

Filozofická fakulta Univerzity Palackého
Katedra anglistiky a amerikanistiky

Intelligibility of Foreign-Accented English

(Bakalářská práce)

Autor: Petra Vybíralová

Studijní obor: Anglická filologie

Vedoucí práce: Mgr. Šárka Šimáčková, PhD.

Olomouc 2014

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a uvedla úplný seznam citované a použité literatury.

V Olomouci dne 10. 12. 2014

.....

I would like to thank Mgr. Šárka Šimáčková, M.A., Ph.D. for her invaluable help in the preparation of this thesis.

CONTENT

1	INTRODUCTION	2
2	INTERLANGUAGE	4
2.1	Language transfer	5
3	FOREIGN ACCENT	6
3.1	Origin of foreign accent.....	7
3.1.1	Age of learning - critical period	7
3.1.2	Length of residence	10
3.2	Assessment of foreign accent	10
4	INTELLIGIBILITY	12
4.1	Interrelation between accentedness, comprehensibility and intelligibility in previous studies	12
4.2	Intelligibility and first language background	15
5	THE INTERLANGUAGE SPEECH INTELLIGIBILITY BENEFIT	20
5.1	Bent and Bradlow 2003	21
5.1.1	Participants and methodology	22
5.1.2	Results and conclusion	24
5.2	Studies following Bent and Bradlow (2003).....	25
6	STUDY PROPOSAL	29
6.1	Introduction	29
6.2	Methodology.....	31
6.2.1	Speakers.....	31
6.2.2	Listeners	32
6.2.3	Stimuli	33
6.2.4	Procedure.....	34
6.2.5	Further suggestions.....	35
7	CONCLUSION	36
8	ZÁVĚR	CHYBA! ZÁLOŽKA NENÍ DEFINOVÁNA.
9	REFERENCES	38
10	ANOTACE	43

1 Introduction

The question of intelligibility of second language speech has usually been discussed in the Second Language Acquisition (SLA) research studies in relation to other features of both native (L1) and non-native (L2) speech. In most of the cases it was the relation between intelligibility, comprehensibility and foreign accent that was being examined (Munro and Derwing 1995, 1999, Munro, Derwing and Morton 2006, Derwing and Munro 1997). Munro and Derwing (1999, 289) defined intelligibility as “the extent to which a speaker’s utterance is actually understood”, accentedness as “the degree to which the pronunciation of an utterance sounds different from an expected pronunciation pattern” and comprehensibility as “the listener’s estimation of difficulty in understanding an utterance” (290).

Experiments proved that native listeners find native speech generally more intelligible (Munro and Derwing 1995) and non-native listeners find non-native speech more (Nash 1969) or as intelligible (van Wijngaarden 2001) as native speech. Bent and Bradlow (2003) further investigated this non-native talker-listener advantage and found an effect they labelled the “interlanguage speech intelligibility benefit” (1600). Their study also proposed definitions of two different realizations of this phenomenon based on participants’ L1 background.

The “matched interlanguage speech intelligibility benefit” (1606) occurred when L2 listeners assigned equal intelligibility scores to both high proficiency L2 speakers from the same L1 background and to native speakers’ utterances. On the other hand, the “mismatched interlanguage speech intelligibility benefit” (1607) ensued when L2 listeners rated high proficiency L2 speakers from various L1 backgrounds equally intelligible as native speakers.

Significantly more experiments in SLA research have been carried out from the perspective of native English listeners responding to foreign-accented speech (Flege and Fletcher 1992, Munro and Derwing 1995, Derwing and Munro 1997) I have decided to take into account the perspective of non-native listeners for the purpose of my study. The chief goal of my thesis is to review existing approaches to L1 and L2 speech intelligibility and attempt to design an experiment which would help to establish the extent to which listeners from a variety of L1 backgrounds share a response to native and non-native speech. The later part of my work should help to

reveal evidence either for or against the previously mentioned notion of “matched and mismatched interlanguage speech intelligibility benefit”. The study will also examine a possible extension of the hypotheses by investigating similarity in responses of a particular group of listeners from a shared language family or geographical area.

The methodology used in the previously mentioned research studies varies to a great degree, but most of the researchers decided to use the method of dictation task in which listeners hear utterances and write them down in standard orthography. Among the other methods used were picture selection in response to the stimulus (Smith and Bisazza 1982) and determination of truth value (Munro and Derwing 1995). Each of these approaches has strengths and drawbacks and it will be my priority to find a reliable combination of previously used methods or to come up with a method completely different. I will also try to suggest areas requiring further investigation.

2 Interlanguage

The process of (second) language learning is accompanied by other underlying psycholinguistic processes and creation of systems in the brain of the speaker. The first researchers to show interest in this topic were two American linguists Pit Corder and Larry Selinker. Their focus was the study of formal distinctions between the first and the second language of a learner and was highly influenced by behaviourism and Chomsky's theory of universal grammar. One of Corder's greatest contributions to the SLA field was the Error Analysis approach, which was a kind of alternative to the previously established Contrastive Analysis. Both of these approaches have worked with the learner's production and used the observable data as an input for predicting errors in second language acquisition. This approach was later on overshadowed by a concept of "interlanguage".

For the purpose of analysing the process of learning of a second language, Selinker (1972) has made a framework of concepts and definitions. Firstly, he extended the generally accepted definition of target language (TL) to a unique norm of "one dialect within the interlingual focus of attention of the learner" (213). He also directed his analytical focus on the individual utterances which the learner produces in order to communicate in the TL and made these "the only observable data upon which [he] can relate theoretical predictions" (213).

This set of utterances for most¹ learners of a second language is not identical to the hypothesized corresponding set of utterances which would have been produced by a native speaker of the TL had he attempted to express the same meaning as the learner (214).

This observation, the fact that the communicated output of native and non-native speaker of a TL is crucially different allows the examiner to presume an existence of "a separate linguistic system" (214) in the mind of the TL learner. Selinker has named this system an "interlanguage" (IL) (214). He has also created a framework of psychologically relevant data for theoretical predictions about L2 learning and the structure of an interlanguage. Within this framework were "utterances in the learner's

¹ Learners who started acquiring L2 after childhood

native language produced by the learner”, “IL utterances produced by the learner” and “TL utterances produced by a native speaker of that TL” (214).

The work of Corder, Selinker and others helped to reveal the fact, that “adult second language learners from a given native language background progress in a relatively consistent way from a monolingual to a bilingual system” as well as that such learners from “different native language background exhibit different deviations from the target language norms” (Bent and Bradlow 2003, 1600) and their findings could later on be implicated in both SLA teaching and learning instruction.

2.1 Language transfer

In connection with the notion of interlanguage, Selinker (1972) differentiates between five central processes of second language learning. One of them is the process of language transfer and occurs when it can be “experimentally demonstrated, that fossilizable items, rules and subsystems which occur in IL performance are a result of the NL” (216). This process of transfer will subsequently be reflected in the foreign-accentedness of the L2 production.

3 Foreign accent

There is no doubt that foreign accent is a very significant phenomenon in Second Language Acquisition (SLA) and serves as a clear evidence that knowledge of first language (L1) influences the acquisition of a second (L2). Even though a lot of attention has been drawn towards it over the past century, there is no precisely stated definition of foreign accent and conceptions of its significance vary with time. Still it might be generally described as a phonological deviation from native form, thus a working definition of foreign accent could be stated as a “deviating realization of corresponding phonological units which can often be traced back to specific features of the respective speakers’ native language” (thus the strong connection between L1 and L2). Munro (1998) defined foreign accent as ‘non-pathological speech produced by second language learners that differs in partially systematic ways from the speech characteristics of native speakers of a given dialect’ (139). One of the first, very disapproving view of this L2 feature appeared in Greene and Wells (1927), who wrote:

Foreign accent, being of the nature of imperfect or defective speech, is the result of incorrect articulation and enunciation and is therefore classified, from out therapeutic viewpoint, as stammering speech (24).

Some years later, Griffen (1980) saw accent as an undesirable feature that has to be eliminated and made this elimination the primary goal of instruction in L2 pronunciation. Nowadays foreign accent is seen as a common aspect of late² second language acquisition and accent-free pronunciation is not a necessary goal for either L2 learners or pronunciation instructors. However, in cases, speaking with a foreign accent does bring unpleasant consequences such as reduced social acceptability, diminished comprehensibility and intelligibility of L2 speech. On the other hand, it can serve as a marker of non-native competence and therefore may eventuate in native interlocutor’s speech adjustment or it can even serve as a communication enhancer, triggering interest in native-foreign interaction.

