The speech intelligibility of English learners of Spanish at Key Stage 4.

Osle Ezquerra, Ángel

The speech intelligibility of English learners of Spanish at Key Stage 4

by

Ángel Osle Ezquerra

Submitted in partial fulfilment of the requirements
of the Degree of Doctor of Philosophy

April 2013

School of Languages, Linguistics and Film
Queen Mary, University of London

Declaration

I hereby declare that this thesis and the materials presented herein have not previously been submitted to this or any other institution, and that they are solely the product of the present author. The copyright of this thesis rests with the author and any information derived from it should be acknowledged.

Date:

Signature:

And the Gileadites took the passages of Jordan before the Ephraimites: and it was so, that when those Ephraimites which were escaped said, Let me go over; that the men of Gilead said unto him, Art thou an Ephraimite? If he said, Nay; Then said they unto him, Say now Shibboleth: and he said Sibboleth: for he could not frame to pronounce it right. Then they took him, and slew him at the passages of the Jordan: and there fell at that time of the Ephraimites forty and two thousand.

(Judges 12: 5-6)

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Christopher Pountain, for his encouragement, continuous advice and insightful comments throughout all stages of this study.

I would also like to thank my examiners, Dr Michael Evans and Dr María Elena Placencia, for their thoughtful and constructive feedback during the examination process.

I would also like to express my gratitude to all the teachers, students and evaluators that participated in this study both in England and Spain. This project would not have been possible without their cooperation.

Nos gustaría dar las gracias a David Rodríguez por la ayuda prestada en el análisis estadístico de los datos y en la trascripción de las muestras de habla.

Igualmente, queremos dar las gracias a la Dr. Elina Vilar por los detallados comentarios a uno de los capítulos de esta tesis doctoral.

Asimismo, nos gustaría agradecer al Profesor Joaquim Llisterri de la Universitat de Barcelona la ayuda prestada en la localización de algunos de los estímulos utilizados en la parte experimental de este estudio.

Gracias también a mi familia en España, en especial a mi madre, por el apoyo incondicional a lo largo de los años.

Last but not least, I would like to thank Richard P. Howard for his unconditional support. Richard, thanks for your patience, understanding and for so many other things that I cannot even put into words. I could never have done this without you!

Abstract

This study offers an assessment of the non-native speech intelligibility of a group of English learners of Spanish at word level and in connected speech. Specifically, we aimed at analysing the impact of certain categories of phonemic errors, as well as three temporal variables of L2 speech (speech rate, pause frequency and pause duration) on intelligibility scores. In addition, the possible correlation between degree of intelligibility and certain individual factors (gender, level of proficiency, motivation, aptitude and L1) was also studied.

Sixty evaluators, native speakers of Peninsular Spanish, transcribed different speech samples belonging to a group of 20 Key Stage 4 English learners of Spanish. The transcription of the different speech samples served to assess intelligibility at word level and in connected speech (sentence, passage and semi-spontaneous production). Results revealed an intelligibility loss at all levels of analysis, as well as a high correlation between intelligibility scores in the single word test and those obtained in connected speech. At a segmental level, deviations affecting vowels, especially unstressed vowels, seemed to play a more important role than inaccuracies affecting consonants. Moreover, correlation analyses underscored the importance of speech rate, pause frequency and pause duration for intelligibility loss. The predictability of our multiple-regression models was high for speech samples obtained at sentence and passage levels. However, multiple-regression models for speech samples obtained through the semi-spontaneous production task exhibited a more limited capability in predicting variation in students' intelligibility scores. Results suggest the existence of additional variables affecting intelligibility at this level of analysis.

All individual differences under study, with the exception of gender, were highly correlated with speech intelligibility.

From a pedagogical perspective, it is argued here that any successful instructional treatment of speech intelligibility will depend on an appropriate integration of temporal aspects of speech within the time devoted to pronunciation instruction in the foreign language classroom.

