UNIVERSITAT DE BARCELONA

Master's Thesis

Interlanguage Speech Intelligibility Benefit for Non-Native Listeners of English

Submitted by
Anja Ludwig

Applied Linguistics and Language Acquisition in Multilingual Contexts (LAALCM)

Universtitat de Barcelona

Supervisor: Joan Carles Mora Bonilla
Date of Submission: 07/2012

UNIVERSITAT DE BARCELONA

UNIVERSITAT DE BARCELONA
Facultat de Filologia
Dept. Filologia Anglesa i Alemanya

Gran Via de les Corts Catalanes, 585
08007 Barcelona

Tel. 93 403 56 86
Fax. 93 317 12 49

Official MA programme in
Applied Linguistics and Language Acquisition in Multilingual Contexts
(LAALCM)
Universitat de Barcelona

Non-Plagiarism Statement

This form must be completed, dated and signed and must be included at the beginning of every copy of the MA Thesis you submit for assessment.

Name and surnames:	Anja Ludwig
MA Thesis title:	Interlanguage speech intelligibility benefit for non-native listeners of English
Supervisor:	Joan Carles Mora Bonilla

I HEREBY DECLARE THAT:

- This MA Thesis that I am submitting for assessment is entirely my own work and I have written it completely by myself.
- I have not previously submitted this work or any version of it for assessment in any other program or institution.
- I have not used any other sources or resources than the ones mentioned.
- I have identified and included the source of all facts, ideas, opinions and viewpoints of others through in-text referencing and the relevant sources are all included in the list of references at the end of my work. Direct quotations from books, journal articles, internet sources or any other source whatsoever are acknowledged and the sources cited are identified in the list of references.

I understand that plagiarism and copying are serious offences. In case of proof that this MA Thesis fails to comply with this declaration, either as negligence or as a deliberate act, I understand that the examiner has the right to exclude me from the assessment act and consequently all research activities conducted for this course will be declared null and the MA Thesis will not be presented for public defence, thus obtaining the lowest qualification.

Date: **Signature:**

ABSTRACT

There are different factors that influence speech intelligibility, including properties of the speech itself, the listeners' linguistic backgrounds and their familiarity with different speech varieties. Regarding the latter, an interlanguage speech intelligibility benefit (ISIB) was hypothesized, which predicts that non-native listeners find speech with a familiar accent more intelligible than native speech.

This study examines whether non-native listeners of English show an ISIB due to a familiar accent and their proficiency in English. Therefore, the intelligibility of six speakers differing in L1 (English, Catalan and German) was assessed by 50 listeners differing in L1 (English, Catalan and German) and proficiency in English (high and low). Reaction time was used to detect differences in intelligibility, which is more sensitive than intelligibility measures used in other studies. Results show that, no matter of the L1, for non-native low proficiency listeners, utterances produced by speakers with a familiar accent were more intelligible than native English speech (ISIB-T). Familiar accented non-native English was more intelligible to high proficiency non-native listeners than to native English listeners (ISIB-L). In conclusion, the L1 as well as the proficiency in the target language are crucial factors in intelligibility and more depending on the latter listeners show different types of ISIB.

TABLE OF CONTENTS

NON-PLAGIARISM STATEMENT iii
ABSTRACT v
1 INTRODUCTION 1
2 FOCUS OF THE STUDY AND RESEARCH QUESTIONS 8
3 METHODS 10
3.1 PARTICIPANTS 10
3.1.1 SPEAKERS 10
3.1.2 LISTENERS 10
3.2 INSTRUMENTS 12
3.2.1 SPEECH MATERIALS 12
3.2.2 TASKS 14
3.3 PROCEDURE 16
3.4 ANALYSES 16
4 RESULTS 18
4.1 INTERRATER RELIABILITY 18
4.2 INTRAGROUP CORRELATIONS 18
4.3 RELATIONSHIP BETWEEN L1 AND FAMILIARITY OF ACCENT 19
4.4 EFFECTS OF LISTENERS' L1 AND PROFICIENCY ON INTELLIGIBILITY 21
4.5 EFFECTS OF LISTENERS' L1 AND PROFICIENCY ON COMPREHENSIBILITY AND ACCENTEDNESS RATINGS 25
4.6 CORRELATIONS BETWEEN THE THREE DIMENSIONS 28
5 DISCUSSION 29
6 CONCLUSION 33
REFERENCES xxxv
APPENDIX xliv

1 INTRODUCTION

To study English as a foreign language can pose varying difficulties. For instance, there is no one-to-one relation between the system of writing and pronunciation (Schmied, 1991). The occurrence of non-native speech makes the situation even more perplexing. Pronunciation is therefore one of the greatest threat to intelligibility between native (NSs) and non-native speakers (NNSs) of English. The issues of intelligibility, comprehensibility and accentedness regarding non-native speech affect many speakers of English, as these days NNSs of English outnumber NSs in the world (Crystal, 2003). "English as a lingua franca", a contact language spoken by people, who do not share a native language, is the most common term used in literature, addressing English in its global context (Mauranen, 2009). NNS of English are generally assumed to speak with a foreign accent, which can decrease their speech intelligibility. Besides the accent, there are many other factors that can influence speech intelligibility positively and negatively. For instance, many recent studies have focused on the issue of how non-native listeners (NNLs) of English benefit in understanding accented English speech from familiarity with their own accent (e.g. Bent & Bradlow, 2003; Ingram & Nguyen, 2007). However, the results of those studies are often contrary which could be due to different measure of intelligibility as well as underestimation of one of the influencing factors other than L1.

The present study is aiming to investigate how listeners' familiarity with accented speech and listeners' proficiency level of the target language (TL) affect their intelligibility of English speech produced by speakers differing in L1 background. Focusing on these factors, a different measure of intelligibility than in most studies on intelligibility was used to detect differences in intelligibility more sensitively.

The **definition of intelligibility** in this study is the extent to which an utterance produced by a NS or NNS is understood by a native listener (NL) or NNL (Munro & Derwing, 1995a). The more words a listener is able to identify accurately when produced by a particular speaker, the more intelligible the speech is (Kenworthy, 1987). This notion has to be distinguished from comprehensibility and accentedness, which do not refer to listeners' actual understanding of an utterance but to listeners' perceptions of how easily they understand an utterance and how closely the pronunciation approaches that of a NS (Munro et al., 2006). Those three dimensions are related but partially independent. It has been shown that strong accented speech tends to be less comprehensible and intelligible than unaccented speech, but in contrast, for instance, Munro and Derwing (1995a) ob-

served that Mandarin-accented English speech samples that were rated as moderately or heavily accented were often perfectly intelligible. Parallel findings have been reported in studies concerning ratings of L2 speakers' vowel production (Flege et al., 1995) and speech intelligibility in general (Major et al., 2002; Munro et al., 2006).

Studies regarding the intelligibility of native and non-native speech for NLs and NNLs have revealed many **factors that may contribute to speech intelligibility**. Those factors can be subdivided into listeners' and speakers' factors.

Non-native speech may contain hesitations, grammatical restructuring and self-correction. One speakers' factor is speech rate (Derwing & Munro, 2001) which can hinder intelligibility when it is too rapid. Additionally idiosyncratic speech habits can confuse listeners, wherefore speakers with unique speech habits are not supposed to be included in a study in which they should represent a given variety. Other speakers' factors are stress, pausing, and intonation (Anderson-Hsieh & Koehler, 1988; Tajima, Port, & Dalby, 1997; Munro & Derwing, 1998, 2001; Trofimovich & Baker, 2006), whether talkers are speaking 'clearly' (Bradlow & Bent, 2002), word frequency (Bradlow & Pisoni, 1999), neighborhood density (Imai, Walley, & Flege, 2005), ascendancies of the context (Mayo et al., 1997) and grammatical errors (Ensz, 1982). Some of these factors are correlated to the degree of accentedness. Research has shown, for instance, that foreign accent ratings correlate with the frequency of segmental and prosodic divergences from typical NS patterns (Anderson-Hsieh et al., 1992; Magen, 1998; Munro & Derwing, 2001).

The main listeners' factors with respect to intelligibility are experience which is defined as the extent of previous exposure to L2 speech, proficiency in the TL, the possibility to use the context, and familiarity of the speech. Smith (1992) observed that L2 proficiency needs to be considered, especially when it comes to comprehensibility. Furthermore, the influence of speakers' proficiency on intelligibility was shown in various studies (van Wijngaarden, 2001; van Wijngaarden et al., 2002a, 2002b; Bent & Bradlow, 2003; Stibbard and Lee, 2006). The context can be helpful for a listener to decode a message in face-to-face interaction (Fry, 1955; Jenkins, 2002; Field, 2003, 2004). Even if most visual and auditory clues can rarely be used when speech samples are tape-recorded and played to listeners in the absence of the speaker, knowledge of language rules can be helpful to decode the speaker's message. Familiarity can be the familiarity of native speech and / or non-native accents in L2 speech or the familiarity of a topic or a specific speech event (Gass and Varonis, 1984). Other listeners' factors are the level of tiredness

(Field, 2003), situation specific factors, such as background noise (Rogers et al., 2004; van Wijngaarden et al., 2002a; 2002b), and (only concerning NNSs) semantics (Kennedy & Trofimovich, 2008).

Regarding the listeners' familiarity of native speech and non-native accents, listeners' **L1 background** plays an important role for intelligibility and comprehensibility of L2 speech. Regarding the familiarity of an accent it can either be with the speech of an individual or within a given variety. In general, native English listeners remember and understand more utterances spoken by a familiar than by an unfamiliar speaker, particularly in noisy conditions (Nygaard, Sommers, & Pisoni, 1994; Goldinger, 1996; Munro, 1998; Bradlow, Nygaard, & Pisoni, 1999; van Wijngaarden 2001). This effect also applies to NNLs. If they are familiar to a given variety of English or a certain accent, this will affect the degree of intelligibility. According to Munro et al. (2006) an accented utterance, which is unfamiliar and differs from the native patterns of oral production, requires a greater processing effort. Bent and Bradlow (2003) reported that non-native English listeners find English spoken with a familiar accent more intelligible than native speech. In other words, they suggested that speakers from a particular L1 background might have a benefit in understanding accented speech from speakers with whom they share the language background. They called this effect the "interlanguage speech intelligibility benefit" (ISIB). In their study, Bent and Bradlow (2003) measured how intelligible English sentences produced by NSs of Chinese, Korean and American English were for listeners from Korea, China or mixed L1 backgrounds. As expected, native English speakers were always most intelligible for native English listeners, and both NLs and NNLs judged low proficiency (LP) L2 speakers to be less intelligible. Thus, they found that NNLs perceived high proficiency (HP) non-native English speakers with whom they shared an L1 more or equally intelligible as NSs, confirming the hyzpothesized "matched ISIB" which was also proven by van Wijngaarden (2001) and van Wijngaarden et al. (2002b). The matched ISIB can be explained by the link between non-native speech perception and native language sound structure (Iverson et al., 2003; Best, 1994, 1995; Strange, 1995; Flege, 1995). Furthermore, the researchers found a "mismatched ISIB", that is an ISIB of NNLs concerning unfamiliar accented non-native English speech. Bent and Bradlow (2003) explained this referring to listeners' and speakers' shared knowledge of the structure of the TL in conjunction with influences of general strategies, applied when a foreign language was learned. Even so, it needs to be considered that the listeners'

L1s investigated in their study (Chinese and Korean) are typologically similar in their sound structure, which is why the observed mismatched ISIB could be interpreted as another manifestation of the matched ISIB rather than a separate phenomenon.