² See section 2.1.2 Critical Period below

3.1 Origin of foreign accent

A lot of research has been carried out in the area of influences on foreign-accentedness in the past century with a view to better understand this phenomenon as well as to improve the instruction in L2 pronunciation. Although different studies produced different, sometimes conflicting outcomes, several factors seem to repeatedly receive a great degree of attention in SLA literature. Among the most reviewed are the age of L2 learning (AOL), gender, length of residency in L2 speaking country (LOR), motivation and amount of native language use. The AOL and LOR are given a separate paragraph below as they were proved to have the largest effect on foreign accentedness of L2 speech.

3.1.1 Age of learning - critical period

There is a common belief that children are better learners of a second language than adults, although the question why and when such learning advantage ceases to exist is rather controversial in the SLA research field. The topic of language learning abilities with respect to the learner's age was first discussed by American neurosurgeons Wilder Penfield and Lamar Roberts (1959), who claim, that "for the purpose of learning languages, the brain progressively becomes stiff and rigid after the age of nine" (236) and suggest a loss of neural plasticity after a certain point in individual's life. Later on, Lenneberg (1967) extends the implication of this hypothesis onto second language acquisition when he claims that

[...] automatic acquisition from mere exposure to a given language seems to disappear [after puberty], and foreign languages have to be taught and learned through a conscious and labored effort. Foreign accents cannot be overcome easily after puberty (176).

The claim that there is an age-related point in life beyond which it becomes impossible to acquire a second language on the level of a native speaker became known as the Critical Period Hypothesis (CPH). Several researchers (Long 1990, Hurford 1991) found out, that there might be several such points in life, after which different language learning abilities are lost, the ability to produce an accent-free L2 speech being the first of them.

Few studies, however, produced contradictory results, reporting better performance by late than early L2 learners. Snow and Hoefnagel-Hohle (1978) for example reported that later L2 learners were better at imitation of Dutch sounds in individual words but this phenomenon turned out to be temporary as after 10 months young L2 learners began to outperform older participants of the experiment. Such short-term experiments would usually favour adult L2 learners as they “proceed through the morphological and syntactic development faster than children (where time and exposure held constant)” (Krashen 1979, 573). Neufeld (1979) tried to falsificate the CPH and through his experiment proved that adult learners who received specialized training in phonology were able to perform like native speakers on certain tasks. Although it was noted that improved and targeted phonological training can dramatically raise L2 learner’s proficiency, such performance on limited tasks is “not equivalent to consistent performance in naturalistic situations” (Gass and Selinker 2008, 407).

Answers for the question when does the point after which ultimate attainment of an L2 occur and therefore when does the critical period end vary from experiment to experiment. Penfield and Roberts’ (1959) empirical study first proposed the age of nine years as the end of the critical period, Lenneberg’s theory (1967) suggested puberty as a crucial point in L2 learning and Scovel (1988) claimed that critical period ends at the age of twelve years. Patkowski (1990) defined CP as an “age based constraint on the acquisition of full native fluency” (74) and suggested its end at fifteen years. According to the CPH, there is a complete discontinuity, a drop-off in the ability to attain native-like pronunciation of an L2 after a certain point, not all of the researchers agree with this proposal though. The outcomes of their experiments suggest that there is no such discontinuity, but rather declension in the ability and that this declension is not any swifter around the age of puberty than at any other age. This observation suggests the so-called “sensitive period hypothesis”. Susan Oyama (1976) was one of the first experimenter who through her research came to a conclusion favouring SPH, followed by Patkowski (1980)

The term ‘critical period’ refers to the notion that the age limitation is absolute. In theory, first language acquisition is not possible past the critical point. The term ‘sensitive period’, on the other hand, refers to the fact that age

limitation is not absolute. It is indeed possible to acquire a foreign language at an adult age, but it is not possible to do so to the extent of being able to 'pass for native' (449).

Long (1990) defined a starting point for several other studies when, through his experiment, he came to a conclusion, that:

There are sensitive periods governing the ultimate level of first or second language attainment possible in different linguistic domains, not just phonology, with cumulative declines in learning capacity, not a catastrophic one-time loss, and beginning as early as age 6 in many individuals, not at puberty, as is often claimed (255).

This claim was supported by both Flege and Fletcher (1992) and Flege et al. (1995). Their studies showed that L2 speakers whose onset of learning occurs at the age of 6 or before are much more likely to speak without a foreign accent than any other L2 speakers. However, Flege et al (1997) also pointed out that even such early AOL as 6 years does not necessarily lead to an accent-free speech, as a group of subjects in his experiment with mean AOL of 3.2 years evinced a detectable foreign accent.

Due to the fact that AOL is not the only factor affecting the degree of foreign accentedness, there is yet no convincing evidence for the crucial effects of CPH on foreign accentedness. It has been confirmed that early learners speak L2 with a lower degree of foreign accent, but the question whether AOL of 6 years or less will certainly lead to accent-free L2 and AOL onset after puberty will, on the other hand, definitely lead to accented speech still remains unresolved.

The present study proposal will suggest late L2 learners (individuals who started learning L2 after the onset of puberty) as speaking participants. This should ensure the presence of a certain degree of foreign accent and the possibility of a correlation test between intelligibility and foreign accentedness, as it has been found that these two features of L2 speech are partly connected. Subsequently, an accent-familiarity test could be included in order to indicate this factor as a possible determiner of the degree of intelligibility.

3.1.2 Length of residence

The length of residence (LOR) refers to a number of years spent in community where an L2 is the predominant language and even though it is the second most reviewed variable influencing the degree of foreign accent, previous research produced conflicting outcomes in terms of its effect. Flege (1988) found no significance in terms of LOR effect on foreign accentedness in groups of Chinese learners who resided in the L2 speaking country for 1.1 and 5.1 years in average. Flege and Fletcher (1992) on the other hand observed a great LOR effect on accentedness of groups of Spanish speakers whose average LOR was 14.3 and 0.7 years. This was, however, eliminated as a significant predictor of degree of L2 by a multiple regression analysis which identified age of learning as the most important determinant. Asher and García (1969) came to a similar conclusion as Flege and Fletcher (1992) when their study showed that 51% Cuban children who have been to a L2 speaking country had a near-native pronunciation and only 10% of these had a definite foreign accent. In the case of children living in the L2 speaking country for 4 years or less, only 15% attained near-native pronunciation and 55% of these were marked for a definite accent. Studies of Oyama (1976), Tahta *et al.* (1981), Piper and Cansin (1988) and Moyer (1999) did not assign LOR any significance whatsoever. Oyama (1976) and Flege (1988) found a possible explanation for the discrepancies in the results and suggested that after a rapid initial phase of learning, LOR has no effect on foreign accent of late L2 learners and therefore the degree of its influence depends on whether the subjects are still in an early phase of L2 learning or not.

3.2 Assessment of foreign accent

Rating techniques used for assessing the degree of foreign-accentedness of L2 speech do not vary to a large extent. Rather than in their nature, they only “differed in resolution” (Piske, Flege and MacKay 2001, 194). Three types of rating methods have been tested so far; the equal appearing interval (EAI) scale, the visual analogue (VA) scale and direct magnitude estimation (DME). The EAI scale, which ranks as the most common technique used for a perceptual voice evaluation, is an itemized attitude rating scale with one extreme point standing for “strongest possible

appreciation of a value” and other extreme point representing “strongest possible depreciation of that value” (Marshall 1998). These are usually marked as “native-like pronunciation” or “no foreign accent” and “heavy foreign accent” or “definite foreign accent” (Piske, MacKay, Flege 2001). As the name suggests, the intervals between individual scaling points should be congruent. The most common type of EAI scale is a five-point scale, but three, four, six, seven, nine and even ten point scales were also applied (see table 1 below). VA scale is another type of scalar rating technique but instead of choosing between fixed points, listeners are prompted to move a lever or a cursor along a continuum upon which only the two extreme points are marked. The position of such lever then returns a value between 0 and 255. The advantage of such approach is that it allows much finer distinction between listeners’ attitudes. Direct magnitude estimation (DME) is a non-scalar method adapted in foreign accent research. While measuring using DME, one sentence is chosen to be a standard stimulus, usually “a good exemplar of “midrange” value” (Weismer, Laures 2002) of foreign-accentedness (or intelligibility, depending on the focus of the study) and is given a numerical value of 100. Listeners then hear other sentences and are prompted to scale them relative to the standard stimulus. If a sentence is perceived twice as accented as the standard stimulus, listeners would give it a value of 200. Southwood and Flege (1999) examined the reliability of DME versus interval rating of foreign accent. The impaired inter-judge reliability within the DME scaling which occurred in this study was influenced by a great variability in number ranges. While one listener’s values ranged from 1 to 400, other listener only used values from 25 to 160 in his rating of most and least accented stimuli. Other possible factor might have been the number of values used for scaling a stimuli or using a previous stimulus as a reference for the one being currently rated, instead of using a standard stimulus. Based on the study outcome, the examiners came to a conclusion that even though frequently used, a seven-point scale might not be sensitive enough for the purpose of rating a foreign accent. Instead, they proposed a nine- or eleven-point scale as being the best choice for improving the listener sensitivity while rating foreign-accentedness (335).