Contents

Tables

Figures

Chapter 1: Introduction

It is widely accepted, within communicative language teaching, that intelligibility represents a sufficient goal for pronunciation instruction (see e.g. Derwing 2008; Kenworthy 1996; Levis 2005; Morley 1991; Munro 2008; Munro and Derwing 2011). Attainment of a native accent, considered as unrealistic for those supporting the existence of a critical or sensitive period in Second Language Acquisition (SLA)[1], has been replaced by the 'intelligibility principle', which posits that communicative effectiveness is a sufficient goal in itself (Levis 2005). As Kenworthy (1996: 13) notes, 'very few teachers today would claim that a pronunciation that is indistinguishable from that of a native speaker is necessary or even desirable for their learners. Instead, it is generally accepted that intelligibility is the most sensible goal'. Furthermore, research has shown that 'accentedness' and intelligibility are not necessarily correlated (Derwing and Munro 2009). In other words, a heavy accent does not always imply a reduction in speech intelligibility. On this issue, Rogers (1997) makes a useful distinction between 'deviance' and 'recognizability', which serves to illustrate the idea that not all deviations from the speech norm result in a lack of perception of the linguistic unit.

The acceptance of the *intelligibility principle*[2] seems to suggest that pronunciation instruction should be firmly grounded in empirical research relating, first and foremost, to speech intelligibility. However, it is undeniable that there exists a great deal of anecdotal and theoretical commentary regarding the teaching of pronunciation that has not been submitted to empirical verification. As Munro and Derwing (2011: 317) rightly put it, 'what has been missing until recently is, first, a conceptualisation of intelligibility that assists teachers in setting priorities and, second, empirical evidence that identifies effective practices'.

[1] In the early 1980s, Krashen (1981) established a distinction between 'acquisition' and 'learning' by suggesting that the former refers to unconscious processes of which an individual is not aware, while the latter points to conscious processes that take place within the framework of formal instruction. This series of conscious processes allow students to achieve an explicit knowledge of grammatical rules and linguistic use, as well as the ability to verbalise this knowledge. The dominant trend in contemporary SLA is to use 'language acquisition' to refer to processes related to both acquisition and learning.

[2] The importance of the notion of intelligibility within Applied Linguistics has not been accepted unanimously (see, for example, Rajagopalan 2010 and a subsequent response by Munro 2011).

Weismer (2008: 2) defines *intelligibility* as the 'relative measure of the degree to which a speaker's speech signal is understood, the relativity depending at a minimum on the identities of the speaker and listener, what is spoken and where it is spoken'. In spite of the seemingly straightforward construct, it has proved challenging to offer a clear definition of speech intelligibility. Different disciplines have used this term in very diverse manners (Rogers 1997). In the field of SLA, Derwing and Munro (2009: 478-479) conceptualise intelligibility 'as the degree of a listener's actual comprehension of an utterance'. This term differs from *comprehensibility,* which focuses on how easy or difficult it is for a listener to understand an utterance. Moreover, a further distinction between these two dimensions and *'accentedness'* must be established. *'Accentedness'* alludes to 'how different a pattern of speech sounds compared to the local variety' (see Chapter 2 for alternative definitions and a detailed analysis of these dimensions).

The complicated task of defining speech intelligibility becomes even more challenging if one considers the intrinsic complexities of oral discourse. Oral discourse is characterised by negotiation of meaning, interactivity, bi-directionality and the importance of the communicative situation (Cantero 1998). It is clear that oral communication possesses a phonetic dimension, given that it cannot exist without the articulation of speech sounds. This is a complex process that cannot be reduced to the utterance of sounds in a linear and isolated fashion (Clark and Yallop 1990).

Pronunciation itself is characterised by a high degree of variability. Speech production may vary depending on numerous objective (sex, age, vocal apparatus, geographical and social situation, etc.) and subjective variables (the speaker's profile, his place in the communicative situation etc.). Speech perception, on the other hand, may depend, for instance, on an individual's hearing or conversational expectations. Authors such as Brown (1991), Cantero (2003) or Iruela (2004) have extended the concept of pronunciation to include not only speech production but also perception of sounds, accent and intonation. In addition to its importance within the communicative process[3], pronunciation provides information about a speaker's identity, as well as his

[3] The importance of mastering L2 pronunciation is stressed by Iruela (2007: 15) when analysing its relevance in connection with the language activities mentioned by the Common European Framework of Reference for Languages. This author concludes that: 'la competencia fónica forma parte de las actividades orales de la lengua: la comprensión auditiva, la expresión oral y la interacción oral. La pronunciación es el soporte de la lengua oral, tanto en su producción como en su percepción, lo que hace que otorgue inteligibilidad a la producción oral del aprendiente y le facilite la comprensión auditiva'.