More recent work by Imai et al. (2005) proved the ISIB for word recognition. They investigated the ability of NLs and NNLs of English to transcribe utterances, produced by a native English speaker and a NNS with L1 Spanish. Results showed that NLs performed better than NNLs on words produced by a NS but NNLs outperformed NLs regarding words produced by NNSs. Restricting this finding a bit, Bradlow and Pisoni (1999) showed that words' recognition by NNLs is better for high frequency words in sparse neighborhoods than for low frequency words in high density neighborhoods. Major et al. (2002) suggested that an ISIB is probably small and not consistently observable. In their study they examined the extent to which NLs and NNLs perform better, when the L1 is shared by speaker and listener. Four different L1 groups of listeners (Chinese, Japanese, Spanish and American English) were tested. Participants listened to brief lectures presented in English by speakers with different L1s and had to answer questions based on the lectures. Major et al. (2002) found out that only Spanish speakers showed a small ISIB when hearing Spanish-accented English speech, whereas other listener groups did not show a parallel ISIB for their accents in English.

Stibbard and Lee (2006) provided evidence against the mismatched ISIB hypothesis of Bent and Bradlow (2003). In their study, sentences containing keywords were read aloud by five talkers (HP and LP Koreans and Saudi Arabians, and native English). The intelligibility was measured using a word recognition test, which was performed by four different L1 groups (Korean, Saudi Arabian, native English and other mixed L1s). It was shown that NNLs perceived LP speakers of English, who did not share the same L1, to be significantly less intelligible. Stibbard and Lee (2006) concluded that there might be an intelligibility problem between mismatched L1 speakers, who were unfamiliar with each others' accent in English. They also questioned Bent and Bradlow's (2003) use of the term 'benefit' and argued that it should only be used in cases in which a speaker receives a higher (and not only equal) intelligibility score than another one.

Hayes-Harb et al. (2008) and Stibbard and Lee (2006) subdivided Bent and Bradlow's (2003) idea of the benefit into two types of ISIB: the benefit for non-native talkers (ISIB-T) and the benefit for NNLs (ISIB-L). The ISIB-T claims that non-native speech is more intelligible than native speech to NNLs, and the ISIB-L that non-native speech is more intelligible to NNLs than to NLs. Hayes-Harb et al. (2008) found evidence for the ISIB-L

but not for the ISIB-T. They investigated the intelligibility of native and Mandarin-accented English speech for NLs and NNLs of English using a forced-choice word identification task, regarding word-final voicing contrast in minimal pairs. Listeners with L1 Mandarin were more accurate than native English listeners at identifying Mandarin-accented English words, but they did not perceive Mandarin-accented speech to be more intelligible than native English speech. In contrast, Rasmussen (2007) found support for the ISIB-T but not for the ISIB-L. Similar to the study of Hayes-Harb et al. (2008) Rasmussen (2007) investigated the intelligibility regarding the contrast between two consonants in minimal pairs (existing in English but not in Arabic) using a forced-choice word identification task. NSs were more accurate in identifying Arabic accented words, but Arabic-accented utterances were more intelligible than native English for native Arabic listeners. It is noticeable that very similar studies investigating ISIB-L and ISIB-T yielded mixed results. Therefore, further studies investigating this difference using diverse and more sensitive measures are needed to clarify the effect of non-native English speech with a (un-)familiar accent for listeners.

In more recent studies the effect of the ISIB is often more subtle than suggested by Bent and Bradlow (2003). Munro et al. (2006) revealed that ISIB group differences are sometimes limited to certain native languages and concluded that listeners' L1 is less predictive regarding intelligibility than factors associated with the speech signal itself. Additionally, the L2 proficiency of listeners (van Wijngaarden et al., 2002b) and speakers (Bent & Bradlow, 2003; Stibbard & Lee, 2006; van Wijngaarden, 2001; van Wijngaarden et al., 2002a) seems to be a critical factor influencing intelligibility and comprehensibility. Van Wijngaarden et al. (2002b) reported that listeners' L2 proficiency was the main factor determining whether listeners find NSs or NNSs more intelligible. In their study, native Dutch listeners, who were more proficient in English than in German demonstrated an ISIB-T for German, but not for English. In summary, the question of whether a match of L1 affects speech intelligibility and if the listeners' L2 proficiency makes a difference, has not been answered satisfactorily by research, yet. Research has shown that the speakers' and listeners' language background plays an important role in speech intelligibility and recently a large number of studies have examined the ISIB (Major et al., 2002; Bent & Bradlow, 2003; Munro et al. 2006; Stibbard & Lee, 2006; Ingram & Nguyen, 2007; Hayes-Harb et al., 2008). However, the results of these studies are contradictory. While a few showed an ISIB (e.g. Bent & Bradlow, 2003; Ingram & Nguyen, 2007), the majority of these studies did not (e.g. Munro et al., 2006; Hayes-Harb et al., 2008).

There are many different ways to **measure and assess intelligibility**. Kenworthy (1987) posited that an impressionistic and subjective assessments of intelligibility seem to be both, accurate and dependable. On the contrary, in the majority of studies, it has been agreed that objective measures should be preferred to assess intelligibility and subjective measures to assess comprehensibility and accentedness (e.g. Kent, Miolo & Bloedel, 1994; Klein & Flint, 2006).

The most common way of assessing speech intelligibility in L2 research is scoring listeners' transcriptions of an utterance (Derwing & Munro, 1997; Bent & Bradlow, 2003 and Burda et al., 2003, 2005). Listeners hear utterances and are asked to write them down in standard orthography. Intelligibility will then be measured counting the percentage of correctly transcribed words. Other possible tasks to assess intelligibility are comprehension questions, dictation tasks, cloze tests, picture selection in response to stimuli, elicitation of summaries, determination of truth value, and others (see Kent et al., 1994 and Clopper et al., 2006 for a review). Major et al. (2002) described 19 possible procedures that could be used to assess intelligibility, which can be categorized into phonetic contrast analysis, phonological process analysis, word identification tests, phonetic indices derived from continuous speech scoring, scaling of continuous speech, and traditional word-level analysis of continuous speech. Speech elicitation techniques and the task type have important effects on the kind of conclusions that can be drawn about intelligibility. For instance, the score gained in a dictation task might not correlate perfectly with the amount of message the listener has actually grasped. Zielinski (2004) reported cases in which all of the words were correctly identified and written down by the listeners, while he was still not sure over what the intended message of the speaker was. Munro and Derwing (1995a) reported that some participants transcribed utterances perfectly, but judged them as difficult to understand. It might be hypothesized that the tendency to assign low comprehensibility scores to some accented speech samples is partly due to increased processing difficulty, which may not manifest itself in transcription but as increased processing time.

Gass and Mackey (2007) created a comprehensive overview of the current broad areas and research methods of L2 research and allocate the measure of reaction time (RT) to studies that use prompted response. RT is a way of determining how quick the response to a stimulus is. Gass and Mackey (2007) argued that RT indirectly reflects processing. It is assumed that the more time it takes a participant to respond to an utterance, the greater processing load is required, which correlates positively with the degree of difficulty of recognizing the stimulus (Munro & Derwing, 1995b; Gass & Mackey, 2007). It can be

further hypothesized that a long RT needed to access a target item correlates with low intelligibility of speech (Hecker et at., 1966; Munro & Derwing, 1995b; Kosinski, 2010 and Schüppert & Gooskens 2011). Additionally, RT measure avoids intervening variables like proficiency in writing and typing errors which can be mistakenly interpreted as results of intelligibility in transcription tasks. There are several reasons to expect that listeners may take longer to process accented speech. For instance, Munro and Derwing (1995b) proposed that the time required for recognition of accented consonants and vowel segments may be greater because those segments differ from category prototypes. Therefore, the time needed to recognize larger units, such as words, is increasing, too. Another reason for the need of more processing time might be a lack of comprehension of utterances, which then needed to be processed top-down and take longer to be decoded. Schüppert and Gooskens (2011) investigated the role of language attitudes for listeners' abilities to decode Danish and Swedish message. As a measure of speech intelligibility they used individual RT of auditorily presented cognate nouns in a multiple-choice picture-pointing task. Even if they did not find a correlation between attitude and intelligibility, the RT measures were quite sensitive. Munro & Derwing (1995b) determined the effect of accentedness and comprehensibility of speech on processing time. Native English listeners heard English true/false statements produced by NSs and NNSs of English, performed a sentence verification task and assigned accentedness and comprehensibility ratings. Accented utterances and those that were perceived to be less comprehensible were found to require more time to process than non-accented and highly comprehensible utterances.

In summary, listeners who are familiar with an accent may be expected to exhibit faster response times to accented utterances than listeners with little or absence of such a familiarity.

2 FOCUS OF THE STUDY AND RESEARCH QUESTIONS

Of all the above mentioned listeners' and speakers' factors that may contribute to intelligibility, speakers' accent and L1, respectively as well as listeners' L1, familiarity of accented and native English speech and L2 proficiency were focuses of this study. All the other listeners' and speakers' factors were eliminated or they were controlled for as much as possible (see 'Appendix 1'). More specifically, the intelligibility of native English, Catalan- and German-accented English speech was assessed by NLs and NNLs, whose L1 is either the same or different from the L1 of the speakers. Regarding the three dimensions (intelligibility, comprehensibility, and accentedness) the sharper focus was laid on intelligibility. Additionally, comprehensibility and accentedness ratings were taken into consideration for sentences, to see how the three perceptual dimensions were related to one another. For words those two dimensions were not investigated because subjective ratings are not sensitive enough and probably not reliable for such short speech samples.

The ISIB is predicted to arouse from shared L1 of speaker and listeners. Here, Stibbard and Lee's (2006) literal definition of 'benefit' has been adopted, and was, therefore, only used when the intelligibility score of a learner was higher and not equal than the score of others. The terms matched and mismatched ISIB are to be distinguished from each other, too. In the present study the investigated languages other than English (Catalan and German) are not very similar in sound structure. Therefore it is assumed that a possible mismatched ISIB can be clearly assigned to one language. Hence, there is no complicacy in differentiating between the matched and mismatched ISIB. The distinction between the benefit for non-native talkers (ISIB-T) and the benefit for NNLs (ISIB-L), proposed by Hayes-Harb et al. (2008) and Stibbard and Lee (2006), was adopted in this study. In summary, there are four types of ISIB as a function of language background that have to be analyzed: matched or mismatched ISIB-L and matched and mismatched ISIB-T.

The following research questions will be addressed in the present study:

1. Will there be an interaction between accents in English speech and the listeners' L1 with regard to intelligibility in terms of an ISIB?
    a. Will accented English speech, produced by native German or native Catalan speakers, be more intelligible to listeners sharing their language background than to native English listeners? (matched ISIB-L)
    b. Will accented English speech, produced by native German or native Catalan speakers, be more intelligible to NNLs of English not sharing the language background than to native English listeners? (mismatched ISIB-L)
    c. Will native German or native Catalan listeners find utterances in familiar accented English (produced by speakers with whom they share the language background) more intelligible than native English speech? (matched ISIB-T)
    d. Will native German or native Catalan listeners find utterances in unfamiliar accented English, more intelligible than native English speech? (mismatched ISIB-T)
2. Does listeners' L2 proficiency level have an influence on the ISIB-L or ISIB-T?
3. Do comprehensibility and accentedness ratings of native English listeners responding to English speech differ from those of German or Catalan NNLs? What is the relationship between comprehensibility and accentedness ratings and the speech intelligibility?

The following two hypotheses were posed concerning the ISIB-L and the ISIB-T:

1. LP non-native listeners of English will show a matched, but not a mismatched ISIB-T. More specifically, they will find utterances in accented English produced by speakers with whom they share the language background more intelligible than native English speech.
2. HP non-native listeners will show a matched, but not a mismatched ISIB-L. More specifically, to them familiar accented non-native English will be more intelligible than to native English listeners.