4 Intelligibility

As I have already stated in the introduction to this work, intelligibility as a feature of foreign language has usually been studied in connection to foreign accent and comprehensibility. The first researchers to mention this phenomenon were Nelson (1982) who defined it as “the apprehension of the message in the sense intended by the speaker” (63) and Munro and Derwing (1995) who also broadly defined it as “the extent to which a speaker’s message is actually understood by a listener” (76). One of the reasons for the increased attention towards this phenomenon was that as early as in 1949, there had been disputes around whether it is a “comfortably intelligible pronunciation” (Abercrombie 1949) or a speech “free of any indication that the speaker is not a clinically normal native” (Griffen 1980) that should stand for the main goal of all pronunciation teaching and therefore for the main goal of all foreign language learning. Even though both of these views had their supporters and objectors, together with a general bias against foreign accented speech they gave rise to the popularity of many accent reduction programs. Similarly to the fact that there is no universally accepted definition of intelligibility, current research field also lacks some universal way of assessing it. Majority of researchers, however, have decided to employ listeners’ orthographic transcriptions as an indicator of the intelligibility of a stimuli.

4.1 Interrelation between accentedness, comprehensibility and intelligibility in previous studies

Even though the working definitions of intelligibility and comprehensibility differ to a great extent, there have been experiments in which these two were interpreted as being the same. In a more recent study of Kirkpatrick, Deterding and Wong (2008), the main goal was to investigate the international intelligibility of Hong Kong English speakers but it actually addressed their comprehensibility as the assessing tool used was a comprehension test. Studies that have examined the correlation between intelligibility, comprehensibility and accentedness include those of Munro and Derwing (1995) Derwing and Munro (1997), Munro, Derwing and Morton (2006), Chen (2011) and Julkowska and Cebrian (2013).

The study of Munro and Derwing (1995), which examined the speech of non-native Mandarin speakers by native English listeners resulted in several significant conclusions. Firstly, it was found that the distribution of listeners' intelligibility scores resembled the distribution of comprehensibility scores rather than accent scores, and as much as 53% of the orthographic transcriptions received intelligibility score of 100%. Significant positive correlations between the perceived comprehensibility and accent ratings were found for all but 1 of the 18 listeners, while such a correlation between accent and intelligibility scores were present in only 5 of the 18 listeners' ratings. Moreover, highly intelligible stimuli were not necessarily rated as low accented the same way as highly accented stimuli did not necessarily receive lower intelligibility scores. These findings suggest that the three types of scores are only partially related, as most listeners' ratings showed correlation between comprehensibility and intelligibility and between comprehensibility and foreign accentedness. The examiners explain the lack of complete congruence between intelligibility and comprehensibility due to factors such as processing time or extra processing resources that had to be taken into account when rating stimuli for comprehensibility. Similarly, when judging accentedness, variables that caused the speech sound deviant had no impact on whether the message was fully understood. Last but not least, familiarity with an accent was also seen as a possible reason for the incongruence of the data. Even though the listener that have reported having regular contact with Mandarin accent, produced intelligibility rating below the mean. Other six people, who reported having fairly frequent contact with other accents produced orthographic transcriptions above the mean. This was one of the findings that later inspired examiners to look further into this phenomenon.

The research mentioned above was later extended in the work of Derwing and Munro (1999). The two experiments differed in some aspects. The 26 native English listeners of the later study were asked about their familiarity with four accents (Cantonese, Japanese, Polish and Spanish) that occurred in the stimuli and they were also prompted to identify the first language background of the talkers. Munro and Derwing (1995) worked with highly-proficient Mandarin speakers, whereas in the consequent study the speakers across all of the four categories showed intermediate level of English proficiency. This adjustment was made solely for the purpose of

comparison of the relationships between these two groups. The results of the study closely replicated those of Munro and Derwing (1995) in the sense that the three dimensions are clearly connected but not equivalent. The accent ratings were again highly skewed, but majority of them fell into the “heavily accented” category, while comprehensibility ratings were less harsh. Concerning intelligibility, the majority of transcriptions were at least 80% correct. Errors within the transcriptions were also categorized into “trivial” (regularizations such as “two mans” instead of “two men”) and “non-trivial” (word omissions and substitutions). While only 26% of the utterances were marked to be transcribed perfectly, this number increased to 60% after trivial errors in the transcriptions were excluded. Mean transcription scores by L1 did not differ to a great extent; Cantonese speakers received mean intelligibility score of 85%, Japanese 84%, Spanish 82% and Polish 81%. The identification of speaker’s L1 was carried out through a forced-choice task and the overall mean correct identification across listeners was 51.5%. The speakers of the two groups, Asian and Indo-European, were most often mistaken for each other, but no language was shown to be more difficult or easier to identify as familiarity with the specific language/accent was a significant predictor of identification. And similarly to Gass and Varonis (1984), there was also positive correlation between familiarity and intelligibility scores.

Another study that examined the relation between accentedness, comprehensibility and intelligibility was that of Julkowska and Cebrian (2013). Unlike in Munro and Derwing (1995) and Derwing and Munro (1999), in this study, foreign-accented (Polish) English was evaluated by both native and non-native English listeners. The outcome provided a support for Munro and Derwing’s (1995) and Derwing and Munro’s (1999) hypothesis that accentedness and intelligibility are rather independent of one another, as data from these two dimensions showed the weakest correlation. The strongest one, on the other hand, was found for intelligibility and comprehensibility ratings. And again, although most of the Polish-accented utterances was rated as moderately or heavily accented, in general they were fairly easy to understand and transcribe by majority of the listeners.

4.2 Intelligibility and first language background

The question of how native and non-native listeners' L1 background affects the intelligibility of foreign-accented English that was previously investigated by Gass and Varonis (1984), Smith and Rafiqzad (1979), Munro and Derwing (1995) and Derwing and Munro (1997) have also caught interest of Chen (2011).³

Chen used speech samples of Cantonese and Mandarin-accented English which were presented to five native and non-native listener groups; NS group, ESL group (Filipino and Pakistani), EFL group (Japanese and Korean), Mandarin group and Cantonese group. The method for assessing intelligibility chosen in this experiment was slightly different than the one in previous studies. The participants were presented with a stimuli containing 15 sentences, but instead of transcribing the whole sentences they were asked to fill in the words in the 70 embedded blanks. Chen (2011) claims that this assessment method “eliminates the less important function words, decreases listeners’ memory loading and enhances their motivation to spell the words” (67). Moreover, most of the blanked content words carried the important information of the passage and were therefore seen as the most significant ones for assessing intelligibility. All groups rated both Cantonese and Mandarin-accented speech at least 70% intelligible, while native English group achieved the highest intelligibility scores of 90% for the Cantonese and 97% for the Mandarin-accented samples. Four of the groups (NS, ESL, EFL and Mandarin) generated higher rates for the Mandarin-accented speech. The exception to this was Cantonese group, which rated their mother tongue as more intelligible than the Mandarin-accented English. In the case of this study, groups of listeners which shared the same L1 background with the speakers did not necessarily show a significant benefit while assessing intelligibility. The native English listener advantage set aside, neither Cantonese, nor Mandarin listeners’ ratings of their own accent were highest among the non-native listener groups, but both of these groups were much more comfortable with their own

³ Only these few are mentioned in connection with the topic above, as several others have also dealt with similar phenomenon, but focused on a particular hypothesis.

accent. The outcome of this experiment suggests that native listeners find non-native speech much more intelligible than non-native listeners do and although listeners from the same L1 background rated their own accent more intelligible, there was no significant correlation that would prove their ultimate advantage over other non-native listener groups.

The Mean Score of the Dictation Task (70 Items)

Group	Cantonese accent	SD	Corr. rate	Mandarin accent	SD	Corr. rate
NS (<i>n</i> = 5)	63.2/70	1.5	90%	68/70	1.2	97%
ESL (<i>n</i> = 10)	49.2/70	12.3	70%	51.1/70	10.6	73%
EFL (<i>n</i> = 4)	55.3/70	4.2	79%	62.5/70	2.6	89%
M (<i>n</i> = 5)	59.8/70	5.5	85%	61.6/70	4	88%
C (<i>n</i> = 5)	51.8/70	2.8	74%	48.8/70	2.6	70%

Note. NS = native speakers; ESL = Filipino + Pakistani; EFL = Japanese + Korean; M = Mandarin C = Cantonese.