or her geographical and social origin. Furthermore, as Abercrombie (1967: 9) notes, pronunciation can reveal changes in a speaker's state of mind through certain 'indexical features which are not present all the time in a person's pronunciation, but come and go according to his physical or mental state'. Following on from this, one can say that the acoustic signal not only provides linguistic but also extralinguistic and paralinguistic information (see e.g. Abercrombie 1968 for discussion). Extralinguistic information belongs to a speaker's personal sphere and can be either transitory or permanent. Extralinguistic features provide information on a speaker's identity, sex and quality of voice, as determined by the special characteristics of his or her articulatory system. Furthermore, the acoustic signal provides paralinguistic [4] information, which serves to convey a speaker's different emotions and mental states and to regulate the rhythm of conversation. Unlike extralinguistic features, paralinguistic features are used intentionally by the speaker. As far as linguistic information is concerned, one can establish a division between a first level of analysis integrated by the phonetic units of speech, and a prosodic level, which points to the systematic use of variations in tone, duration and intensity. It must be borne in mind that the distinction between linguistic, extralinguistic and paralinguistic features is not as straightforward as it may seem and, in fact, some authors (e.g. Clark and Yallop 1990; Crystal and Quirk 1964; Gil 2007) advocate the idea of a gradual continuum between the three dimensions[5].

Given the importance of pronunciation in the communicative process and the relevance of the 'intelligibility principle' for pronunciation instruction, it is rather disappointing to see the paucity of systematic investigation that this issue has generated within SLA. This is especially evident if one looks at empirical research conducted outside the area of English as a Second Language (ESL), or even if we compare it to the great deal of attention that intelligibility studies have received in the field of speech pathology. The potential benefits of research centred on speech intelligibility and the reasons behind the present study are set out below.

[4] As Abercrombie (1968: 55) points out, the term *paralanguage* can be potentially misleading, since 'it can give the impression that, because there exists a (more or less) homogeneous entity called language, there must be, existing beside it, a comparably homogeneous entity called *paralanguage*. I believe this is not so…These non-verbal, though conversational activities to which the word paralanguage refers are far too diverse, too little codified, too uninvestigated and too insufficiently understood, to be given the air of unity which a noun confers on them'.

[5] It is important to note that the terminology varies from author to author. See e.g. Schoetz (2002) for a review.

1.1 Justification of this study

In the last few decades, there has been a clear increase in the number of studies devoted to Second Language[6] (L2) phonology. Several models have attempted to account for how L2 listeners tackle the difficult task of speech perception. Three of these models have received a great deal of attention in the literature: Flege's (1995, 1999 and 2002) Speech Learning Model (SLM), Best's (1995) Perceptual Assimilation Model (PAM) and Kuhl's (1993, 2008) Native Language Magnet (NLM). All of them have in common the idea that the First Language (L1) sound system plays a major role in L2 phonological acquisition. Furthermore, they seek to predict the degree of difficulty in the acquisition of L2 sounds on the basis of the learner's L1 sound system[7].

Flege's (1995, 1999 and 2002) SLM tries to account for how learners succeed or fail in perceiving and producing phonetic segments in a foreign language. The SLM puts forward the idea that L2 learners are, in principle, able to distinguish the phonetic properties of L2 sounds, even though the process of 'learning' these phonetic properties is lengthy and heavily influenced by variables such as quality and quantity of input. This model proposes that the mechanisms that help acquire L1 speech remain accessible to the L2 learner. Furthermore, L1 and L2 phonetic elements coexist in a 'common phonological space' and influence each other. The SLM has generated several hypotheses that are continually being tested. Among these, Flege (1995) has hypothesised that a new phonetic category is more likely to form when there is a great dissimilarity between the L2 sound and the closest L1 sound. Moreover, if a new phonetic category cannot be formed due to proximity between L1 and L2 sounds, assimilation of L1 and L2 categories will ensue.

---

[6] The terms 'foreign language' and 'second language' refer to the acquisition of a language other than the native language. A distinction can be established between a second language, which is learned in an area where the language is spoken, and a foreign language that is learned in a context where it is not used as an official or co-official language. This study focuses on the acquisition of Spanish by English native speakers in a secondary school setting in England, and, therefore, examines the acquisition of Spanish as a foreign language. Nevertheless, it is important to indicate that the term "second language" will be used here as a cover term to refer to an additional language learnt in any contextual situation. See Ortega (2009) for a discussion on the dangers of using the dichotomy L1-L2.

