykulacji w przypadku przydechu i dźwięczności oraz rodzaj samogłoski w przypadku długości samogłoski) będzie miał wpływ na zastosowane strategie językowe.

Analiza akustyczna i statystyczna zebranych próbek mowy ujawniła, że wymowa uczestników badania uległa zmianom pod wpływem zetknięcia się z mową rodzimego i nierodzimego użytkownika języka angielskiego, tj. zjawisko konwergencji fonetycznej zostało wykryte w wymowie Polaków uczących się języka angielskiego. Hipoteza pierwsza została zweryfikowana przez rezultaty badania. Wyniki wskazują, że uczestnicy badania stosowali inne strategie językowe po zetknięciu się z wymową polskiego mówcy niż po zetknięciu się z wymową angielskiego mówcy. Hipoteza druga została częściowo potwierdzona przez wyniki badania. Większość studentów była negatywnie nastawiona do wymowy z wyraźnym polskim akcentem co, jak sugerują wyniki badania, skłoniło ich do przybliżenia się do sposobu mówienia angielskiego mówcy i zastosowaniu strategii podtrzymania i dywergencji w przypadku polskiego mówcy. Tendencja ta nie została jednak zaobserwowana w przypadku wszystkich badanych zmiennych językowych. Rezultaty badania wskazują, że na strategie językowe stosowane przez studentów miał również wpływ stopień przyswojenia danej cechy wymowy. Wydaje się, że uczestnicy badania dostosowywali swój sposób mówienia do wymowy rodzimego użytkownia tylko w przypadku tych cech wymowy, które zostały przez nich opanowane. Cechy, które nie zostały jeszcze w pełni opanowane przez uczestników badania były natomiast bardziej podatne na wpływ języka ojczystego. Hipoteza trzecia nie została potwierdzona przez wyniki badania, tj. kontekst fonetyczny okazał się nie wywierać znaczącego wpływu na zastosowane przez uczestników badania strategie językowe. Podsumowując, wyniki badania nad konwergencją językową w wymowie polskich użytkowników języka angielskiego potwierdzają wniosek, że zjawisko to ma miejsce w mowie nierodzimych użytkowników języka i wskazują, że proces ten podlega pewnym uwarunkowaniom psychologicznym i psycholingwistycznym.

Appendix A

Slides from the PowerPoint presentation used in the baseline task.