Figure 1: Figure adapted from Chen (2011).

Chen’s (2011) experiment also included a foreign accent assessment task, but due to the fact that the accent was rated on a basis of speech properties like “honesty”, “smartness”, “appearance” and “education”, I did not find this set of data significant for the purpose of my work.

Smith and Rafiqzad (1979) designed an empirical study involving more than 1300 people from eleven countries in which they would “compare the degree of intelligibility between native and non-native varieties of educated English” (371). They used taped recordings of nine speakers from Hong Kong, India, Japan, Korea, Malaysia, Nepal, the Philippines, Sri Lanka and the United States which were presented to eleven groups of listeners, seven of which had a shared L1 background with the speakers. The examiners expected an equal intelligibility in the ratings of native speaker’s utterance and an utterance of a speaker with whom would the listener share an L1 background. This would then be followed by a score assigned to a speaker from a close geographical area. The results of the study did not provide support for these expectations though. Firstly, the native speakers’ intelligibility scores were always among the lowest, averaging 55%. Secondly, only two out of

seven non-native groups found their “fellow countrymen” (375) more intelligible than the other ones. This was the case of Japanese and Korean participants. The examiners themselves remarked, that the results might have been influenced by a great variability in the difficulty of individual passages, which later turned out to be the case. For this reason, their work has lost its significance within the second language research field.

The non-native listener advantage in rating other non-native speech samples for intelligibility was first mentioned by Nash (1969):

“A [non-native] speaker who cannot make himself understood when speaking English to a native English speaker will have no difficulty conversing in English with another [non-native] speaker” (4).

Van Wijngaarden (2001) examined this relationship while trying to quantify the loss of intelligibility during non-native (Dutch) and native English speaker interaction. Unlike the previous studies I have mentioned, this one used different assessment methods and focused on measuring speech intelligibility in noisy conditions. The speech intelligibility was measured on two levels; sentence intelligibility level and phoneme intelligibility level. The first one was measured through test method known as speech reception threshold (SRT), while the phoneme intelligibility test was related to the equally balanced CVC test. The highest sentence intelligibility appeared to show in the native speaker and native listener interaction. Although the results produced no significant support for the L2 speaker – L2 listener advantage, the most intelligible two (based on their ranking from native listeners) out of the four non-native talkers were more intelligible to other non-native listeners than to any of the four native English listeners.

Other studies examined the shared language background on the listener’s ability to recognize individual words (Imai et al 2003, Bradlow and Pisoni 1999). Both of these studies have worked with the notion of lexical neighbourhood density (ND) when measuring word intelligibility. Imai et al (2003) investigated the accuracy of word recognition between non-native (Spanish) and native English speakers and native English and non-native listeners who shared the (Spanish) L1 background. The

examiners expected the “phonological mismatch” (846), meaning the “different phonological representation of English words influenced by listener’s (Spanish) L1 background” (847) to significantly impair the ability of word recognition. They also hypothesized that this impair would be greater on “hard” words. Hard words would in this context mean low frequency words from a dense lexical neighbourhood (i.e. low frequency words that have many similar sounding “neighbors” and could therefore be easily confused with these).

Their “phonological mismatch” hypothesis was supported, as the native listeners recognized more words while listening to a native talker, but non-native listeners outperformed native listeners while listening to a non-native talker only for the words that came from dense lexical neighbourhood.

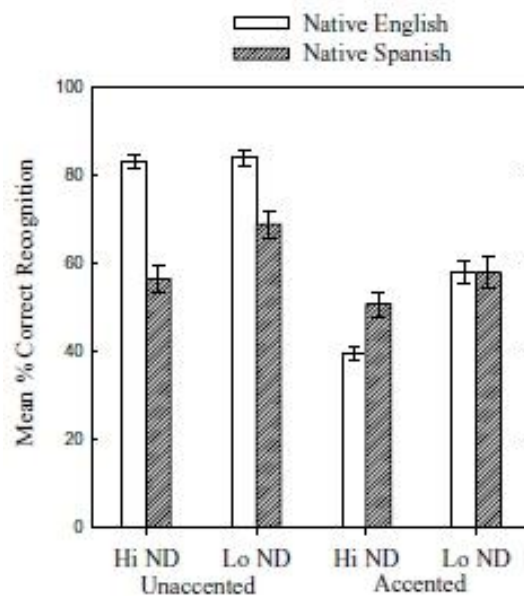


Figure 2: Figure adapted from Imai et al (2003).

These findings are consistent with those of Bradlow and Pisoni (1999). In their explanatory study, they have used a stimuli produced by ten native speakers of General American English and presented them to two different groups of listeners in two separate experiments. This study has also incorporated the notion of “easy” and “hard” words according to their lexical neighbourhood density, speaking rate and familiarity with a specific speech production as possible predictor of overall speech

intelligibility. In the first experiment, native listeners were recognizing words in native speech samples. It turned out that the overall intelligibility of this interaction was highly affected by lexical discriminability; easy words were more intelligible than hard words. Another significant factor for native word recognition was overall speech rate; faster utterances were less intelligible than the ones produced at a lower pace. These two factors, could however be overcome by listener's familiarity with the speech of the particular talker. Findings of the second experiment were more or less consistent with the previous one, but the effect of lexical neighbourhood density was much more prominent in the case of non-native listeners though.

5 The interlanguage speech intelligibility benefit

As I have mentioned in one of the previous chapters, there have been many studies which have dealt with the intelligibility of both foreign-accented and native speech. One of the studies in this SLA topic has examined this “nature of talker-listener interaction during speech communication” (Bent and Bradlow 2003) in much detailed way and proposed a hypothesis that would later on become a starting point for several others. Bent and Bradlow (2003) have focused on “the phenomenon of native and target language interaction during the acquisition of second language sound structure at the level of overall sentence intelligibility” (1600) and through their findings wanted to demonstrate, that “any measure of overall speech intelligibility must take into account both talker- and listener-related factors”(1600).

Even though there have been studies before Bent and Bradlow (2003) that aimed to measure the intelligibility of non-native interaction, all of these but one⁴ have worked with non-native participants that shared their first language background. Most of these studies reported that for non-native listener, the speech of fellow non-native talker is easier to understand than the speech of a native talker. Bent and Bradlow (2003) made this finding one of their starting points for further hypothesis:

Since individuals from the same native language background who are in the process of acquiring a given target language all share an “interlanguage”, we predicted that target language intelligibility between non-native talkers and listeners from the same native language background would be enhanced relative to intelligibility between a native talker and non-native listener (1601).

The reason Bent and Bradlow also included talkers and listeners from different native language background was to examine whether the benefit which appears in case of a shared interlanguage between a non-native talker and listener is “entirely dependent on the talker and listener sharing the native language or if non-native speech is

⁴ See the study of Smith and Rafiqzad in chapter 4.2

generally more intelligible than native speech for all non-native listeners regardless of native language background” (1601).

Besides the fact they worked with the notion of an interlanguage, they also acknowledged that the first language has a significant impact on the perception and production of non-native contrasts and the possible difficulties non-native listeners encounter when recognizing non-native speech are related to “the relationship between the status of sounds in the overall systems of phonemic contrasts in the first and second language” (1601).

5.1 Bent and Bradlow 2003

The aim of the experiment was to investigate how the previously mentioned talker-listener match or mismatch with respect to their first language, but also their second language proficiency affects the degree of non-native talker intelligibility. The whole study is based on two main predictions, both of which assume that high proficiency non-native talker, either from the same or from different L1 background will be at least as intelligible to other non-native listeners as a native speaker.

...a non-native talker with relatively high proficiency in English speech production would be at least as intelligible as a native talker for nonnative listeners from the same native language background (1602).

This first prediction is based on the notion of shared linguistic and phonetic knowledge among non-native learners from the same native language background. However, it also suggests the loss of overall intelligibility in case of less-proficient non-native talkers for both native and non-native listeners.

“a relatively high proficiency nonnative talker will also be at least as intelligible as a native talker for non-native listeners from a different native language background”(1602). The second prediction assumes that the absence of some reduction phenomena, such as alveolar flapping or unreleased stops might have a beneficial effect on non-native speech recognition for any non-native listener.

5.1.1 Participants and methodology

Bent and Bradlow (2003) have used the Northwestern University Foreign Accented English Speech Database (NUFAESD) as the only source of the speech recordings for their experiment. The database included talkers from the Northwestern University International Summer Institute and English as a Second Language program and demonstrated a high level of proficiency with written English, but their experience with spoken English communication was limited. Apart from the speech recording itself, the database also included demographic information about each talker and an overall intelligibility score of his production as rated in a perception test with native English listeners. Recordings were made in a sound-attenuated booth and the stimuli contained sentences with highly familiar and syntactically simple words to non-native listeners. “The digitized recordings were embedded in white noise, yielding a speech-plus-noise file with a +5dB signal-to-noise ratio” (1603).