[7] Additional models of L2 speech perception have been put forward, among others, by Brown (1998), Escudero and Boersma (2004) and Major (1987). The vitality of research in this area is also highlighted by the number of literature reviews available (see e.g. Escudero 2007; Strange and Shafer 2008).

Best's (1995) PAM assumes that 'when listening to an unfamiliar non-native phone (phonetic segment), naïve listeners are likely, due to their native language experience, to perceptually assimilate the non-native phone to the most articulatory similar native phoneme' (Best and Tyler 2007: 22). Best's model is based on direct realism and articulatory phonology. Perceptual similarity is therefore understood in terms of articulatory information. With regard to the patterns of assimilation of L2 segments to L1 sounds, this model posits that L2 sounds can be perceived as exemplars of non-linguistic sounds (e.g. clicks), as uncategorised sounds, i.e. recognised as speech sounds but not related to an L1 category, or as exemplars of L1 sounds (as either good, acceptable or notably deviant exemplars of that category). In terms of discrimination between L2 contrasts, this model hypothesises that discrimination should be excellent if both L2 phones are assimilated to different L1 categories (Two-Category Assimilation). It also predicts a moderate level of discrimination when both L2 phones are assimilated to a single L1 category but are perceived as different good exemplars of this category (Category-Goodness Difference). Serious discrimination problems will surface if both L2 phones are perceived as good exemplars of a single L1 category (Single-Category Assimilation)[8].

Kuhl's (1993, 2008) Native Language Magnet (NLM) was initially developed to account for L1 acquisition. This model is based on the idea that speech perception is not governed solely by functional principles of the auditory system, but also by perceptual prototypes stored in the listener's long-term memory. These categories are organised around abstract representations, i.e. prototypes, which include the most representative features of a specific category. According to this theory, listeners make comparisons between the incoming acoustic signal and their internal prototypical representations. If the signal is sufficiently similar to the prototype, then it is classified as a member of that specific phonetic category. Depending on the degree of similarity, some stimuli will be considered as better exemplars of a category than others. The NLM posits that prototypes act as 'magnets' by attracting the perceptual space around them. Intracategorical differences are reduced as a result of the 'shrinking' of the perceptual space. On the other hand, intercategorical differences are emphasised due to the enlargement of the perceptual space that can be found between different

[8] A detailed analysis of the PAM and SLM, as well as an attempt to extend the PAM beyond naïve listeners can be found in Best and Tyler (2007).

prototypes. The establishment of the aforementioned prototypes in the child's mind arises as a consequence of the emergence of perceptual maps, at a neurological level, after exposure to the specific distributional patterns of the speech signal. The difficulty for L2 learners stems, therefore, from the existence of a specific L1 filter that distorts the perceptual space surrounding the prototypes. In other words, the shape of the L1 mappings seem to act as a filter that renders the acquisition of L2 sounds much more difficult. It is important to note that this model does not provide any explanation as to how new perceptual mappings are established during the L2 learning process.

As far as L2 speech production is concerned, studies have largely focused on L1 and L2 phonetic differences with regard to the production of vowels (e.g. McAllister, Flege and Piske 2002), stops (e.g. Flege 1987, 1991; Schmidt and Flege 1996) and liquids (e.g. Major 1986). Some studies have examined the substitution patterns exhibited by L2 learners (Eckman 1977; Lombardi 2003). On a suprasegmental level, research has focused on L2 learners' deviations in terms of syllabic structure (e.g. Broselow and Finer 1991; Broselow, Chen and Wang 1998) and stress (e.g. Archibald 1994). It has also served to uncover the influence of certain universal developmental factors such as the Universal Canonical Syllable Structure, Minimal Sonority Distance Parameter or the Interlanguage Structural Conformity Hypothesis, to name a few (see e.g. Iruela 2004 for a review). Furthermore, results are regularly used to emphasise the influence of the L1 or in support of specific models of speech perception and production.