3 METHODS

3.1 PARTICIPANTS

3.1.1 SPEAKERS

Thirteen speakers recorded words and sentences in English. Four were NSs of English, four NSs of German and five NSs of Catalan. Out of those 13, two speakers per L1 (one male and one female) were chosen on the basis of overall intelligibility and speech rate; the others were excluded. None of them had noticeable idiosyncratic speech habits and the correlation between individuals' speech rates for words (Pearson *r*: from .947 to .985) as well as sentences (Pearson *r*: from .931 to .979) were very high. Their age ranged from 23 to 29 (M=25.33 years) and all of them were highly educated.

Table 1: Participants - Speakers

Speakers	**L1 English**	**L1 German**	**L1 Catalan**
Number	n=2	n=2	n=2
Age (in years)	M=25.5; SD=0.71	M=26; SD=0	M=24.5; SD=0.71
Country of birth	United Kingdom	Germany	Catalonia, Spain
LoR[1] in country of birth	M=21.5; SD=0.71	M=25; SD=0	M=20.5; SD=0.71
Native language	English	German	Catalan / Spanish
LoE[2] to English		M=9; SD=0	M=8; SD=0
Self-perceived proficiency in English[3]		M=4; SD=0	M=3.5; SD=0.71

The NSs of English, German and Catalan were born and raised in the UK, Germany and Spain, respectively. Neither of the NNS groups had ever lived in an English-speaking country. They learned English for eight to nine years and except of Spanish, in which the Catalan speakers perceived themselves to have a native(-like) proficiency, English was perceived to be their most proficient L2.

3.1.2 LISTENERS

Fifty listeners participated in this study; ten NSs of English, twenty NSs of German and twenty NSs of Catalan. All of them were students and either holder of a Bachelor's

[1] LoR = Lenth of Residence (in years).
[2] LoE = Length of Exposure (in years).
[3] Participants scored themselves on a self-perception scale from 1 (no knowledge) to 5 (native(-like) proficiency.

(n=39) or a Master's degree (n=11). With native Catalan listeners the study was conducted in Barcelona, Spain; with native German listeners in Cologne and Düsseldorf, Germany and with native English listeners the study in Linköping, Sweden and Brighton, UK.

Table 2: Participants - Listeners

Listeners		**L1 English**	**L1 German**		**L1 Catalan**	
Proficiency		HP	HP	LP	HP	LP
Sex	**females**	n=5	n=5	n=5	n=5	n=5
	males	n=5	n=5	n=5	n=5	n=5
Age (in years)		M=24.3; SD=1.19	M=24.4; SD=1.31		M=24.2; SD=1.06	
Country of birth		UK	Germany		Spain	
LoR in home country		M=22.8; SD=2.03	M=21.75; SD= 2.05		M=23.5; SD=2.07	
Native language		English	German		Catalan / Spanish	
Sef-perceived proficiency in English		M=5; SD=0	M=4.5; SD= 0.53	M=2.5; SD=0.707	M=4.1; SD=0.57	M=2.3; SD=0.48
Familiarity to Spanish-accented English[4]		M=1.6; SD=0.39	M=1.3; SD=0.48	M=1.1; SD=0.32	M=4.7; SD=0.48	M=4.9; SD=0.32
Familiarity to German-accented English		M=1.5; SD=0.79	M=4.8; SD=0.42	M=4.7: SD=0.48	M=1.2; SD=0.42	M=1; SD=1
Score in vocabulary size test		M=9750; SD=156.4	M=6685; SD=460.7	M=3300; SD=143.37	M=6620; SD=481.4	M=3215; SD=189.7
LoE to English			M=9.2; SD=0.63	M=9; SD=0	M=7; SD=0.84	M=6.2; SD=1.03

The group of native English listeners (see Table 2), which functioned mainly as a control group, had never lived in a German- or Spanish-speaking country and, therefore, most of them were not used to listen to Spanish- or German-accented English. They were all born and raised in the UK. Foreign languages they have learned are French (n=6), Swedish (n=4), German (n=2), and Spanish (n=1).

The two NNL groups of this study with L1 German or Catalan (see Table 2), were initially selected based on their suspected proficiency level in English to ensure that HP and LP listeners were evenly distributed. They were assigned to a proficiency groups according to their score in the vocabulary size test - to the HP group if the score was higher than

[4] The participants rated their familiarity with German- or Spanish-accented English speech on a five-point scale ranging from 1 ("not at all familiar) over 3 ("from time to time") to 5 ("very much familiar") in the background questionnaire.

6000 and to the LP group if the score was lower than 3500 (adding the scores of X_Lex and Y_Lex). The German participants were all born and raised in Germany. German listeners of the both proficiency groups of were barely familiar to Spanish-accented speech in English, while most of them stated to be very familiar to German-accented English. Participants of the HP German listener group rated their proficiency in English to be near native-like and four of them had lived in an English-speaking country while most participants of the LP group state to be better in another foreign language than English. Other foreign languages besides English that they speak were French (n=15), Italian (n=4), Spanish (n=4), Russian (n=3), Turkish (n=2), Dutch (n=2), Portuguese (n=2), Polish (n=3), Swedish, and Japanese.

The Catalan listeners were all born and raised in Spain. Although only three participants stated to be bilingual in Catalan and Spanish, all of them perceived to have a native-like proficiency in both languages. They speak various other foreign languages (French (n=8), German (n=3), Portuguese (n=2), Italian (n=2), Russian, Japanese and Arabic). Most of them were very familiar with the Spanish accent and rarely with the German accent in English. The HP Catalan listener group perceived English to be their strongest foreign language and two of them have had stayed for half a year each in an English-speaking country while the LP group perceived themselves to have a low proficiency in English.

3.2 INSTRUMENTS

3.2.1 SPEECH MATERIALS

Accented non-native English speech is attributed to many *typical derivations* related to phonemic inventories of speakers' L1. Altering of vowels and consonants, epenthesis, wrong word stress or rhythm are common phenomena in non-native speech (Derwing & Munro, 2005). There are also characteristics of accented speech that are sub-phonemic, resulting of differences in the phonetic implementation of phonemes in the speaker's L1 and L2 (Flege and Eefting, 1987). Language-specific derivation of Spanish and German speakers of English (see Kenworthy, 1987 for a more detailed description) could cause problems in the speech intelligibility. Therefore, they were taken into account, when

choosing the utterances for the tasks[5] to ensure that all words and sentences have a comparable degree of difficulty in pronunciation for both NNS groups of English.

Twenty words (ten animated and ten non-animated, see 'Appendix 2') and ten sentences, adapted from Munro and Derwing's (1995b) bank of single-clause true–false sentences, (five true and five false, see 'Appendix 3) were chosen controlling for frequency and cognate status. To test the frequency of the items a vocabulary profile feature available from [http://www.lextutor.ca/vp/] was used. All items were chosen from the 2000, and if possible even from the 1000, most frequent English words, to control for the effect of lexical knowledge. The cognate status was checked for English-German and English-Catalan, so that either the English word is a cognate in both languages or in neither of them. In total there were six cognates but no cognate effects were found.

After performing the task the main effect of learning, that if a sentence is heard more than once, people are likely to know on the nature of it after the initial words, which would diminish any intelligibility effects, was detected. Therefore, all sentences should be different, which is a matter of changing the content words, to force participants to listen to the whole sentence. 60 different sentences (see 'Appendix 3') were created in Praat by combining different phrases out of the original 10 sentences. Thereby, ten (5 true and 5 false) sentences were spoken by each speaker. The sentences' length varied from 6 to 10 syllables (M=8 syllables). They were all similar in structure and still fulfill the criteria mentioned above.

All of the speakers were recorded in a sound treated booth at the University of Barcelona. The printed list of speech material consisted of the originally 20 words and 10 sentences. Each word and sentence was read aloud at least twice. The speech samples were separated into single items or sentences using Praat. Out of the chosen six speakers, samples matching the following criteria were selected: no unique speech habits, no errors, similar speech rate, stress, pausing, and intonation, and clear speech. All recorded utterances were normalized by rescaling to full amplitude range using GSU tools in Praat. DC contamination was removed by subtracting mean of waveform, AC contamination was removed using highpass filtering, and the sampling frequency was resampled to 11025 Hz. For the comprehensibility rating task, the speech samples were additionally mixed with cafeteria noise to make them less comprehensible and to prompt a bigger distinction using the program GoldWave.

[5] Although the speakers' and listeners' L1 of this study is Catalan or they are bilingual in Catalan and Spanish, the typical derivations from native English are very similar for both languages.

To ensure the same or at least a similar degree of foreign accent of the non-native speech samples, four NSs of English were asked to rate the accentedness of the speech samples on a seven-point scale ranging from 1 (“no foreign accent”) to 7 (“very strong foreign accent”). The Catalan speakers obtained a mean rating of 4.75 which was slightly higher than the German speakers’ rating (M=4.0), but in general all speech samples have a similar degree of accentedness.

3.2.2 TASKS

3.2.2.1 Vocabulary Size Test

As a measure for overall proficiency in English two vocabulary size test available at [http://www.lognostics.co.uk/tools/] were used. X_Lex v2.05 is a vocabulary size test covering the 5000 most frequent English words and Y_Lex v2.05 is the advanced test sampling vocabulary in the 5000 to 10000 word range. Both programs test vocabulary breadth by presenting a set of words, one at a time, in a context-free environment.

Those tools are in the first instance designed for vocabulary acquisition research, but a significant relationship between vocabulary size and overall proficiency was shown (e.g. Meara and Milton, 2003; Meara, 2005; Milton, 2006, 2007 and Nemati, 2010). Milton & Alexiou (2009) stated that vocabulary size and examination level link quite closely and can be credibly tied into the Common European Framework of Reference (CEFR). Meara and Milton (2003) explicitly link X-Lex scores to CEFR levels (see Table 3). Since this correlation has not been shown for Y_Lex scores, yet, X_Lex scores are primarily used to assign a participant to a proficiency level and Y_Lex scores are used to adjust this result up or down.

Table 3: Vocabulary size and the CEFR framework (adapted from Meara & Milton, 2003)

CEFR level	X_Lex
A1	< 1500
A2	1500 – 2500
B1	2500 – 3250
B2	3250 – 3750
C1	3750 – 4500
C2	4500 – 5000

3.2.2.2 Intelligibility Task

The intelligibility task was designed with DMDX (Forster & Forster, 2003) especially for this study and consists of two parts.

The first part is a forced-choice word identification task. 120 lexical items were presented aurally and listeners had to decide whether a word is living or non-living by pressing one of two keys on the keyboard. If the key was not pressed for 5 seconds the next items was provided automatically. Participants were asked to identify the word they hear as fast as possible. A few examples, which will not be included in data analysis, were provided for practice to make sure that listeners were familiar with the type of speech samples and the distinction they have to make.

After the word identification task a 60-item sentence verification task (true – false) started automatically providing the same instruction as for the first task. Participants heard one sentence at a time, and had to decide whether the sentence they hear is true or false.

3.2.2.3 Comprehensibility rating task and accentedness rating task

In the accentedness and comprehensibility rating tasks listeners rated the degree of both dimensions for 60 sentences, one at a time, on a 7-point scale ranging from 1 (“no foreign accent” or “very easy to understand”) to 7 (“very strong foreign accent” or “very difficult to understand”). Raters listened to each utterance maximally twice, without seeing a transcription of it.

The rating data complement the intelligibility results regarding the sentences with the closely related perceptional dimensions of comprehensibility and accentedness.

3.2.2.4 Background questionnaire and word familiarity test

All speakers and listeners filled out a background questionnaire on personal details and language history (see Appendix 4).

Participants also filled out a word familiarity test listing the 20 words and all 39 content words from the sentences to ensure that they were able to understand the meaning of the utterances and to control for lexical knowledge. Participants had to choose between three levels of familiarity: “I know the word”, “I have seen the word before, but I'm not entirely sure of its meaning” and “I don’t know the word” (modeled after Smith, 2004).