Five talkers participating in the experiment came from three different language backgrounds; one from English, two from Chinese and two from Korean. Four of the non-native talkers were selected from the NUFAESD database, based on their first language, gender and the intelligibility score assigned for the purpose of the database compilation. The level of English proficiency was also a decisive factor as one in the pair of talkers was, in both cases, rated to be highly proficient while the other showed a low level of English proficiency. The attempt to select talkers with a complete match in both levels of proficiency was not successful. The only monolingual English talker was recorded the same way as the participants of the NFUAESD recording session. The total number of 65 listeners were divided into four groups according to their language background, three of which matched the language backgrounds of the talkers (Korean, Chinese and English) and one contained a mixture of non-native listeners from various L1 backgrounds (Bulgarian, Dutch, French, German, Greek, Hindi, Japanese, Serbian, Spanish and Tamil). The listeners were also recruited from the same NWU program. The table below shows additional variables concerning non-native listeners’ experience with L2.

General information about non-native listener groups. Mean and standard deviations (in parentheses) are shown. The Northwestern University Graduate School minimum for TOEFL is 560 (paper) or 220 (computer).

	NN-Chinese (n = 21)	NN-Korean (n = 10)	NN-Mixed (N = 12)	Significant differences
Age (years)	23.60 (1.60)	28.00 (4.72)	25.18 (2.35)	Korean > Chinese ^a Mixed > Chinese ^b
TOEFL ^c — computer	n/a	(n = 1) 253	(n = 8) 269.63 (22.06)	n/a
TOEFL— paper	(n = 17) 641.82 (19.63)	(n = 7) 602.43 (18.80)	(n = 4) 629.25 (37.25)	Chinese > Korean ^a
Time in US (months)	1.12 (1.92)	8.25 (9.48)	5.23 (10.34)	Korean > Chinese ^a
AOA ^d (years)	10.95 (2.22)	12.25 (1.39)	10.00 (4.22)	none
LOA ^e (years)	10.11 (2.26)	7.00 (1.41)	8.17 (4.37)	Chinese > Korean ^a

^a = $p < 0.01$.

^b = $p < 0.05$.

^cTOEFL = Test of English as a Foreign Language.

^dAOA = age of acquisition.

^eLOA = length of acquisition (i.e., English study).

Figure 3: Figure adapted from Bent and Bradlow (2003).

The intelligibility assessment was carried out through a sentence transcription. Each sentence contained three to four key words. The final intelligibility score was assigned according to a number of key words transcribed correctly. Morpheme omission and addition were considered errors while obvious spelling mistakes were not. The score assigned by each listener was converted to percent correct and then to rationalized arcsine units (rau). A separate word familiarity test was administered for the non-native listeners after the speech perception test was over. The subjects were presented with all the keywords from the sets of sentences and also with another set of 75 filler items to ensure that the full range of the 7-point familiarity scale was represented. The items were presented in random order.

5.1.2 Results and conclusion

All of the listeners gave a rating of 5 or more to 94% of the items, while 79% of the words received a highest familiarity rating of 7. The ratings of the filler items were consistent with those assigned by native listeners in Bradlow and Pisoni (1999), the non-native listeners' familiarity ratings were therefore considered very reliable. Furthermore, no correlation emerged between the listeners' average familiarity scores and their speech perception scores. Based on the outcome of this test, experimenters could assume that all of the non-native listeners were sufficiently familiar with the keywords and that they could perceive sentences in noise regardless of word familiarity.

Percent keywords correctly transcribed for talkers and listener groups in rau. Standard deviations are shown in parentheses. Scores in bold are significantly higher than the other scores in the row and are not significantly different from one another except for the underlined scores. Specifically, for the NN-Chinese listeners, the high-proficiency Korean talker is significantly more intelligible than the native English, and for the NN-Korean listeners the high-proficiency Korean talker is significantly more intelligible than the low proficiency Korean talker.

Listener group	Talker					All talkers
	Chinese high proficiency	Korean high proficiency	Native English	Chinese low proficiency	Korean low proficiency	
NN-Chinese (n = 21)	64 (10.8)	66 (11.7)	<u>56</u> (10.4)	30 (12.1)	41 (9.8)	51 (17.4)
NN-Korean (n = 10)	60 (15.5)	74 (15.8)	60 (11.7)	22 (11.4)	53 (12.0)	54 (21.7)
NN-Mixed (n = 12)	62 (11.0)	70 (7.7)	67 (15.8)	19 (20.9)	41 (14.8)	52 (24.1)
Native English (n = 21)	77 (12.2)	91 (8.4)	109 (14.7)	38 (13.8)	60 (12.1)	75 (27.7)
All listeners	67 (13.9)	76 (15.2)	76 (26.8)	29 (15.9)	49 (14.5)	

Figure 4: Figure adapted from Bent and Bradlow (2003).

Three main conclusions could be drawn upon the sentence recognition data presented in the table above:

- 1) For native listeners, intelligibility of the native talker was greater than the intelligibility of any of the non-native talkers.
- 2) For non-native listeners, intelligibility of a high proficiency non-native talker and in one case a low proficiency talker from the same native language

background was equal to the intelligibility of the native talker. This is the “matched interlanguage speech intelligibility benefit.”

- 3) For non-native listeners, intelligibility of a high proficiency non-native talker from a different native language background was greater than or equal to the intelligibility of the native talker. This is the “mismatched interlanguage speech intelligibility benefit.” (Bent and Bradlow 2003)

Both of Bent and Bradlow’s (2003) predictions were therefore confirmed. The “matched interlanguage speech intelligibility benefit” occurs because “non-native speech production and perception are both systematically linked to native language sound structure. (1607). This means that the shared linguistic knowledge of non-native speakers from the same native language background is much more extensive than between any other speakers, as it covers (for example phonetic and phonological) aspects of both the native and the target languages. Due to the language transfer between their native and second language, the native-language-matched non-native listener is able to interpret certain acoustic-phonetic features of the non-native speech, even though they may deviate markedly from the target language norm (1607). The “mismatched interlanguage speech intelligibility benefit” observed here is a clear support to the findings of Smith, Bradlow and Bent (2003) where non-native listeners from various backgrounds were better at identifying words in minimal pairs. This might be due to a general strategies non-native learners apply when acquiring a new language. In the case of recognizing minimal word pairs, non-native listeners produced word-final stop consonant releases which served as possible clues for other non-natives and to which native listeners did not attend.

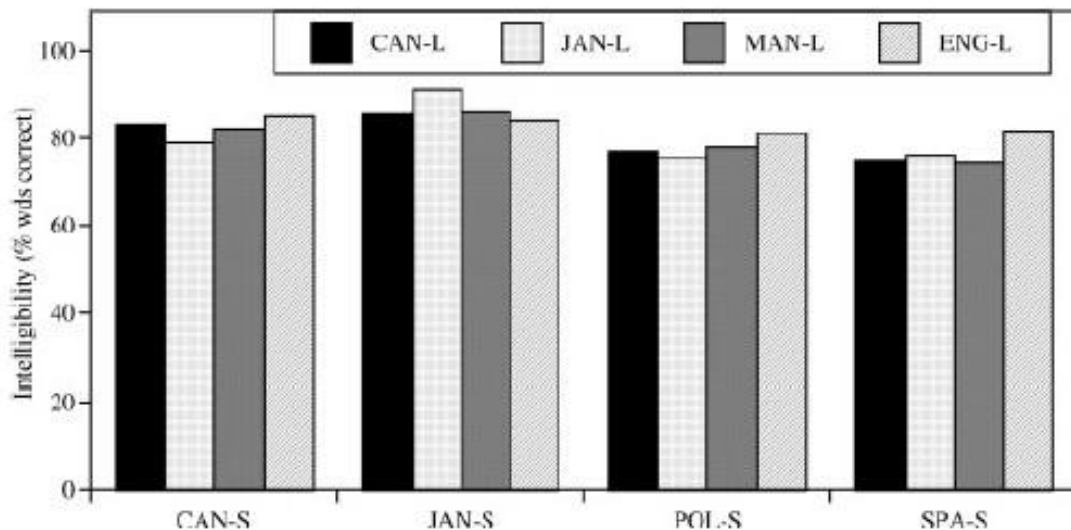
5.2 Studies following Bent and Bradlow (2003)