It is difficult, however, to see how this surge of research on L2 speech perception and production can provide useful insights immediately applicable by teachers and instructors in a classroom setting. Furthermore, the majority of studies have been carried out with subjects from different linguistic backgrounds that are learning English as a Second Language. Therefore, data on the acquisition of additional languages is more limited. Nevertheless, it is undeniable that instructors could use findings related to the production difficulties of certain L2 sounds when implementing their corrective techniques. Moreover, the issue of the relationship between speech production and perception could obviously have a major impact on the teaching of pronunciation. If perception and production are interrelated and perception is considered to precede production, then specific perceptual work should be introduced before students are asked to focus on production. Llisterri (1995)

carried out a review on experimental studies dealing with the relationship between perception and production and concluded that a straightforward relationship between both dimensions cannot be claimed. In fact, this relationship is quite complex and seems to be influenced by a variety of factors, e.g. training, L2 experience, contextual variables, social factors, linguistic similarity etc. On this issue, Leather (1999) points out that the 'mutually facilitative' relationship between perception and production is, however, not constant and does not appear under all contextual circumstances. Escudero (2007) favours the priority of perception over production and criticises, on methodological grounds, those empirical studies that have shown otherwise (it is indeed difficult to find tasks that measure both production and perception in an equivalent manner). This author also notes that 'the difficulty that adult learners experience producing L2 sounds has a perceptual basis, such that incorrect perception leads to incorrect production' (Escudero 2007: 111). Regardless of whether we accept the priority of perception over production, it seems evident that both are intrinsically related and, thus deserve to receive appropriate attention in the foreign language classroom.

The aforementioned analysis underscores that the surge in the number of studies exploring the different dimensions of L2 phonological acquisition has not necessarily translated into more research focusing specifically on intelligibility issues. As noted above, in spite of the widely accepted *intelligibility principle* (Levis 2005) for pronunciation instruction, very little has been done to actually get a clear picture of the variables that may affect L2 speech intelligibility (some exceptions can be found in Hahn 2004; Jenkins 2000; Zielinski 2008). In fact, we agree with Derwing (2008) and Derwing and Munro (2009) in pointing out that a more detailed exploration of those factors that could reduce speech intelligibility is required. This is especially true in the case of Spanish as a Foreign Language, since the overwhelming majority of research in pronunciation instruction and L2 phonological acquisition has been undertaken within the field of English as a Second Language.

This lack of empirical research extends to the general area of pronunciation instruction. In the history of foreign language teaching, the importance attached to pronunciation instruction has greatly varied over the years. Grammar-translation methods largely disregarded pronunciation issues, while audiolingual methodologies considered pronunciation as an extremely important component. After the demise of the audiolingual method, pronunciation was largely forgotten during the 70s and 80s.

However, one can say that in the past two decades pronunciation has enjoyed a certain pedagogical revival. In spite of this, pronunciation instruction still remains the least developed area within the communicative approach. Furthermore, the constant changes experienced in the field of foreign language didactics during the last thirty years have not necessarily impacted the field of pronunciation instruction [9]. Consequently, teachers find it very difficult to integrate the teaching of pronunciation in the foreign language classroom. Furthermore, the lack of appropriate teaching materials does not help in improving this situation. In fact, it is widely acknowledged that current textbooks do not pay enough attention to pronunciation issues and, when they do, they seem to do so in an isolated way (Torres 2006). Derwing and Munro (2005) argue that, as a consequence of its marginalised status, many teachers do not receive any training on how to teach pronunciation. Studies confirming this point can be found in Breitkreutz, Derwing and Rossiter (2002), Lambacher (2001) or MacDonald (2002).

The marginalisation of pronunciation instruction can also be observed if we look at the amount of research conducted on this issue in the field of Applied Linguistics. Deng *et al*. (2009) reviewed 14 academic journals in search of articles devoted to pronunciation issues. They calculated the percentage of pronunciation articles for the past 10 years and concluded that, in spite of repeated calls for an increase of research in this area, pronunciation receives little attention in SLA studies. This marginal status is also confirmed if we look at research assessing the effects of pronunciation instruction. While numerous reviews can be found on the effects of formal instruction on several areas of language acquisition, including syntax, morphology, vocabulary and even pragmatics (e.g. De Graaff and Housen 2009; Doughty 2003), literature reviews on the effects of formal instruction related to the acquisition of pronunciation are very scarce (notable exceptions can be found in Barrera Pardo 2004 and Elliott 2003). It seems to be the case that empirical research is not abundant and consistent enough to definitely show the importance of specific