3.3 PROCEDURE

All tasks were administered in one session and performed on laptops in the place of participants' residence. Four people were helping to conduct the study in Germany, Sweden, and England. Each of them was provided with information how to configure the programs needed for the tasks and a short script including how to instruct the participants, the order of the tasks, and some general issues they have to pay attention at.

Each participant was met individually at a reserved quiet room at the university or at another suitable place. The tasks were administered in the following order:

1. vocabulary size test (X_Lex and Y_Lex)
2. word identification task and sentence verification task to test intelligibility
3. accentedness rating task
4. comprehensibility rating task
5. word familiarity test
6. background questionnaire

For the listening tasks the test taker was equipped with headphones to ensure that possible background noises would not disturb or interrupt him or her. Between the different tasks short breaks were provided to avoid tiring effects.

For the two rating tasks the test taker was instructed orally to explicitly rate either the degree of accentedness or comprehensibility. Before running the second rating task the researcher drew the attention to the difference between both dimensions. Since participants are likely to rate the accentedness in a comprehensibility rating task, the order of tasks helped to distinguish between the two dimensions. Raters were encouraged to use the whole scale when rating, and to listen to the whole utterance before making a decision. Total time for the session was approximately 45 minutes.

3.4 ANALYSES

To test consistency between participants' scores, preliminary the interrater reliability for rating data (comprehensibility and accentedness ratings) and a correlation analysis for intelligibility scores were assessed. Additionally the hypothesized relationship between *L1* and familiarity of *accent* was examined.

To run the main analyses mean RTs (measured from word onset and calculated on the basis of correct identification only) and mean rating scores were used. To determine effects of listeners' L2 *proficiency* and the *L1* two different mixed between-within analyses

of variance (ANOVAs) were run for each of the three dimensions: intelligibility, comprehensibility, and accentedness. The first one included all *L1* groups, but *proficiency* as a between-subject variable was excluded. *L1* was the between-subject variable consisting of three languages (Catalan, German, and English). Within-subjects factors were the different *accents* (Spanish-accented, German-accented, and native English) and the *stimulus types* (words or sentences). The second mixed between-within ANOVA included *proficiency* as an additional between-subjects factor and excluded the group of English listeners because they are not comparable in terms of *proficiency* with the NNL group. To adjust the two types of ANOVAs for comprehensibility and accentedness ratings, which were only assessed for sentences, *stimulus type* was excluded as a within-subjects factor. If needed, additional one-way and univariate ANOVAs were conducted to explain occurring interactions or to clarify the effect of the *L1* and *proficiency* on the three dimensions.

4 RESULTS

4.1 INTERRATER RELIABILITY

Interrater reliability was assessed for accentedness and comprehensibility rating scores. Table 4 presents intraclass correlations (ICC) for the two ratings for each listener group. These correlations ranged from .981 to .996, indicating good interrater reliability for all listener groups on both ratings. The members of each listener group strongly agreed with one another on the relative comprehensibility and accentedness of the speakers.

Table 4: ICCs for accentedness and comprehensibility ratings by listener groups

Listener groups	Accentedness	Comprehensibility
	Sentences	
L1 Catalan	.995	.977
L1 German	.996	.983
L1 English	.994	.981

4.2 INTRAGROUP CORRELATIONS

Intelligibility scores of each listener for each stimulus were used to examine the intragroup correlations on intelligibility. The correlations, expressed as Pearson *r* values, were calculated for all possible pairings of participants within one listener group (see Table 5). Each of the correlations was significant at a $p<.01$ level. The correlations (r) ranged from .927 to .999, thus, the percentage of variance ranged from 85.93% to 99.8%.

Table 5: Pearson *r* for all individuals within each listener group

Listener groups	Proficiency in English	**Pearson *r***		**Percentage of variance**	
		from	to	from	to
L1 Catalan	HP	.968	.999	93.70	99.80
	LP	.962	.999	92.54	99.80
	Total	.930	.999	86.49	99.80
L1 German	HP	.935	.999	87.42	99.80
	LP	.977	.999	95.45	99.80
	Total	.927	.999	85.93	99.80
L1 English	Total	.945	.998	89.30	99.60

4.3 RELATIONSHIP BETWEEN L1 AND FAMILIARITY OF ACCENT

Results from the word familiarity test showed that all participants were 100% familiar with all target items. Therefore, a difference in participants' lexical knowledge can be excluded and all items were included in the analyses.

Familiarity ratings of Spanish- and German-accented speech were submitted to one-way ANOVAs with *L1* (Spanish, German, English) as the between-subjects factor detecting a significant main effect of *L1* ($p<.001$). The Tukey HSD post-hoc test showed that *L1* German as well as *L1* Spanish listeners were significantly more familiar to their own *accent* ($p<.001$) in English while *L1* English listeners rated both *accents* similarly at a low level of familiarity (Figure 1).

Figure 1: Mean familiarity ratings for different accents in English

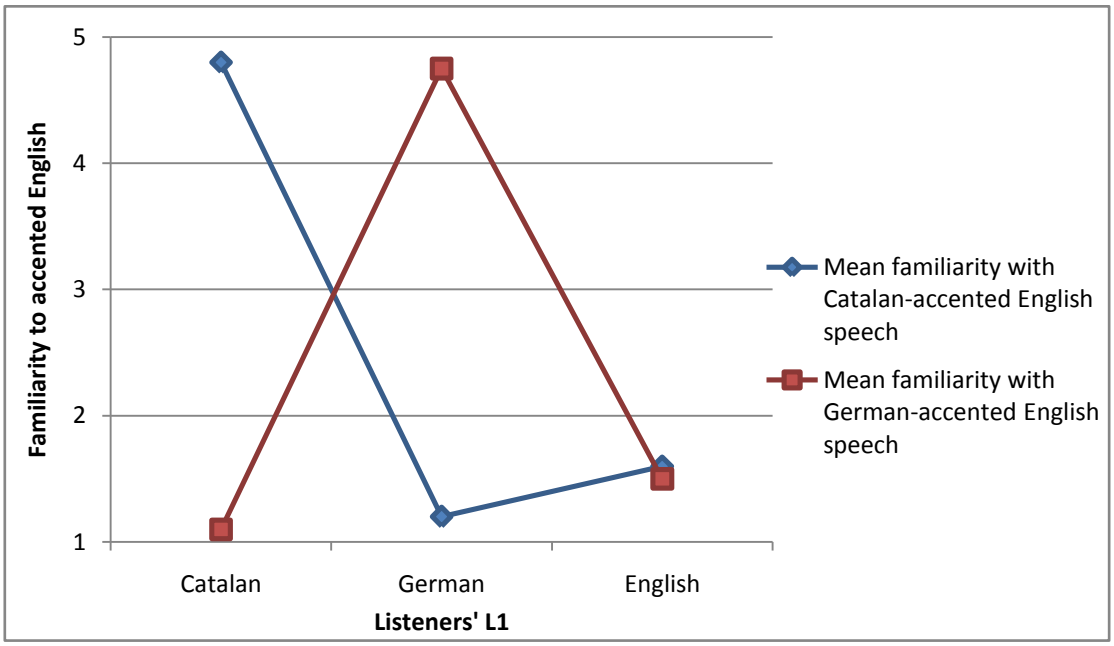


Table 6: Mean results for intelligibility, accentedness, and comprehensibility

Subjects	Proficiency in English	**Intelligibility**						**Accentedness**			**Comprehensibility**		
		words			sentences			sentences					
		Spanish accent	German accent	Native English	Spanish accent	German accent	Native English	Spanish accent	German accent	Native English	Spanish accent	German accent	Native English
L1 Catalan	HP	1052.53	1194.56	1092.75	2092.83	2616.33	2121.91	5.96	5.08	1.17	3.95	3.81	1.58
	LP	1146.99	1254.90	1262.69	2512.99	2867.94	2780.28	5.31	4.71	1.40	3.80	5.19	2.56
L1 German	HP	1166.55	1048.17	1044.59	2608.46	2152.17	2145.24	6.05	4.92	1.22	4.8	3.88	1.56
	LP	1409.11	1233.67	1367.22	2953.20	2503.82	2790.69	6.04	4.56	1.42	5.69	4.03	2.78
L1 English		1186.24	1205.45	1016.16	2474.12	2460.62	2065.23	6.16	5.06	1.05	4.15	3.50	1.23

4.4 EFFECTS OF LISTENERS' L1 AND PROFICIENCY ON INTELLIGIBILITY

Table 6 shows mean RTs (in ms), mean comprehensibility and accentedness ratings for all five listener groups: native English listeners and LP and HP German and Catalan listeners. The following analyses were based on this data.

A mixed between-within ANOVA was conducted to explore the impact of *L1* as a between-group variable and *accents* and *stimuli types* as within-group variables on intelligibility, as measured through RTs. The results of the ANOVA are presented in Table 7.

Table 7: Results for intelligibility of a mixed between-within ANOVAs

Mixed between-within ANOVA including all L1 groups		INTELLIGIBILITY		
		F	$\eta 2$[6]	*p*
Effects	*Accent*	23.12	.501	<.001
	Stimulus type	2062.54	.978	<.001
Main effect	*L1*	2.17	.022	=.126
Interactions	*Accent* and *L1*	61.67	.728	<.001
	Accent and *stimulus types*	12.16	.346	<.001
	Accent, *stimulus types* and *L1*	26.85	.539	<.001
Additional results for a mixed between-within ANOVA including proficiency				
Main effect	*Proficiency*	119.17	.841	<.001
Interactions	*Proficiency* and *accent*	46.65	.727	<.001
	Proficiency and *Stimulus types*	33.55	.428	<.001
	Proficiency, *accent* and *stimulus types*	12.26	.412	<.001

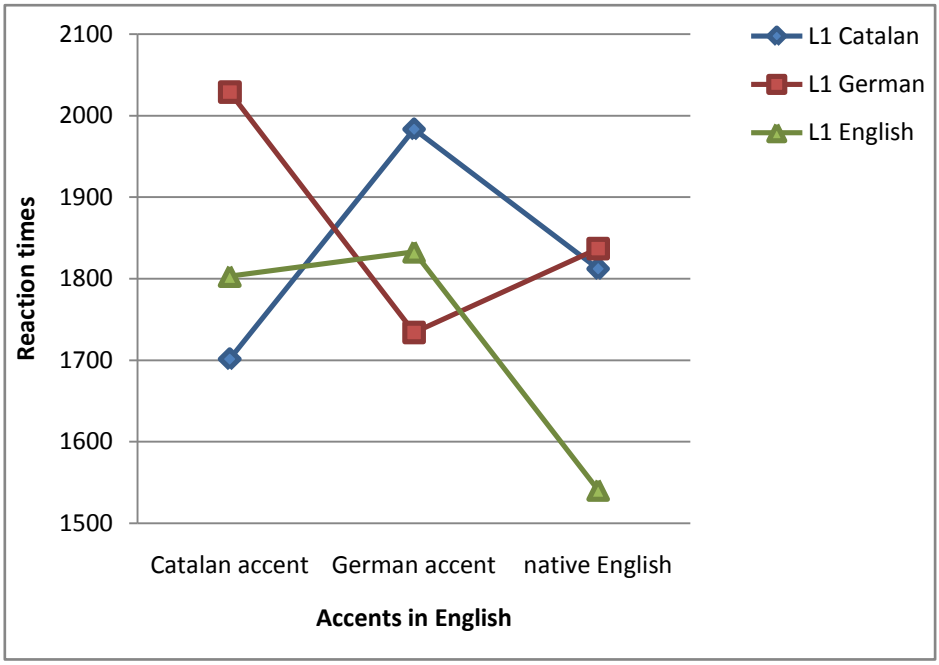
[6] Values for $\eta 2$can range from 0 to 1. To interpret the strength of $\eta 2$values the following guidelines are used (from Cohen, 1988): .01=small effect; .06=moderate effect; and .14=large effect.