Munro and Derwing have already conducted several studies concerning L2 speech intelligibility, comprehensibility and accentedness. Their most recent one (Munro, Derwing and Morton, 2006) was again an extension of their previous research (Derwing and Munro 1997) on L2 speech intelligibility, comprehensibility and

accentedness. While Munro and Derwing (1995) worked with native English subjects listening to both native and non-native (Mandarin) speakers, Derwing and Munro (1997) included speakers from four different language background, while retaining the native English listener group as the only one. Their last paper (Munro, Derwing and Morton 2006) was focused on the shared speaker – listener L1 background advantage (i.e. the interlanguage speech intelligibility benefit), as one of their research question states as follows:

To what extent do nonnative English listeners from different L1 backgrounds share a response to L2 speech when they assess it for intelligibility, comprehensibility and accentedness? (115)

They used the same speech samples of Cantonese, Japanese, Polish and Spanish speakers as in the preceding study, but presented them to three different groups of listeners. The Cantonese and Japanese group was included due to a shared L1 background with the two of the speaker groups and the Mandarin group represented no such shared L1 background with any of the speaker groups. The evaluations from native English group were also taken from Munro and Derwing (1997). The interrater reliability was measured for all groups on all measures and showed that “the members of each group tended to agree with one another on the relative intelligibility, comprehensibility and accentedness of the speakers” (119). Concerning the shared L1 intelligibility benefit, although an advantage was found in the case of Japanese listeners’ intelligibility scores, no such parallel advantage was found for Cantonese speaker-listener interaction.



Mean intelligibility scores (percent words correct) for each group of speakers. Each cluster shows the scores of the four listener groups for one speaker group.

Figure 5: Figure adapted from Munro, Derwing and Morton (2006).

This outcome provides support for the claim, that “L2 listeners sometimes-but not always-understand speech better when it is produced with their own accent than with another accent” (125). The shared interlanguage speech intelligibility benefit hypothesis could therefore not be confirmed. In this study of non-degraded speech perception⁵, the properties of the speech itself were “a potent determinant of the listeners’ responses and despite listeners’ disparate linguistic backgrounds, they tended to share a response to the speech” (125).

Another disfavoured view of Bent and Bradlow’s (2003) hypothesis and the very last study to be mentioned appeared in Stibbard and Lee (2006). The aim was to deconstruct the “mismatched speech intelligibility benefit”, as it appeared in Bent and Bradlow’s (2003) study of Korean and Chinese subjects. Stibbard and Lee (2006) suspect this phenomena being a mere effect of similarity between the two languages, leading to between-group familiarity with each other’s accent (434). For this reason, the language groups included in their study (Arabic and Korean) were chosen to differ in many phonological aspects. The term “benefit” was also re-defined to

⁵ As opposed to Bent and Bradlow’s (2003) speech-in-noise perception study

“exclude cases of equal intelligibility scores and include only those cases where higher intelligibility scores were given” (434). The intelligibility of the speech stimuli was assessed through a key word-recognition task which was followed by a word-familiarity test. As predicted, the results have shown no evidence for the “mismatched interlanguage speech intelligibility benefit”. The evidence against this hypothesis was only found in the case of low-proficiency L2 speaker-listener interaction and a notion of “mismatched interlanguage speech intelligibility detriment” was introduced (440).

“...non-native listeners found the low-proficiency non-native talker with whom they did not share the same first language the most difficult of all the talkers to understand. This provides evidence for a mismatched interlanguage speech intelligibility *detriment* in those cases involving the low-proficiency non-native talkers, and indicates that low-proficiency learners may find difficulty in being understood either by non-native listeners who do not share their first language or by native English listeners.” (440)

6 Study proposal

In this part of my work I will design methodology for a laboratory controlled experiment, investigating the intelligibility of foreign-accented speech on the basis of talker-listener interaction. Bent and Bradlow's (2003) "interlanguage speech intelligibility benefit" hypothesis will be the starting point of the investigation. The study itself will be designed in order to either confirm or reject the proposed "matched" or "mismatched" non-native talker-listener advantage and to explore its possible extensions with respect to variations in participants' L1 backgrounds. These variations will include the combination of geographical closeness and language family group. The work will also aim to eliminate an over-extensive range of foreign accent among the speakers and accent familiarity with L2 speakers' accents as possible precursors of degree of L2 intelligibility.

6.1 Introduction

Outcomes of various studies have confirmed that non-native listeners find non-native speech more intelligible than native speech, while the opposite applies to native listeners. Since the publication of Bent and Bradlow (2003) there has been an offspring of SLA studies investigating how L1 background affects the intelligibility and overall perception of both native and non-native speech. While there have been studies which did agree with the "matched interlanguage speech intelligibility benefit", the notion of its counterpart, the "mismatched interlanguage speech intelligibility benefit" has either not been confirmed or it was completely rejected (Stibbard and Lee 2006).

The present study will focus on both of these propositions and will also be designed in order to find evidence for its extension. So far there have been no experiments which would investigate the variable of geographical closeness, or a language family group as a potential determinant of the "mismatched" advantage. In other words, no attempt has been made to break the boundaries of these two phenomena.

The following two conclusions of Bent and Bradlow's (2003) study will serve as hypotheses for the current experiment:

- 1) For non-native listeners, intelligibility of a high proficiency non-native talker and talker from the same native language background was equal to the intelligibility of the native talker. This is the "matched interlanguage speech intelligibility benefit." (1607)
- 2) For non-native listeners, intelligibility of a high proficiency non-native talker from a different native language background was greater than or equal to the intelligibility of the native talker. This is the "mismatched interlanguage speech intelligibility benefit." (1607)

In addition to these predictions, the shared language family group or geographical closeness is expected to affect the degree of non-native speech intelligibility in the following way:

- 3) For non-native listeners, intelligibility of a high proficiency non-native talker from the same native language family will be lower than intelligibility of a non-native speaker with a shared native language, but higher than intelligibility of non-native speaker from different native language background and family.

In other words, a kind of intelligibility downgrade should be observed respectively to a decreasing native language and family closeness. Geographical closeness should be observed to affect the intelligibility of non-native speech in a similar way. Combining these two variables will also allow to examine a possible correlation between them.

There is no prediction concerning intelligibility of native speaker in terms of the last hypothesis.

6.2 Methodology

6.2.1 Speakers

Total number of five speaker groups will be included in the study, one native and four non-native groups. All non-native participants will be highly proficient⁶ late L2 learners of English (based on a Cambridge exam or TOEFL score) with no reported stay in an L2 speaking environment and no received instruction in English phonology. The four non-native groups will be Czech, Ukrainian, Romanian and Tibetan. The rationale behind this is that in order to examine the language family-intelligibility relation, European (Slavic), European (non-Slavic) and a distant Sino-Tibetan language families have to be represented. Each of these language families represent a different value of closeness to the family of the listeners. The representative group of the Slavic language family is also designed to be geographically closer to the listener group than the non-Slavic language family. In this way, the study will aim to combine both the geographical closeness and language family prediction, which has not been attempted before. A listener-speaker accent familiarity test will also be included in order to ensure that the majority of listeners have no or minimal experience with any of the proposed foreign accents. The range of the number of speakers varied widely across experiments, from 2 (Chen 2011) to 5 (Bent and Bradlow 2003) and 48 (Munro and Derwing 1995,1999). The current experiment will include 4 speakers in each group, 2 male and 2 female representatives, giving a total number of 20 speakers in the whole study.

Previous research has shown that accentedness and intelligibility are two partially independent dimensions and out of the three dimensions studied (comprehensibility, accentedness and intelligibility), the accent-intelligibility correlations always turned out to be the weakest. This means that even heavily accented speech may be perfectly intelligible.⁷ Nevertheless, a foreign accent evaluation task will be carried out prior to the stimuli recording in order to provide equally accented speech samples, eliminating foreign accent as a possible determiner

⁶ The language proficiency will not be examined here as the possible determinant of speech intelligibility.

⁷ See Chapter 4.1 for references

of intelligibility. Final 30 speakers will then be chosen based on this accent evaluation data. A short spontaneous dialogue with a native speaker will be designed and individual speakers will be prompted to react, answer simple questions or contribute to the conversation. Three native English observers will be present to assess the degree of foreign accent on a 7-point Likert scale. Mean accent scores will be computed for each subject and speakers with accent ratings within the middle range (3 to 5) will be marked as suitable for the purposes of the current study.

Even though Southwood and Flege (1999) proposed at least 9-point scale to be the most suitable one, due to the number of native evaluators, it would be difficult to select enough speakers with desired mean accent rating.