Slide 1

Dan	tan

Slide 2

goat	coat

Slide 3

pun	bun

Slide 4

pat	bat

Slide 5

seat	seed

Slide 6

bed	bet

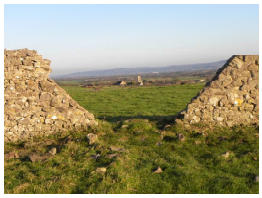
Slide 7

sit	Sid

Slide 8

bed	bet

Slide 9



gap	cap

Slide 10

bid	bit

Slide 11

goat	coat

Slide 12

bad	bat

Slide 13

Dutch	touch

Slide 14

got	cot

Slide 15

said	set

Slide 16

sad	sat

Slide 17

bead	beat

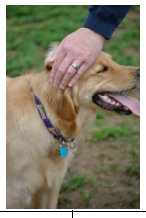
Slide 18

mat	mad

Slide 19

bid	bit

Slide 20



pat	bat

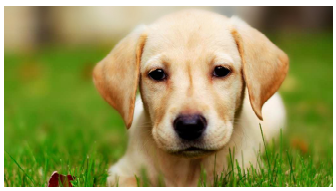
Slide 21

bed	bet

Slide 22

met	med

Slide 23



tog	dog

Slide 24

mitt	mid

Slide 25

dip	tip

Slide 26

beat	bead

Slide 27

Dan	tan

Slide 28

met	med

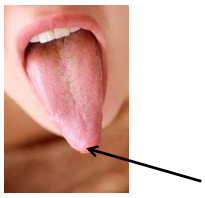
Slide 29

Sid	sit

Slide 30

gut	cut

Slide 31



tip	dip

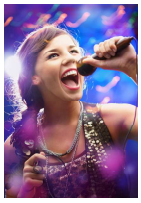
Slide 32

tog	dog

Slide 33

set	said

Slide 34



bop	pop

Slide 35

seat	seed

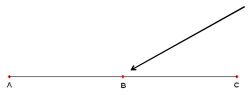
Slide 36

mead	meat

Slide 37

pun	bun

Slide 38



mid	mitt

Slide 39

gut	cut

Slide 40

mead	meat

Slide 41

gap	cap

Slide 42

pop	bop

Slide 43

sad	sat

Slide 44

got	cot

Slide 45

Dutch	touch

Slide 46

pat	bat

Slide 47

pet	bet

Slide 48

mat	mad

Appendix B

The exercise sheet used is the imitation tasks.

1	
Dan	tan

2	
goat	coat

3	
pun	bun

4	
pat	bat

5	
seat	seed

6	
pet	bet

7	
Sid	sit

8	
bed	bet

9	
gap	cap

10	
bid	bit

11	
goat	coat

12	
bad	bat

13	
Dutch	touch

14	
got	cot

15	
said	set

16	
sad	sat

17	
bead	beat

18	
mat	mad

19	
bid	bit

20	
pat	bat

21	
bed	bet

22	
met	med

23	
tog	dog

24	
mitt	mid

25	
dip	tip

26	
beat	bead

27	
Dan	tan

28	
met	med

29	
sit	Sid

30	
gut	cut

31	
tip	dip

32	
tog	dog

33	
set	said

34	
bop	pop

35	
seat	seed

36	
mead	meat

37	
pun	bun

38	
mid	mitt

39	
gut	cut

40	
mead	meat

41	
gap	cap

42	
pop	bop

43	
sad	sat

44	
got	cot

45	
Dutch	touch

46	
bad	bat

47	
bet	pet

48	
mat	mad

Imię .....

Imię drugiej osoby .....

Czy łatwo było Ci zrozumieć osobę, której słuchałaś/eś?

a) bardzo łatwo
b) łatwo
c) średnio
d) trudno
e) bardzo trudno

Appendix C

Questionnaire

Imię …… **Wiek** ……

NASTĘPUJĄCA CZĘŚĆ ANKIETY DOTYCZY WYMOWY POLAKA, KTÓREGO SŁUCHAŁAŚ/EŚ

Zaznacz w jakim stopniu zgadzasz się z poniższymi twierdzeniami a potem odpowiedz na pytania.

1. Poprawnie wymawiał słowa.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

2. Jego wymowa sprawiała, że przyjemnie mi się go słuchało.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

3. Jego wymowa sprawiała, że brzmiał inteligentnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

4. Jego wymowa sprawiała, że brzmiał profesjonalnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

5. Jego wymowa sprawiała, że brzmiał jak osoba wykształcona.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

6. Jego wymowa sprawiała, że brzmiał przyjaźnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

7. Jego wymowa sprawiała, że brzmiał atrakcyjnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

8. Czy któreś z podanych cech wymowy **zwróciły Twoją uwagę** kiedy słuchałaś/eś tej osoby?

a) brak aspiracji

b) nieodpowiednia długość samogłosek przed dźwięcznymi i bezdźwięcznymi spółgłoskami

c) używanie /i/ w słowach takich jak *bid* czy *sit*

d) używanie /e/w słowach takich jak *bat* czy *sat*

e) inna cecha: ……

9. Czy starałaś/eś się zmienić lub dostosować swoją wymowę kiedy czytałaś/eś słowa dla tej osoby? Dlaczego? W jaki sposób?

……

……

……

……

……

……

……

NASTĘPUJĄCA CZĘŚĆ ANKIETY DOTYCZY WYMOWY <u>BRYTYJCZYKA</u>, KTÓREGO SŁUCHAŁAŚ/EŚ

Zaznacz w jakim stopniu zgadzasz się z poniższymi twierdzeniami a potem odpowiedz na pytania.