The effects for *accent* and *stimulus types* have to be interpreted with regard to several interactions between the variables. The two-way interaction between *accent* and *L1* is consistent with the hypothesized effect that one or more listener groups will understand a particular *accent* better than the other listener groups.

Univariate ANOVAs for each type of *accent* showed that *L1* has a main effect on each *accent*: Catalan-accented English [$F(2, 47)=10.95$, $p<.001$], German-accented English [$F(2, 47)=16.72$, $p<.001$] and native English [$F(2, 47)=6.95$, $p=.003$]. The Tukey HSD post-hoc test showed that regarding non-native *accents*, the listener group that was familiar with the *accent* showed significantly lower RTs ($p<.001$ for both), while the listener group that was unfamiliar with the *accent* showed significantly higher RTs ($p<.001$ for both) than native English listeners (Figure 2). For native English speech both NNL groups scored similar ($p=.948$) and significantly slower that the NL group ($p<.001$ for both).

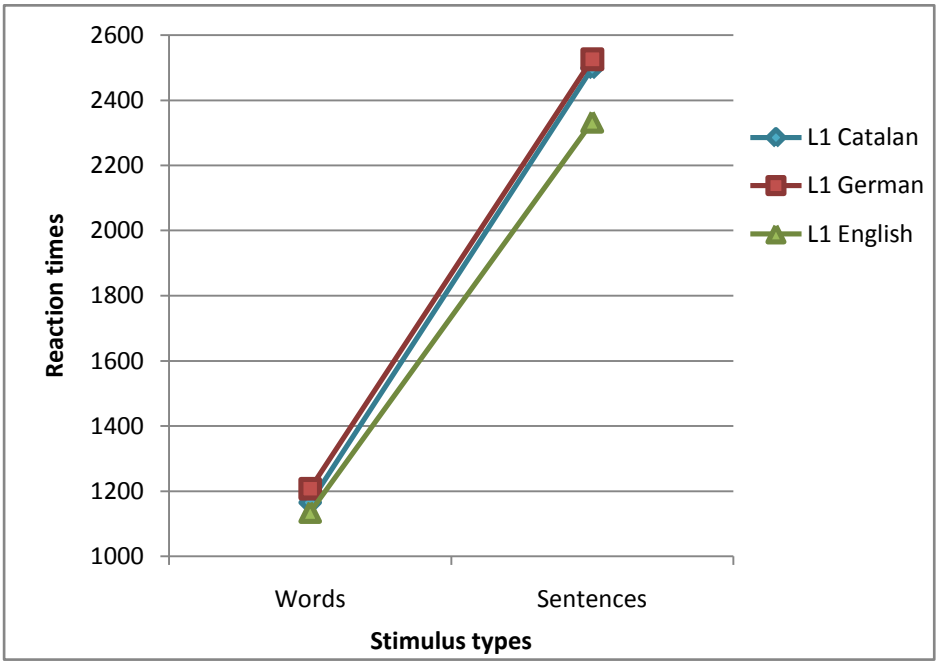
Looking at each *L1* separately it can be stated that both NNL groups were significantly faster in understanding the familiar than the unfamiliar *accent* ($p<.001$). Additionally they understood the unfamiliar *accent* significantly worse than native English ($p<.001$ for both) but only L1 Catalan listeners understood their own *accent* significantly better than native English ($p=.002$). English listeners understood native English significantly faster ($p<.001$) than both non-native *accents* between which they did not show a difference in RT.

Figure 2: Interaction between *L1* and *accent* for RTs



The interaction between *accent* and *stimuli types* was probably due to the significant difference in RTs (Figure 3). *Sentences* required higher RTs than *words* (p<.001). Simple main effect analyses showed a main effect for *L1* for all *stimuli types* and *accents* except for German-accented *words* for the German *accent* (p=.139), explaining the three-way interaction between *accent*, *stimuli types*, and *L1*. Additionally, some of the above mentioned tendencies were not significant for *words* (the difference between Catalan and German listeners for German-accented *words* and the difference between English and Catalan listeners for native English *words*), while they were all significant for *sentences*. Comparing the different *L1s* it can be said that the above describes tendencies were not that clear for *words*. For instance, the difference between German-accented speech and native English was not significant for both non-native English listener groups.

Figure 3: RTs of listener groups for different stimuli types

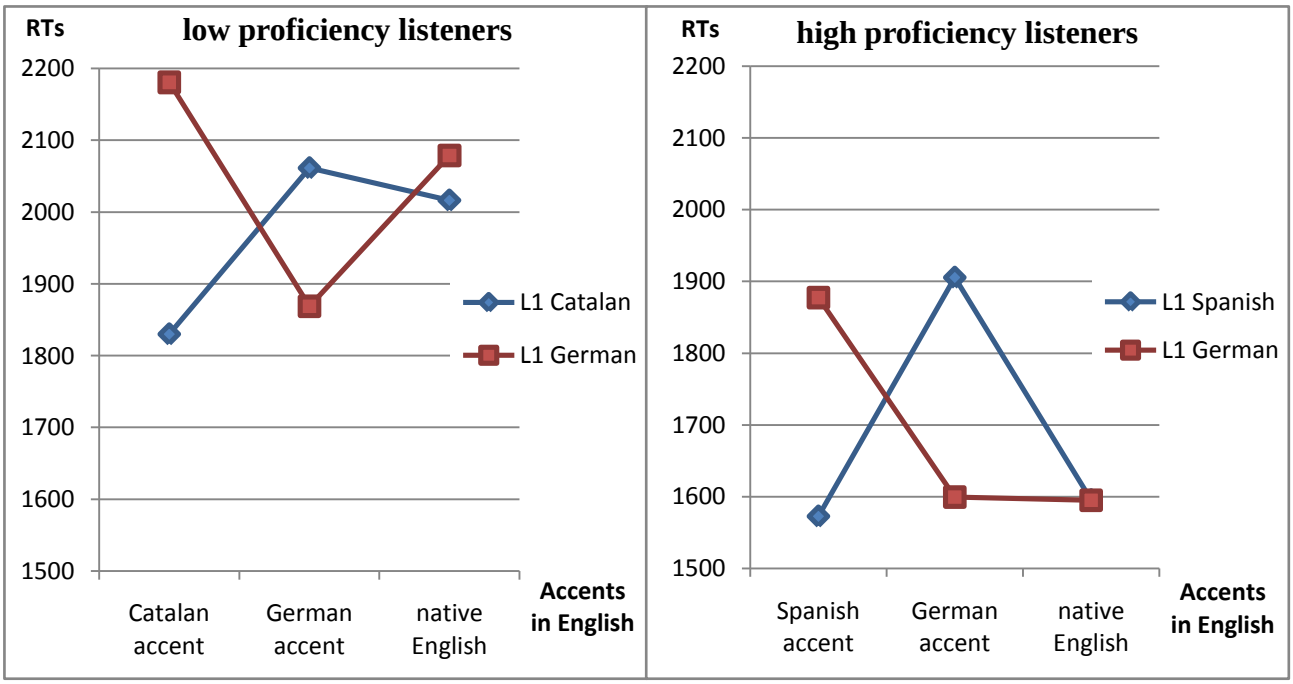


To examine additionally the effect of *proficiency* an ANOVA, excluding the native English group, was run. As a preliminary point it can be stated that the proficiency scores of HP and LP groups were very similar between Catalan and German listeners [t(38)=1.79, p=.197].

The mixed between-within ANOVA with *proficiency* being an additional between-subject variable showed a significant main effect of *proficiency* and various interactions (Table 7). To explore them, the two *proficiency* levels were analyzed independently. HP

listeners were always significantly faster in rating speech samples than LP listeners (p<.001 for all three *accents*, Figure 4). Regarding the *accents* there was a main effect of *L1* for German- or Catalan-accented speech (p<.001) but not for native English (p=.678). For the German *accent* German listeners responded faster than Catalan listeners and for the Catalan *accent* vice versa. Regarding the different *L1s* it can be stated that both Catalan and German listeners understood familiar accented speech significantly better than non-familiar accented speech no matter of their level of *proficiency*. Both LP groups found their own *accent* more intelligible than native English (p<.001 for both), which did not apply to the HP groups that found the unfamiliar *accent* significantly less intelligible than native English (p<.001 for both).

Figure 4: RTs of HP and LP L1 Spanish and L1 German listeners for different accents



Additionally, for sentences HP listeners were significantly faster in responding than LP listeners (p<.001 for all three accents), while for words this effect was not significant for the German accent (p=.022).

4.5 EFFECTS OF LISTENERS' L1 AND PROFICIENCY ON COMPREHENSIBILITY AND ACCENTEDNESS RATINGS

The same two types of ANOVAs as for intelligibility were run for comprehensibility and accentedness with *sentences* being the only *stimulus type*. To determine the effects of *listeners' L1 on comprehensibility ratings* a mixed between-within ANOVA was run with *L1* being the between-subject variable and *accents* the within-subjects factor. As for intelligibility the expected interaction between *accent* and *L1* occurred. Nevertheless, there was a significant main effect for *L1* and a significant effect of *accent* (Table 8).

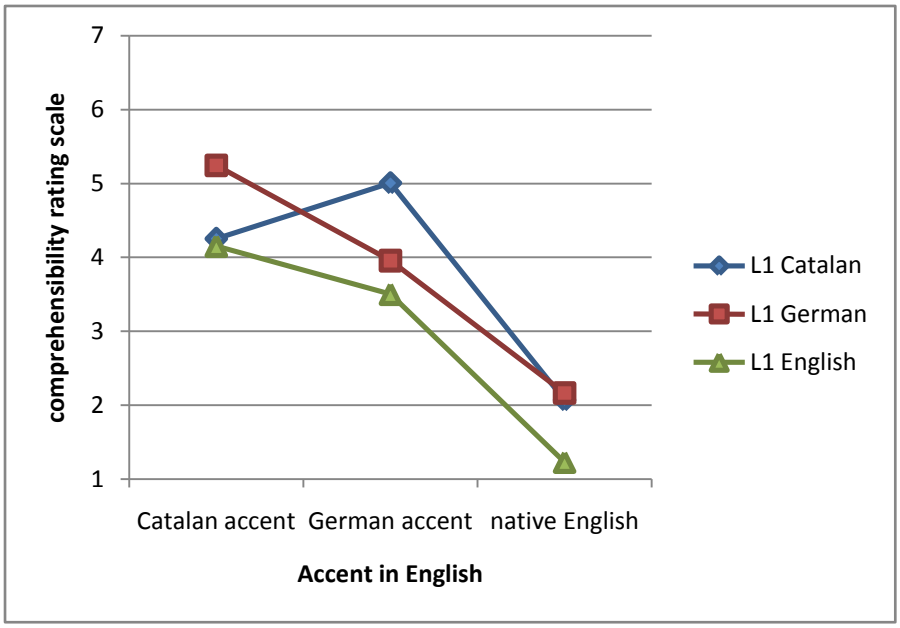
Table 8: Results for comprehensibility and accentedness of mixed between-within ANOVAs

Mixed between-within ANOVA including all L1 groups		Comprehensibility			Accentedness		
		F	*η2*	*p*	*F*	*η2*	*p*
Effects for	*Accent*	222.75	.906	<.001	2243.8	.990	<.001
Main effect of	*L1*	16.21	.411	<.001	.817	.017	=.448
Interaction	*Accent* and *L1*	10.59	.315	<.001	6.5	.220	<.001
Additional results for a mixed between-within ANOVA including proficiency							
Main effect of	*Proficiency*	19.60	.352	<.001	7.4	.032	=.15
Interactions	*Proficiency* and *L1*	8.71	.195	=.006			
	Proficiency and *accent*	10.50	.375	<.001			
	Proficiency, *L1* and *accent*	8.00	.314	=.001			