6.2.2 Listeners

All hypotheses are designed upon the non-native speaker perspective and this will therefore be the only one taken into account. The number of listeners is usually higher than the one of speakers, and for that reason, 45 Czech listeners will participate in the study. All of these will be highly proficient late learners of English with no received training in phonology and no reported familiarity with any of the other foreign accents. Czech is an European Slavic language and will be shared with the Ukrainian group, the partial language family mismatch will occur with the Romanian group and complete language family mismatch will be present in relation to the Tibetan group of speakers. The shared Czech L1 background between the speaker-listener groups allows examination of the “matched interlanguage speech intelligibility benefit”. Listeners will be divided into two groups with respect to their L2 experience. The criterion for “experienced” will in this case be a reported stay in an L2 (English) speaking country of at least one month. The use of unsophisticated speakers and listeners is important due to the fact that it allows a valuable insight into how intelligible an L2 speaker-listener interaction is in commonly occurring situations. Moreover, phonetically trained evaluators may not respond to L2 speech in the same way as unsophisticated listeners.

6.2.3 Stimuli

The stimuli collection procedures have varied across intelligibility studies. Lane (1963) used stimuli containing individual words, Smith and Rafiqzad (1979) and Smith and Bisazza (1982) used passages, sentences and paragraphs as their speech material similarly to Brodkey (1972), who introduced the today commonly known method of dictation task. The studies of Munro, Derwing and Morton are known to be using a speech stimuli of either spontaneous (Munro and Derwing 1995, 1999) or slightly premeditated cartoon description. Some studies have worked with a specially designed pronunciation databases. Bent and Bradlow (2003) made a selection of four BKB-R (Bamford-Kowal-Bench Sentence Test) lists, which were part of the NUFAESD (The Northwestern University Foreign Accented English Speech Database) and Chen (2011) used asset of sentences from a pronunciation teaching reference book.

For the purpose of this study, a set of sentences similar to those of Bent and Bradlow (2003) will be used. Such sentences are syntactically simple and most of the words are reported to be highly familiar to non-native speakers. All of the participants of the study are highly proficient English learners and therefore should have no problems with comprehension and reproduction of these simplified utterances. The length of each stimulus will be controlled to avoid length variability being a possible influence on listeners' judgements. Each talker will be represented by three sentences, giving a total number of 60 sentences in the whole stimuli.

List 3

- (1) The book tells a story.
- (2) The young boy left home.
- (3) They are climbing the tree.
- (4) She stood near her window.
- (5) The table has three legs.
- (6) A letter fell on the floor.
- (7) The five men are working.
- (8) The shoes were very dirty.
- (9) They went on a vacation.
- (10) The baby broke his cup.
- (11) The lady packed her bag.
- (12) The dinner plate is hot.

List 4

- (1) A dish towel is by the sink.
- (2) She looked in her mirror.
- (3) The good boy is helping.
- (4) They followed the path.
- (5) The kitchen clock was wrong.
- (6) Someone is crossing the road.
- (7) The mailman brought a letter.
- (8) They are riding their bicycles.
- (9) He broke his leg.
- (10) The milk was by the front door.
- (11) The shirts are hanging in the closet.
- (12) The chicken laid some eggs.

Figure 6: Figure adapted from Bent and Bradlow (2003).

Recordings will be carried out in a sound-treated room. Sentences will be presented individually on a computer screen. The digital speech recordings will then be segmented into sentence-length files with added 300ms of silence at the beginning of each.

6.2.4 Procedure

Intelligibility assessment methods vary from experiment to experiment. Some employed picture selection in response to the stimulus (Smith and Bisazza 1982) and Perlmutter (1989) asked the listeners to summarize the idea of L2 speakers' short presentations. However, Munro (2008) claims that if intelligibility is defined as "the extent to which a speaker's utterance is actually understood" (Munro and Derwing 2006, 112), it is necessary to distinguish between the message intended and message received.

“...the researcher can be certain of the content of the intended message only if that content is pre-determined, as in tasks in which the L2 speaker reads or repeats words, sentences or longer texts (Munro 2008, 202).

The tasks which Munro (2008) is referring to are the now commonly used transcription tasks, but these have been carried out in several different ways. The first type of the transcription task relied on transcription of whole utterances (Munro and Derwing 1995, 1999). In another kind of transcription task, listeners were asked to fill in the words in the embedded blanks (Chen 2011). Bent and Bradlow (2003) prompted the listeners to transcribe the whole utterances, but perception scores were determined by strict keyword-correct count. In the present study, the whole-sentence transcription method will be employed.

Listeners will be seated in front of a computer and will listen to all 60 speech recordings through a high-quality headphones. After hearing an utterance, they will transcribe exactly what they heard in standard orthography on a specially prepared answer sheet. Each sentence will be played only once and after the transcription is completed, the listener will press a key on a computer keyboard to elicit the next trial. After the perception test, a separate word familiarity session will be held.

6.2.5 Further suggestions

The current study aims at evaluating native and non-native speech intelligibility in laboratory controlled conditions, using noise-free speech samples. A good extension would be the use of samples with added noise in order to compare how non-native listener speech perception changes in adverse conditions.

7 Conclusion

Investigation of relevant literature and previous studies were crucial for designing an experiment that would either support or reject the theory of “matched and mismatched interlanguage speech intelligibility benefit.”

Due to several extensions of the original Bent and Bradlow study included, the present proposal gives a reasonable amount of space for the investigation of possible correlations that would occur in the case of “matched interlanguage speech intelligibility benefit”. Among these are the correlations between intelligibility of L2 and geographical closeness of the L1 language family of the speaker and listener and the combination of these two thereof. The accuracy of the outcome should also be increased compared to previous experiments due to elimination of some possible intelligibility determiners.

8 Shrnutí

Bakalářská práce se zabývá tématem cizího přízvuku, srozumitelnosti řeči s cizím přízvukem a teoriemi o vlivu rodného jazyka posluchače na hodnocení míry srozumitelnosti. Stěžejní teorií je „matched and mismatched interlanguage speech intelligibility benefit“, jež byla formulována profesorkami Tessou Bent a Ann Bradlow. Na základě shrnutí dosud publikované literatury má práce za cíl navrhnout experiment, který by tuto teorii potvrdil či vyvrátil.

První část se pokrývá problematiku pojmu Interlanguage, který je součástí stěžejní hypotézy. Následuje souhrn literatury a studií zabývajících se původem a hodnocením cizího přízvuku jakožto možného prediktoru míry srozumitelnosti. Další kapitola je věnována samotné srozumitelnosti; ta je zde definována a na základě předešlých experimentů je prozkoumán její vztah k přízvuku a jeho dalším aspektům. Ve zvláštní podkapitole je popsán vztah srozumitelnosti k rodnému jazyku mluvčího a posluchače. Ve druhé polovině je detailně popsán experiment Bent a Bradlow; zvláštní pozornost je věnována jejich hypotézám a metodologii, která je posléze částečně implikována do vlastní studie. Závěrečnou částí práce je návrh vlastního experimentu, jehož výsledek má potvrdit nebo vyvrátit hypotézu „matched and mismatched interlanguage speech intelligibility benefit“. Metodologie je navržena tak, aby bylo možno zkoumat tuto hypotézu v závislosti na dosud nezkoumaných charakteristikách mluvčích a posluchačů. Pro zvýšení věrohodnosti výsledku byly při návrhu metodologie eliminovány faktory, jež byly v ostatních experimentech při hodnocení míry srozumitelnosti označeny jako nežádoucí.

9 References

- Abercrombie, D. 1949. "Teaching pronunciation." *English Language Teaching*, 3.1: 113–122.
- Asher, J., and Garcia, G. 1969. "The optimal age to learn a foreign language." *Modern Language Journal* 38.1: 334-341.
- Behney, Jennifer, Susan M Gass, and Luke Plonsky. *Second Language Acquisition: An Introductory Course*. 4th ed. New York ; London: Routledge, 2013.
- Bent, Tessa, and Ann R. Bradlow. 2003. "The interlanguage speech intelligibility benefit." *The Journal of the Acoustical Society of America* 114: 1600-1610.
- Bradlow, A. R. & Pisoni, D.B. 1999. "Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors." *Journal of the Acoustical Society of America* 106.1: 2074-2085.
- Brodkey, D. 1972. "Dictation as a Measure of Mutual Intelligibility: A Pilot Study." *Language Learning* 22.1: 203-217.
- Chen, H. C. 2011. "Judgments of intelligibility and foreign accent by listeners of different language backgrounds." *The Journal of Asia TEFL* 8.4: 61-83.
- Flege, J. E. 1988. "Factors affecting degree of perceived foreign accent in English sentences." *The Journal of the Acoustical Society of America* 84.1: 70-79.
- Flege, J. E. & Fletcher, K. L. (1992) "Talker and listener effects on degree of perceived foreign accent." *Journal of the Acoustical Society of America*, 91.1: 370-389.
- Flege, J. E., Munro, M. & MacKay, I. (1995). "The effect of age of second language learning on the production of English consonants." *Speech Communication* 16.1: 1-26.
- Flege, J. E. et al. (1997). "Effects of experience on non-native speakers' production and perception of English vowels". *Journal of Phonetics* 25, 437-470.
- Gass, S. & Varonis, E. 1984. "The effect of familiarity on the comprehensibility of nonnative speech." *Language Learning* 34.1: 65-89.