10. Poprawnie wymawiał słowa.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

11. Jego wymowa sprawiała, że przyjemnie mi się go słuchało.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

12. Jego wymowa sprawiała, że brzmiał inteligentnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

13. Jego wymowa sprawiała, że brzmiał profesjonalnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

14. Jego wymowa sprawiała, że brzmiał jak osoba wykształcona.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

15. Jego wymowa sprawiała, że brzmiał przyjaźnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

16. Jego wymowa sprawiała, że brzmiał atrakcyjnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

17. Czy któreś z podanych cech wymowy **zwróciły Twoją uwagę** kiedy słuchałaś/eś tej osoby?

a) aspiracja

b) dłuższe samogłoski przed dźwięcznymi spółgłoskami

c) jakość samogłosek

d) inna cecha: ……

18. Czy starałaś/eś się zmienić lub dostosować swoją wymowę kiedy czytałaś/eś słowa dla tej osoby? Dlaczego? W jaki sposób?

……

……

……

……

……

……

……

NASTĘPUJĄCA CZĘŚĆ ANKIETY DOTYCZY PRZEBIEGU EKSPERYMENTU

Odpowiedz na poniższe pytania.

19. Podczas rozpoznawania słów **na obrazkach**

a) skupiałam/em się na swojej wymowie i starałam/em się wymawiać słowa tak jak zrobiłby to native speaker

b) nie zwracałam/em uwagi na swoją wymowę

c) inne: ……

20. Podczas rozpoznawania słów, **które wypowiadał Polak**

a) skupiałam/em się na swojej wymowie i starałam/em się wymawiać słowa tak jak zrobiłby to native speaker

b) nie zwracałam/em uwagi na swoją wymowę

c) inne: ……

21. Podczas rozpoznawania słów, **które wypowiadał Brytyjczyk**

a) skupiałam/em się na swojej wymowie i starałam/em się wymawiać słowa tak jak zrobiłby to native speaker

b) nie zwracałam/em uwagi na swoją wymowę

c) inne: ……

NASTĘPUJĄCA CZĘŚĆ ANKIETY DOTYCZY OPINII NA TEMAT WYMOWY

Odpowiedz na poniższe pytanie a potem zaznacz w jakim stopniu zgadasz się z podanymi twierdzeniami.

22. Kiedy mówię po angielsku, chciałabym/chciałbym brzmieć jak:

a) Amerykanin/Amerykanka

b) Brytyjczyk/Brytyjka

c) Polak/Polka

d) nie ma dla mnie znaczenia jak brzmię dopóki jestem w stanie się porozumieć z innymi osobami

d) inne: ……..

23. Nie chciałabym/chciałabym mówić po angielsku z wyraźnym polskim akcentem.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

24. Chciałabym/chciałbym brzmieć jak native speaker kiedy mówię po angielsku, nawet jeśli kosztowałoby mnie to dużo czasu i wysiłku.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

25. Wymowa, jaką mamy mówiąc po angielsku jest tak samo ważna jak słownictwo i struktury gramatyczne, których używamy.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

26. Ludzie powinni starać się mówić jak native speakerzy kiedy uczą się języka angielskiego.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

27. Nie podoba mi się kiedy Polacy mówią po angielsku z wyraźnym polskim akcentem.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

28. Nie lubię słuchać wymowy osób, które mówią po angielsku z wyraźnym polskim akcentem.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

29. Mówienie po angielsku z wyraźnym polskim akcentem sprawia, że brzmi się śmiesznie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

30. Mówienie po angielsku z wyraźnym polskim akcentem sprawia, że brzmi się mało inteligentnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

31. Mówienie po angielsku z wyraźnym polskim akcentem sprawia, że brzmi się nieprofesjonalnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

32. Mówienie po angielsku z wyraźnym polskim akcentem sprawia, że brzmi się jak osoba niewykształcona.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

33. Mówienie po angielsku z wyraźnym polskim akcentem sprawia, że brzmi się mało atrakcyjnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

34. Mówienie po angielsku z wyraźnym polskim akcentem sprawia, że brzmi się nieprzyjaźnie.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

35. Mówienie po angielsku z wyraźnym polskim akcentem to używanie niepoprawnej wymowy.

5	4	3	2	1
zdecydowanie tak	tak	nie wiem	nie	zdecydowanie nie

DZIĘKUJĘ ZA POMOC!