Simple main effect analyses showed a main effect of *L1* (p<.001 for accented speech and p=.002 for native English) for all *accents*. Non-native accents were perceived to be less comprehensible by the NNL not sharing the *L1* than by NNL sharing the *L1* and native English listeners (p<.001 for both, Figure 5). Native English was perceived to be more comprehensible by English listeners than by German (p=.002) and Catalan (p=.007) listeners. Regarding the different *L1s* NNLs perceived their familiar accent to be signifi-

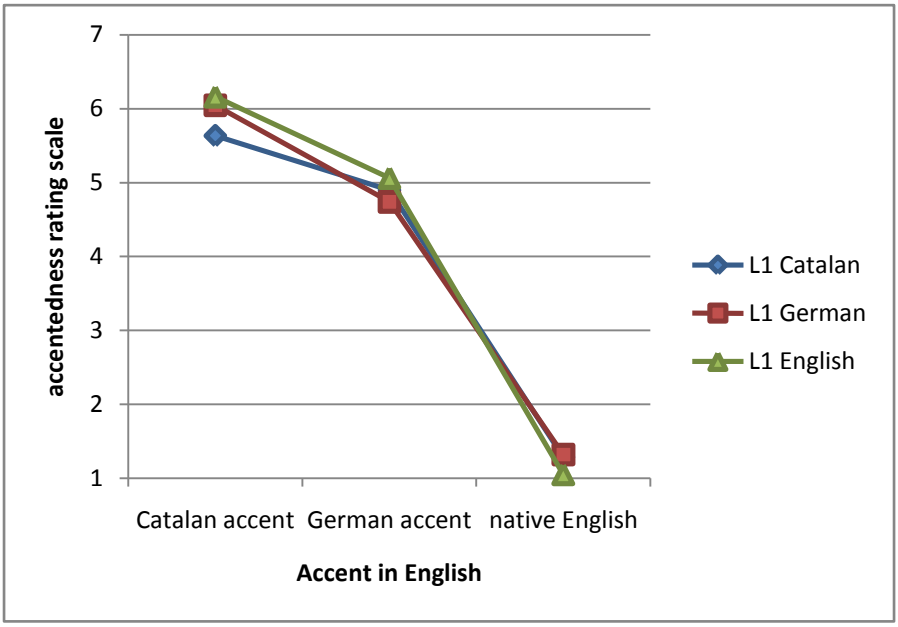
cantly more comprehensible than the unfamiliar accent (p<.001 for both accents) and as comprehensibility as native English (p=.934 for the Catalan accent and p=.119 for the German accent).

Figure 5: Mean ratings for comprehensibility of different listener groups



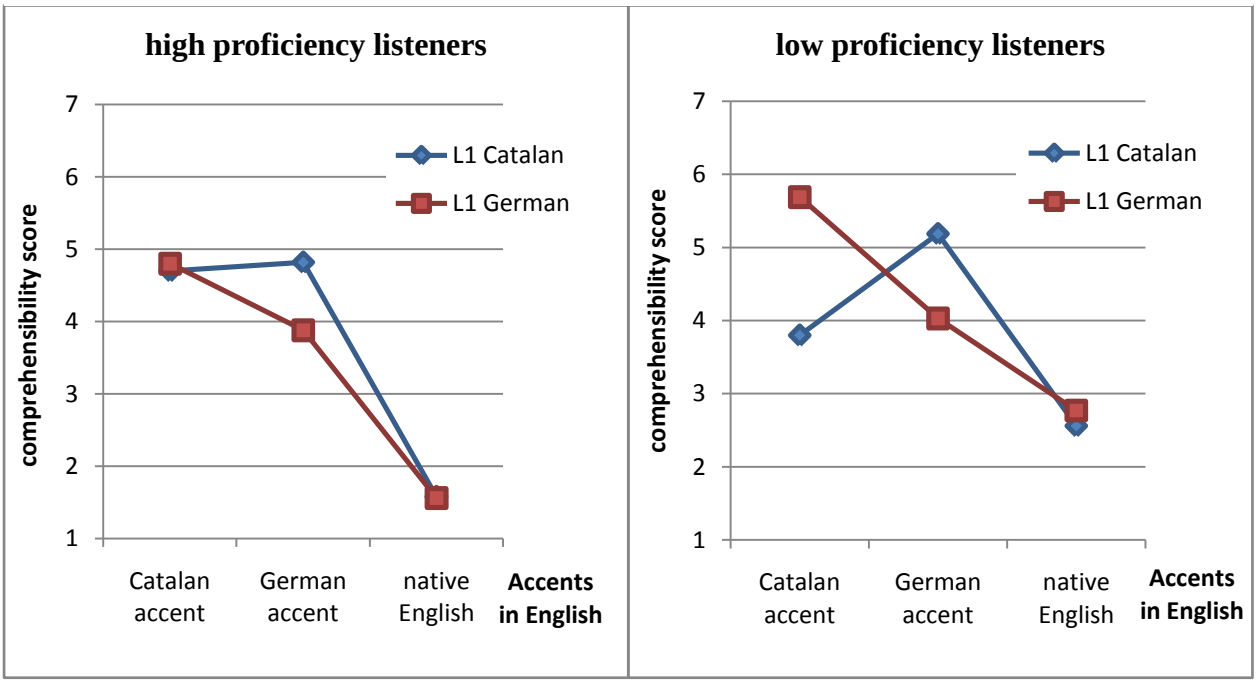
To determine the **effects of listeners' L1 on accentedness ratings** the same mixed between-within ANOVA as for comprehensibility was run. There was no main effect for *L1*, but a significant effect of *accent* (Table 8). For all listeners Spanish-accented *sentences* were perceived as being (not significantly) slightly more accented than German-accented *sentences*. As expected, unaccented utterances were perceived as native-like. Accentedness ratings seem to be more coherent regarding the different *L1s* than comprehensibility ratings and intelligibility scores (Figure 6). Simple main effect analyses showed only a significant main effect of *L1* for Catalan-accented sentences (p=.003) due to the fact that Catalan-accented English was perceived to be less accented by Catalan listeners than by native English listeners (p=.009).

Figure 6: Mean ratings for accentedness for different listener groups



The purpose of the second set of analyses was to examine the **effect of *proficiency***, excluding the group of English listeners. The mixed between-within ANOVA for comprehensibility showed a significant main effect of *proficiency* and the same interactions as for intelligibility: between *L1* and *proficiency*, *accent* and *proficiency* and a three-way interaction between *L1*, *accent* and *proficiency* (Table 8).

Figure 7: Comprehensibility scores of HP and LP Spanish and German listeners for different accents



Both proficiency groups were analyzed separately (Figure 7). The interaction between *L1* and *accent* occurred only for the LP group (p<.001), due to the described effect that

NNL groups understand the familiar *accent* better than the other NNL group. For the HP group only German-accented *sentences* were perceived as being less comprehensible by Catalan listeners (p<.001), while the other *accents* were perceived in a similar way by both listener groups. Listeners' *proficiency* (p=.15) did not influence accentedness ratings.

4.6 CORRELATIONS BETWEEN THE THREE DIMENSIONS

The relationship between intelligibility scores, and comprehensibility and accentedness ratings was investigated using Pearson product-moment correlation coefficient.

Table 9: Pearson correlation between the three dimensions for sentences (r value and percentage of variance)

Accents	Dimensions	Intelligibility	Comprehensibility	Accentedness
Spanish	Intelligibility	1** (100%)	.458* (20.9%)	.024
	Comprehensibility		1 (100%)	.894** (79.9%)
	Accentedness			1** (100%)
German	Intelligibility	1** (100%)	.589** (34.7%)	.411* (16.9%)
	Comprehensibility		1** (100%)	.969** (93.9%)
	Accentedness			1** (100%)
English	Intelligibility	1** (100%)	.931** (86.7%)	.956** (91.4%)
	Comprehensibility		1** (100%)	.927** (85.9%)
	Accentedness			1** (100%)

The correlations, expressed as Pearson r values, were calculated for all pairings of intelligibility, comprehensibility and accentedness scores for each *accent* (Table 9). Each of the correlations, which was significant at a p<.01 level was asterisked (intermediate correlations with one asterisk and strong correlations with two). Positive correlations indicate that a strong foreign *accent*, low comprehensibility and low intelligibility are associated. For all *accents* the correlation between accentedness and comprehensibility ratings were very strong, while the strength between intelligibility and comprehensibility varied regarding the different accents. The correlation between intelligibility and accentedness was the weakest and did not reach significance for the Spanish accent.

5 DISCUSSION

Previous studies have shown that listener responses to L2 speech may be influenced by a variety of factors (Lippi-Green, 1997; Stibbard and Lee, 2006). However, the most important outcomes of this study were the differences in intelligibility, comprehensibility and accentedness scores across listener groups due to their *L1* and *proficiency* in English.

The **first research question** was asking whether the combination of different *accents* and listeners' *L1* influences speech intelligibility. The strong interaction between both proved that different combinations of them have various influences on speech intelligibility. In case that listener's and speaker's *L1* matched, the speech was more intelligible, which can be explained by the fact that non-native speech perception and production are linked to L1 sound structure (Strange, 1995; Best, 1994, 1995; Flege & Fletcher, 1992; Flege, 1995).

This study was aiming to examine the ISIB-T and ISIB-L as separate phenomena. It was found evidence for both effects with respect to the listeners' *L1* and L2 *proficiency*. Answering the research question concerning the matched ISIB-T, it can be stated that native Catalan and native German listeners found utterances in accented English, produced by speakers with whom they share the *L1,* more intelligible than native English speech, although for German listeners the effect did not reach significance. This finding is similar to that of Major et al (2002), who concluded that NNLs of English sometimes, but not always, found speech produced with their own accent more intelligible than with another accent (see also Smith & Rafiqzad, 1979; van Wijngaarden, 2001; van Wijngaarden et al., 2002a; Bent & Bradlow, 2003; Smith et al., 2003; Munro et al., 2006). In accordance with other studies, NLs of English always found NSs most intelligible (Munro, 1998; van Wijngaarden 2001).

Answering the research question concerning the matched ISIB-L, it can be stated that accented English speech, either if it is produced with a German or Catalan *accent*, is more intelligible to listeners sharing the *L1* than to native English listeners, even if the difference did not reach significance for all groups. This could be due to the fact that overall shared phonetic knowledge of L1 and L2 features between NNS and NNL sharing the same *L1* is likely to be more extensive than between NNS and NL. The ISIB-L pattern has been reported previously in the literature by Weinreich (1953), Imai et al. (2005), and Munro et al. (2006), and was shown especially by LP listeners. The role of proficiency will be discussed more in detail below.

In accordance with the results of Stibbard and Lee (2006), this study provides evidence against the mismatched ISIB hypothesis of Bent and Bradlow (2003). Regarding the ISIB-T and ISIB-L, the non-familiar *accent* was always least intelligible and significantly less intelligible than native English to both NNL groups. The native English control group perceived Spanish- and German-accented English similarly. NNLs, not sharing the L1 with the talkers, showed a big disadvantage in understanding the speech, which provides evidence for a "mismatched interlanguage speech intelligibility detriment" (Stibbard and Lee, 2006, p. 440).