- Greene, J., & Wells, E. 1927. *The cause and cure of speech disorders: A textbook for students and teachers on stuttering, stammering and voice conditions*. New York: The Macmillan Company.
- Griffen, T. 1980. "A nonsegmental approach to the teaching of pronunciation". *Revue de Phonétique Appliquée* 54: 81–94.
- Hurford, James R. 1991. "The evolution of the critical period for language acquisition." *Cognition* 40.3: 159-201.
- Imai, S., Flege, J.E., & Walley, A. 2003. "Spoken word recognition of accented and unaccented speech: Lexical factors affecting native and non-native listeners." *Proceedings of the International Congress on Phonetic Science* 15.1: 845-848.
- Jułkowska, I. A., & Cebrian, J. 2013. "Intelligibility, comprehensibility and foreign accentedness in the speech of Polish L2 English speakers." *Proceedings of the Third International Conference on English Pronunciation: Issues & Practices (EPIP3)*: 45-48. Murcia, Spain: University of Murcia.
- Kirkpatrick, A., Deterding, D., & Wong, P. M. J. 2008. The international intelligibility of Hong Kong English. *World Englishes*, 27.3-4: 359-377.
- Krashen, S. D., Scarcella, R. C., & Long, M. H. 1979. "Age and rate and eventual attainment in second language acquisition." *TESOL Quarterly*, 13.1: 573-78.
- Lane H. 1963. "Foreign accent and speech distortion." *Journal of the Acoustical Society of America* 35.1: 451–453
- Lenneberg, Eric H., Otto Marx, and Noam Chomsky. n.d.1967. *Biological foundations of language / Erich H. Lenneberg ; with appendices by Noam Chomsky and Otto Marx*. New York : John Wiley and Sons.
- Long, Michael H. 1990. "Maturational Constraints on Language Development." *Studies In Second Language Acquisition* 12.3: 251-285.
- Marshall, Gordon. "equal appearing intervals." A Dictionary of Sociology. 1998. *Encyclopedia.com*. (January 22, 2014).
- Moyer, A. 1999. "Ultimate attainment in L2 phonology The critical factors of age, motivation and instruction." *Studies in Second language Acquisition* 21.1: 81-108.

- Munro, Murray J., and Tracey M. Derwing. 1995. "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners." *Language Learning* 45.1: 73-97, reprinted as Munro, Murray J., and Tracey M. Derwing 1999. "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners." *Language Learning* 49.1: 285-310.
- Munro, Murray J. 1998. "The effects of noise on the intelligibility of foreign-accented speech." *Studies in Second Language Acquisition* 20.1: 139-154.
- Munro, Murray J., Tracey M. Derwing, and Susan L. Morton. 2006. "The mutual intelligibility of L2 speech." *Studies in Second Language Acquisition* 28.1: 111-131.
- Munro, M. J. 2008. Foreign accent and speech intelligibility. In J. G. Hansen Edwards & M. L. Zampini (eds.), *Phonology and second language acquisition*, Philadelphia: John Benjamins.
- Nash, Rose. 1969. "Intonational Interference in the Speech of Puerto Rican Bilinguals: An Instrumental Study Based on Oral Readings of a Juan Bobo Story." *J Engl Second Lang* 4.1: 1-42.
- Nelson, C. 1982. "Intelligibility and non-native varieties of English." In B. B. Kachru (Ed.). *The other tongue: English across cultures*: 58-73. University of Illinois Press.
- Neufeld, G. 1979. "Towards a theory of language learning ability." *Language Learning* 29.2: 227-241.
- Oyama, Susan. 1976. "A sensitive period for the acquisition of a nonnative phonological system." *Journal of Psycholinguistic Research* 5.3: 261-83.
- Patkowski, Mark. 1980. "The sensitive period for the acquisition of syntax in a second language." *Language Learning* 30.1: 449-472.
- Patkowski, Mark. 1990. "Age and Accent in a Second Language: A Reply to James Emil Flege." *Applied Linguistics* 11.1: 73-89.
- Penfield, Wilder, and Lamar Roberts. 1959. *Speech and brain mechanisms*. Princeton, NJ, US: Princeton University Press.

- Piper, Terry and Dilek Cansin. 1988. "Factors Influencing The Foreign Accent." *The Canadian Modern Language Review*, 44.2: 334-342.
- Piske, T., MacKay, I. R. A. & Flege, J. E. 2001. "Factors affecting degree of foreign accent in an L2: a review." *Journal of Phonetics* 29.1: 191-215.
- Perlmutter, M. 1989. "Intelligibility rating of L2 speech pre- and post-intervention." *Perceptual and Motor Skills* 68.1: 515-521.
- Scovel, Thomas. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. New York: Newbury House and Harper and Row.
- Selinker, Larry. 1972. "INTERLANGUAGE." *International Review Of Applied Linguistics In Language Teaching (IRAL)* 10. 1-4: 209-232.
- Smith, Larry E., and John A. Bisazza. 1982. "The comprehensibility of three varieties of English for college students in seven countries." *Language Learning* 32.2: 259-269.
- Smith, L. E., and K. Rafiqzad. 1979. "English for cross-cultural communication: the question of intelligibility." *TESOL Quarterly* 13.3: 371-80.
- Smith, B., Bradlow, A. R., and Bent, T. (2003). "Production and perception of temporal contrast in foreign accented English." In M. J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth International Congress of Phonetic Sciences*, Barcelona, Spain, 519-522.
- Snow, Catherine E., and Hoefnagel-Hohle, Marian. 1978. "The Critical Period for Language Acquisition: Evidence from Second Language Learning." *Child Development* 49.1: 1114-1128.
- Southwood, M.H., and J.E. Flege. 1999. "Scaling foreign accent: Direct magnitude estimation versus interval scaling." *Clinical Linguistics And Phonetics* 13.55: 335-349.
- Stibbard, R. M. & J. I. Lee. 2006. "Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis." *Journal of the Acoustical Society of America* 120.1: 433-442.

- Tahta, S., Wood, M., and Lowenthal, K. 1981. "Foreign Accents: Factors Relating to Transfer of Accent from the first Language to a Second Language." *Language and Speech*, 24.1: 265-272.
- van Wijngaarden, Sander J. 2001. "Intelligibility of native and non-native Dutch speech." *Speech communication* 35.1: 103-113.
- van Wijngaarden, Sander J., Herman J. M Steeneken, and Tammo Houtgast. 2002. "Quantifying the intelligibility of speech in noise for non-native listeners." *The Journal of the Acoustical Society of America* 111: 1906-1916.
- Weismer, G. & Laures, J. S. 2002. "Direct magnitude estimation of speech intelligibility: Effects of a chosen standard." *Journal of Speech, Language, and Hearing Research* 45.1: 421-433.

10 Anotace

Příjmení a jméno	Vybíralová Petra
Katedra a fakulta:	Katedra anglistiky a amerikanistiky, Filozofická fakulta
Název česky:	Srozumitelnost cizích přízvuků v angličtině
Název anglicky	Intelligibility of foreign-accented English
Vedoucí práce:	Mgr. Šárka Šimáčková, PhD.
Počet stran:	44
Přílohy:	CD
Klíčová slova:	osvojování cizího jazyka, cizí přízvuk, srozumitelnost, míra cizího přízvuku
Klíčová slova v AJ:	second language acquisition, foreign accent, intelligibility, comprehensibility, accentedness, interlanguage speech intelligibility benefit
Anotace:	Tato bakalářská práce se zabývá tématem cizího přízvuku a srozumitelností angličtiny s cizím přízvukem. Jejím záměrem je poskytnout komplexní souhrn příslušné literatury, zvláště pak té se zaměřením na „interlanguage speech intelligibility benefit.” Má také za cíl shrnout předešlé postoje a výsledky studií zabývajících se tímto jevem a navrhnout metodologii pro experiment v této oblasti.

Anotace v AJ:

The present thesis deals with the notion of foreign accent in and intelligibility of foreign accented English. It aims to give a comprehensive overview of relevant literature with a particular focus on the notion of “interlanguage speech intelligibility benefit”, which appeared in an experiment of Tessa Bent and Bradlow. The main goal of the work is to review previous approaches to this phenomenon as well as varying outcomes of different studies and to provide a reliable methodological base for a future research in this field.