There are several explanations for the ISIB-T and ISIB-L not reaching significance for all listener groups. As mentioned earlier, *stimulus type* and listeners' *proficiency* in English interacted with *accent* and were, therefore, likely to influence speech intelligibility. Taking the *stimulus type* into account the described tendencies regarding the ISIB-L and ISIB-T are stronger and clearer for *sentences* than for *words*, as some of them are not significant for the latter. More important is listeners' English *proficiency*, which had a main effect on speech intelligibility. A goal of this research was to investigate its role in mediating the ISIB-T and ISIB-L. Answering the **second research question** it can be stated that for HP listeners all speech samples were significantly more intelligible than for LP listeners. Regarding the ISIB-L, both, LP and HP listeners responded faster to familiar accented English than to the unfamiliar accent. To be able to conclude whether there was an ISIB-L for one of the *proficiency* groups, their results were compared (only in terms of intelligibility and not *proficiency*) to the intelligibility scores of L1 English listeners (bearing in mind that for the analyses native English listeners were excluded because they are not comparable to the NNLs in terms of *proficiency*). The ISIB-L held only for HP non-native English listeners, to whom familiar accented English speech was more intelligible than to native English listeners. LP listeners scored always lower than native English listeners and even if there was no significant difference for their familiar accent, it was not interpreted as a benefit in this study.

Regarding the ISIB-T, to non-native HP listeners native English was as intelligible as the familiar accent, while the unfamiliar accent was significantly less intelligible. LP listeners found native English speech significantly less intelligible than the familiar accent but still more intelligible than the unfamiliar accent. Therefore, the results for ISIB-T differed with regard to the *proficiency* level, as LP listeners showed a matched ISIB-T, but the HP groups did not.

Accented speech may be more intelligible than native English speech to non-native English listeners with limited proficiency, because NNLs were exposed to less non-accented speech and were, therefore, sensitive to cues available in their own accent in English, which might offset their relative lack of experience with English. In contrast, native English listeners are likely to be less familiar to the language specific cues of accented English, although it had been demonstrated that NLs improve their ability to comprehend non-native speech with experience (Bradlow & Bent, 2008). As learners reach higher levels of L2 proficiency, they may adapt their speech more to native English standards and, therefore, perceive their own accent to be as intelligible as native English. On the other hand, HP listeners show an ISIB-L in contrast to LP listeners, which is probably due to the fact that they scored at a similar level of RTs as native English listeners, while LP listeners were generally slower in understanding L2 speech. The ISIB-L for HP NNLs is presumably due to their high familiarity and experience with that accent. The fact that Hayes-Harb et al. (2008) only found evidence of the ISIB-L and, in contrary, Rasmussen (2007) only found support for the ISIB-T, could be due to differences in listeners' proficiency in their studies, which has been shown to effect intelligibility substantially in the present study. Both authors discussed that results would be clearer when listeners' proficiency would have been taken into account (see also van Wijngaarden et al., 2002b).

Regarding comprehensibility and accentedness and the **third research question,** it can be stated that the results were different for both perceptual dimensions. Comprehensibility results looked similar to those of intelligibility; German and Catalan LP listeners rated their own *accent* to be more comprehensible than the unfamiliar *accent*. Nevertheless, all listeners rated native English to be most comprehensible, even if the difference to the familiar *accent* was not significant for NNLs. HP listeners did not differ in their comprehensibility perception of non-native accents, but perceived both less comprehensible than native English. Results of LP listeners supported the evidence against the mismatched ISIB-L and ISIB-T. Nevertheless, comprehensibility ratings did not support the matched ISIB-T or ISIB-L, because for the English listeners group all *accents* were more comprehensible than for NNL groups and native English was always most comprehensible to all listeners.

L2 listeners' *proficiency* seems to be an important factor when it comes to comprehensibility and intelligibility, while it does not make a difference for accentedness (Smith, 1992). Accentedness seems to be more independent from *L1* background and L2 *profi-*

ciency than the other dimensions, as the ratings showed no effect for *proficiency* and the same tendency for all *L1s*, which is descending accentedness from Spanish-accented over German-accented to native English. All listener groups assigned generally harsher accentedness than comprehensibility ratings, which is a common finding in many studies (Derwing & Munro, 1997; Derwing, Munro, & Carbonaro, 2000; Munro & Derwing, 1995b, 1999, 2001).

It should be noted that NNLs of English appear to apply the examined dimensions scales in different ways than native English listeners, which suggests that different underlying processes due to *L1* background and L2 *proficiency* are at work in both types of listener. Although the strength of foreign *accent* is correlated with comprehensibility ratings and intelligibility scores, it does not necessarily reduce one of them. The correlation between intelligibility and comprehensibility showed clearly that less comprehensible utterances tended to be less intelligible than highly comprehensible utterances, which can be explained by influences of *L1* and *accent* on both dimensions. In accordance, Munro and Derwing (1995a) showed that although some utterances were highly intelligible and comprehensible, the accentedness ratings varied widely.

6 CONCLUSION

An examination of listeners' intelligibility, comprehensibility, and accentedness of L2 speech is important in L2 research, testing, and pedagogy.

In the field of language testing statements are made about how intelligible or comprehensible a particular learner's speech is. In this study it has been shown that the response to L2 speech differs due to the L2 proficiency and L1 of listeners. Therefore, it can be doubted that oral test scores can have predictive value. Nevertheless, these findings need to be replicated in other work involving different listening conditions and different stimuli types.

In foreign language acquisition it is commonly recommended that teachers focus on those specific aspects of the learner's speech that most affect comprehensibility and intelligibility rather on those aspects simply associated with accent or phonological and grammatical accuracy (Breitkreutz, Derwing & Rossiter, 2001; Derwing & Munro, 2005). Due to the low correlations between accentedness and the dimensions of intelligibility and comprehensibility it can be doubted that foreign accent reduction automatically leads to a higher intelligibility. There is a need for further research on which aspects may have the strongest impact in comprehensibility and intelligibility. Furthermore, in foreign language teaching contexts learners are exposed largely to English of their L1 speakers. Due to their ability to communicate successfully with them (which is due to the ability to adjust their production and compensate their perception on the basis of shared knowledge), they may be convinced of a false sense of their intelligibility, which may not be applicable to other NNSs or native English speakers. Thus, it is important that learners are exposed to a range of accents beyond their own and that pronunciation teaching eradicates disruptive features of their speech to facilitate international intelligibility (Derwing et al., 2002).

It has been proven that even an incorrect assumption that a speaker comes from a non-native background might sometimes reduce a listener's comprehension (Rubin, 1992), which might have been a factor in this study in some subjective ratings, although it could not have been a large factor given the correlation between ratings and intelligibility. Therefore, a bias against foreign-accented speech does not necessarily seem to interact with the underlying ability to comprehend accented speech. Nevertheless, this study has not addressed the problem of bias, which might be more likely to occur in real-life encounters. Likewise, sociolinguistic factors, which might interact in real communication

with the L1, were not taken into consideration and have to be investigated in further field studies.

In previous studies the use of different measures for intelligibility may be responsible for variation in findings. Different to most studies, which have used word identification as a measure of intelligibility (e.g., Bent & Bradlow, 2003), this research used RT. Because RT indirectly referred to processing time and it was controlled for linguistic content, the ISIB presumably arose from the differences in the acoustic signal and not from differences in lexical choices or syntactic structures. Therefore, it is very likely that the ISIB occurred at an early, phonetic stage of processing. Nevertheless, it possibly also operates at higher levels of processing, which needs to be investigated by further research.

Regarding the participants some limitations of the study have to be mentioned. There were only two proficiency groups investigated, which is why further research with intermediate proficient listeners is needed to explore the turning point from ISIB-T to ISIB-L. Additionally, research needs to investigate listeners from a wider variety of L1s, other TLs than English and a wider variety of data types. The fact that only three different L1 groups were represented (similar in age and academic background) may limit the study, because they may not be an accurate reflection of the L1 groups. Furthermore, the sample of speakers was very small after it was controlled for various influencing factors on intelligibility, which could cause a bigger influence of individual speech characteristics than expected.

Influences of speakers' proficiency, which was mostly operationalized as accentedness, on intelligibility were shown in various studies (van Wijngaarden, 2001; van Wijngaarden et al., 2002a; Bent & Bradlow, 2003; Stibbard and Lee, 2006). But only very few concerned the proficiency of the listener (e.g. van Wijngaarden et al., 2002b). The present study indicates that listeners' L2 proficiency is one of the main factors determining whether an ISIB-L or ISIB-T occurred. Evidence for the ISIB-T was provided for LP listeners and for the ISIB-L for HP listeners. In conclusion, the present study shows that a distinction between two types of the ISIB, which was often not made in previous studies, is essential to clarify the different influences of *L1* and *proficiency* on intelligibility of non-native speech.

REFERENCES

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38*, 561–593.

Anderson-Hsieh, J.; Johnson, R. & Koehler, K. (1992). The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentais, Prosody, and Syllable Structure. *Language learning, 42*, 529-555.

Bent, T. & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America 114 (3)*, 1600-1610.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In Goodman, J. & Nusbaum, H. C. (Eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In Strange, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 171-204). Baltimore: York Press.

Bradlow A. R. & Bent T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America 112*, 272-284.

Bradlow A. R. & Bent T. (2008). Perceptual adaptation to nonnative speech. *Cognition, 106*, 707–729.

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*, 2074–2085.

Bradlow A. R., Nygaard L. C. & Pisoni D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Attention, Perception, & Psychophysics 61 (2)*, 206-219.

Breitkreutz , J. A., Derwing, T. M, & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal, 19*, 51–61.

Burda, A. N., Overhake, D. R., & Thompson, K. K. (2005). Familiarity and older adults' transcriptions of native and nonnative speech. *Perceptual and Motor Skills, 100,* 939–942.

Burda, A. N., Scherz, J. A., Hageman, C. F., & Edwards, H. T. (2003). Age and understanding speakers with Spanish or Taiwanese accents. *Perceptual and Motor Skills, 97,* 11–20.

Clopper, C.; Pisoni, D. & Tierney, A. (2006).Effects of open-set and closed-set task demands on spoken word recognition. *Journal of the American Academy of Audiology, American Academy of Audiology, 17,* 331-349.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Crystal, D. (2003). *English as a Global Language* (second edition). Cambridge: Cambridge University Press.

Derwing, T. M. & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition, 19,* 1-16

Derwing, T. M. & Munro, M. J. (2001). What speaking rates do non-native listeners prefer? *Applied Linguistics, 22,* 324-337.

Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39,* 379–797.

Derwing, T. M.; Munro, M. J., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly, 34,* 592–603.

Derwing, T. M.; Rossiter, M. J.; & Munro, M. J. (2002). Teaching native speakers to listen to foreign- accented speech. *Journal of Multilingual and Multicultural Development, 23,* 245–259.

Ensz, K. (1982). French attitudes toward typical speech errors of American speakers of French. *The Modern Language Journal, 66,* 133-139.

Field, J. (2003). The fuzzy notion of 'intelligibility': A headache for pronunciation teachers and oral testers. *IATEFL Special Interest Groups Newsletter,* 34–38.

Field, J. (2004). An insight into listeners' problems: Too much bottom-up processing or too much top-down? *System, 32*, 363–377.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language Research* (pp. 233-277). Baltimore: York Press.

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

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*, 370–389.

Flege, J. E.; Munro, M. & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Forster, K. & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, 35*, 116-124.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. The Journal of the Acoustical Society of America 27, 765–769.

Gass, S. & Mackey, A. (2007). *Data elicitation for second and foreign language research.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. & Varonis, E. (1984).The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34*, 65-87.

Goldinger, S. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166-1183.

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

Hecker, M. H. L., Stevens, K. N., and Williams, C. E. (1966). Measurements of reaction time in intelligibility tests. *The Journal of the Acoustical Society of America*, 39, 1188-1189.

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

Ingram, J. & Nguyen, T. (2007). Vietnamese accented English: Foreign Accent and Intelligibility judgement by listeners of different language backgrounds. *Proceedings of a TESOL in the internationalization of higher education in Vietnam Conference. Hanoi, Vietnam, May 12, 2007*

Iverson, P.; Kuhl, P.; Akahane-Yamada, R.; Diesch, E.; Tohkura, Y.; Kettermann, A. & Siebert, C. (2003).A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition 87 (1),* B47-B57.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*, 83–103.

Kennedy, S. & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review/La Revue canadienne des langues vivantes, 64*, 459-489.

Kent, R.; Miolo, G. & Bloedel, S. (1994). The intelligibility of children's speech: A review of evaluation procedures. *American Journal of Speech-Language Pathology, ASHA, 3*, 81-95.

Kenworthy, J. (1987). *Teaching English pronunciation*. Longman Harlow.

Klein, E. & Flint, C. (2006). Measurement of intelligibility in disordered speech. *Language, Speech, and Hearing Services in Schools, 37*, 191-199.

Kosinski, R. (2010). *A literature review on reaction time*. Retrieved from: http://biology.clemson.edu/bpc/bp/Lab/110/reaction.htm [05/23/2012].

Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States.* London: Routledge.

Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics, Elsevier, 26*, 381-400.

Major, R.; Fitzmaurice, S.; Bunta, F. & Balasubramanian, C. (2002). The effects of non-native accents on listening comprehension: Implications for ESL assessment. *TE-SOL Quarterly 36*, 173–190.

Mauranen, Anna (2009). Spoken rhetoric: How do natives and non-natives fare? In Suomela-Salmi, Eija & Dervin, Fred (Eds.), *Cross-linguistic and Cross-cultural Perspectives on Academic Discourse (pp. 199-218).* Amsterdam: John Benjamins.

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

Meara, P. (2005). Designing vocabulary tests for English, Spanish and other languages. In Butler, C. S.; Gómez-González, M. & Doval-Suárez, S. M. (Eds.), *The Dynamics of Language Use* (pp.271-285). Amsterdam: Benjamins.

Meara, P. & Milton, J. (2003). *X_Lex, The Swansea Levels Test*. Newbury, UK: Express Publishing.

Milton, J. (2006). Language lite? Learning French vocabulary in school. *Journal of French Language Studies, 16 (2)*, 187-205.

Milton, J. (2007). Lexical profiles, learning styles and the construct validity of lexical size tests. In Daller, H., Milton, J. & Treffers-Daller, J. (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 47-58). Cambridge, UK: Cambridge University Press.

Milton, J. & Alexiou, T. (2009). Vocabulary size and the Common European Framework of Reference for Languages. In Richards, B.; Daller, H. M.; Malvern, D.; Meara, P.; Milton, J. & Treffers-Daller, J. (Eds.), *Vocabulary Studies in First and Second Language Acquisition: The interface between theory and application* (pp. 194-211). Basingstoke: Palgrave.

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

Munro, M. & Derwing, T. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning, 45*, 73-97.

Munro, M. & Derwing, T. (1995b) Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*, 289-306.

Munro, M. & Derwing, T. (1998). The Effects of Speaking Rate on Listener Evaluations of Native and Foreign-Accented Speech. *Language Learning, 48*, 159-182.

Munro, M. & Derwing, T. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*, 285–310.

Munro, M. & Derwing, T. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition, 23*, 451-468.

Munro, M.; Derwing, T. & Morton, S. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*, 111-131.

Nemati, A. (2010). Active and passive vocabulary knowledge: The effect of years of instruction. *The Asian EFL Journal Quarterly, 12*, 30-46.

Nygaard, L.; Sommers, M. & Pisoni, D. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*, 42-46.

Rasmussen, Z. B. (2007). The interlanguage speech intelligibility benefit: Arabic-accented English. *Unpublished senior honours thesis.* Utah: University of Utah.

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

Rubin , D. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education, 33*, 511–531.

Schmied, J. (1991). *English in Africa: An Introduction*. London: Longman.

Schüppert, A. & Gooskens, C. (2011). Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia, 7*, 119-140.

Smith, B. L.; Bradlow, A. R.; & Bent, T. (2003). Production and perception of temporal contrasts in foreign-accented English. In Sole, M. J.; Recasens, D.; Romero, J., (Eds.). *Proceedings of the XVth international congress of phonetic sciences* (pp. 519-522). Barcelona, Spain: Causal Productions.

Smith, D. B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition, 26*, 365-398.

Smith, L. (1992). Spread of English and issues of intelligibility. In Kachru, B.B. (Ed.). *The other tongue: English across cultures* (pp.75-90). Urbana and Chicago: University of Illinois Press.

Smith, L. & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly, 13*, 371–380.

Stibbard, R. & Lee, J. (2006). Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. *The Journal of the Acoustical Society of America, 120*, 433-442.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-45). Baltimore: York Press.

Tajima, K.; Port, R. & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25*, 1-24.

Taylor, D. (1996). Demystifying word stress. *English today, 48 (12)*, 46-52.

Trofimovich, P. & Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*, 1-30.

van Wijngaarden, S. J. (2001). The intelligibility of non-native Dutch speech. *Speech Communication, 35*, 103–113.

van Wijngaarden, S. J., Steeneken, H. J. M. & Houtgast, T. (2002a). Quantifying the intelligibility of speech in noise for non-native talkers. *Journal of the Acoustical Society of America 111 (4)*, 3004–3013.

van Wijngaarden, S. J., Steeneken, H. J. M., and Houtgast, T. (2002b). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America 111 (4)*, 1906–1916.

Weinreich, U. (1953). *Languages in contact: Findings and problems*. The Hague, Mouton.

Zielinski , B (2004). Measurement of speech intelligibility: What are we actually measuring? *Paper presented at the annual meeting of the American Association for Applied Linguistics*, Portland, OR.

APPENDIX

APPENDIX 1

FACTORS INFLUENCING THE INTELLIGIBILITY OF SPEECH

SPEAKERS' FACTORS:	
speech rate	controlled for after recording (nearly the same for all chosen speech samples)
pausing	controlled for after recording (nearly the same for all chosen speech samples)
stress	controlled for after recording (nearly the same for all chosen speech samples)
errors	controlled for after recording (no errors in chosen speech samples)
intonation	controlled for after recording (speech samples with falling intonation were chosen)
clear speech	controlled for after recording (only speakers with clear speech were chosen)
neighbourhood density	controlled for before recording (the same for all words and sentences)
unique speech habits	controlled for after recording (only speakers with less unique speech habits were chosen)
word frequency	controlled for before recording (words and sentences are high frequent (withing the first 2000))
ascendancies of context	controlled for before recording (no context for words, for sentences the infuence of context is nearly the same)
non-native accent	**focus of this study -> native English, German- and Catalan-accented English (same degree of accent)**

LISTENERS' FACTORS:	
experience / previous exposure to L2 speech	controlled for in selection of participants (personal questionnaire)
proficiency in TL	**focus of this study -> different proficiency levels (vocabulary size test)**
usage of context	same context for all of them
familiarity of topic or speech event	same familiarity for all of them (unknown topic and speech event)
level of tiredness	controlled for with randomization of speech samples presented
situation specific factors	same situtation for everybody
semantics (for NNSs)	for sentences: real-world expectations (same for everybody)
background noise	no noise for intelligbility test and accentedness rating but same degree of noise for comprehensibility test
familiarity of native speech / non-native accents	**focus of this study -> more familiar with accent of speakers who share the same L1**

APPENDIX 2

SPEECH MATERIAL - WORDS

ANIMATED WORDS	NON-ANIMATED WORDS
1. Chicken	1. Chair
2. Monkey	2. Pencil
3. Zebra	3. Picture
4. Bunny	4. Pistol
5. Bird	5. Radio
6. Horse	6. Rope
7. Dog	7. Scissors
8. Snake	8. Spoon
9. Giraffe	9. Telephone
10. Elephant	10. Toothbrush

APPENDIX 3

SPEECH MATERIAL - SENTENCES[7]

L1	TRUE SENTENCES		FALSE SENTENCES	
	FEMALE SPEAKER	MALE SPEAKER	FEMALE SPEAKER	MALE SPEAKER
L1 ENGLISH	*Some people like to watch television.*	Most people have scissors.	Some people have fourteen eyes.	*Most dolphins have fourteen eyes.*
	Some people like to smoke cigars.	Many children like dogs.	Most horses like to smoke cigars.	Some people walk on their lips.
	Many children like to watch television.	Some people like their lips.	Most dolphins enjoy playing basketball.	*Most people wear watches on their lips.*
	Some people enjoy watching basketball.	Some people have many horses.	Chocolate is a healthy food to eat.	Some people have fourteen eyes.
	Many children enjoy playing basketball.	Some people like their ears.	Many dogs walk on their ears.	Many children like to eat scissors.
L1 GERMAN	Some people like to eat chocolate.	Some people like dogs.	Most dolphins like to smoke cigars.	Many dogs like to eat chocolate.
	Many children like to eat chicken.	Some people enjoy cigars.	Most horses like to eat chocolate.	Many children like to eat cigars.
	Some people enjoy playing basketball.	Most horses live on dry land.	Most people wear watches on their ears.	Many dogs wear watches on their lips.
	Some people like playing basketball.	Some people walk on dry land.	Most horses enjoy playing basketball.	Most horses walk on their ears.
	Some people like to wear watches.	Most horses walk on dry land.	Many dogs enjoy playing basketball.	Most horses like to eat cigars.
L1 SPANISH	Some people like to watch basketball.	Many dogs live on dry land.	Most horses have fourteen eyes.	*Many dogs like to smoke cigars.*
	Some people smoke many cigars.	Chicken is a healthy food to eat.	*Most dolphins live on dry land.*	Most horses like to watch television.
	Many children like to eat chocolate.	*Salad is a healthy food to eat.*	*Some people walk on their ears.*	Many children like to smoke cigars.
	Some people like healthy food to eat.	Some people have many dogs.	Most dolphins like to watch television.	Most horses wear watches on their lips.
	Some people enjoy watching television.	Some people like to eat chicken.	Many dogs have fourteen eyes.	Most dolphins like to eat chocolate.

[7] The nine originally recorded sentences are italicized.

APPENDIX 4

BACKGROUND QUESTIONNAIRE

1. What is your sex?
   - ☐ Male
   - ☐ Female

2. How old are you? _____

3. What is the highest degree or level of school you have completed? If currently enrolled, mark the previous grade or highest degree received.
   - ☐ No schooling completed
   - ☐ High school graduate
   - ☐ Bachelor's degree
   - ☐ Master's degree
   - ☐ Doctorate degree

4. Are you currently...?
   - ☐ Employed for wages
   - ☐ Self-employed
   - ☐ Out of work
   - ☐ A student
   - ☐ Retired

5. Where were you born?
   - Country:
   - City/region:

6. How long did you live in your home country? _____ years

7. Did you also live in an English speaking country? If so, how long and what language was most important there for your daily life?

   ____ years / month in _____ Language:_____

   ____ years / month in _____ Language:_____

   ____ years / month in _____ Language:_____

8. What is your native language?

9. Could you put all the languages you have ever learnt / spoken from the one you have spoken first to the one you have learnt recently into the order?
    1.
    2.
    3.
    4.
    5.
    6.

10. Could you put all the language you have ever learnt / spoken from most proficient to less proficient into the order? And estimate your knowledge from 1 (no knowledge anymore) to 5 (native-like proficiency).

    1. 1 – 2 – 3 – 4 – 5
    2. 1 – 2 – 3 – 4 – 5
    3. 1 – 2 – 3 – 4 – 5
    4. 1 – 2 – 3 – 4 – 5
    5. 1 – 2 – 3 – 4 – 5
    6. 1 – 2 – 3 – 4 – 5

11. How many years have you learnt English? _____ years

12. Are you used to listen to English with a Spanish or Catalan accent?

Not at all		From time to time		Very much
1	2	3	4	5

13. Are you used to listen to English with a German accent?

Not at all		From time to time		Very much
1	2	3	4	5