[9] Iruela (2007: 2) notes the following on this issue: 'A pesar de que cumple un papel relevante en la comunicación oral, la atención a la pronunciación ha sido relegada de la enseñanza de L2 durante largo tiempo. Y mientras la didáctica ha evolucionado constantemente y ha incorporado nuevas perspectivas, la enseñanza de la pronunciación se ha mantenido inalterada durante décadas. Nos parece necesario que la pronunciación adopte, en la medida en que sus singularidades lo permitan, los principios y las prácticas que se han aplicado en los últimos años en la enseñanza de L2. Para lograrlo, es fundamental que asuma la visión que en nuestros días se tiene de la lengua y de su aprendizaje'.

instructional techniques for the acquisition of pronunciation. Certain studies have not confirmed the benefits of pronunciation instruction altogether. Yule and MacDonald (1995) examined the effects of four different types of activities on the acquisition of L2 pronunciation and concluded that individual differences seem to play a far more crucial role than instruction itself. Purcell and Suter (1980) concluded that variables accounting for the effects of formal instruction did not seem to be significantly correlated with good pronunciation. Hardy (1993) underlined that the effects of instruction are beneficial but only on a short-term basis. On the other hand, Couper (2006: 46), in one of the few studies focusing on both short and long-term effects of formal instruction, concluded that 'appropriately focused instruction can lead to changes in learners' phonological interlanguage even where this may appear to have fossilised'. More recently, Saito and Lyster (2012) assessed the effects of Form-Focused Instruction (FFI) and corrective feedback (CF) on the acquisition of L2 pronunciation. Specifically, they focused on the acquisition of /ɹ/ by Japanese learners of English. Participants were divided into three groups: one receiving FFI plus CF, another one FFI and a control group that received instruction but without FFI. Results revealed statistically significant changes in F3 values for the FFI+CF group. The other two groups, i.e. control group and FFI only, did not show any statistically significant changes in F3 values as revealed by the corresponding acoustic analyses. If we turn our attention to the teaching of Spanish pronunciation, studies reporting gains, in either production or perception, as a result of explicit instruction can be found, among others, in Castino (1996); Elliott (1997); González Bueno (1997); Lord (2005).

As far as the impact of formal instruction on intelligibility, comprehensibility or 'accentedness' is concerned, the number of studies assessing improvements in these dimensions of L2 speech is very scarce, especially when compared to the amount of research devoted to measuring improvements in perception or production of L2 sounds. Perlmutter (1989) examined a group of ESL students over a six month period following a course that placed specific emphasis on pronunciation. Results showed significant improvements in speech intelligibility. However, the absence of a control group makes it difficult to determine if those gains can be attributed to formal instruction or to other variables. Moreover, no information was provided on the nature of the teaching programme. Dowd, Smith and Wolfe (1997) examined the production of French vowels by inexperienced English native speakers after receiving

pronunciation training either through a traditional method of pronunciation instruction or through real-time visual feedback. Results revealed that the group that had received visual feedback reached significantly higher levels of intelligibility in their oral productions. Derwing, Munro and Wiebe (1997) found that a 12-week teaching course emphasising prosodic features and general speaking habits resulted in improvements in both intelligibility and comprehensibility for a group of ESL learners. In a subsequent study (Derwing, Munro and Wiebe 1998), these authors assessed three groups of ESL learners following three different pronunciation treatments: pronunciation instruction focused on segmental features, instruction focused on prosodic features or no special focus on pronunciation instruction. Elicited speech samples consisted of a series of sentences and narratives. Results revealed significant gains at the sentence level, in terms of comprehensibility and 'accentedness', for those students who had received explicit pronunciation instruction. However, only those students who had followed instruction on suprasegmentals and speaking habits reported an improvement in fluency and comprehensibility. Derwing, Munro and Wiebe (1998: 406) explain the aforementioned results by stating that 'speakers who had had instruction emphasizing prosodic features such as rhythm, intonation, and stress could apparently transfer their learning to a spontaneous production'. Burleson (2007) provided computer-assisted instruction to a group of 5 Mandarin learners of English regarding the production of 6 phonemic contrasts. Instruction resulted in significant improvements in intelligibility as judged by a group of native English speakers (correct identification of the selected segments rose from 50% to 89%). Trofimovich *et al.* (2009), on the other hand, examined the effects of a comprehension-based programme, as opposed to a regular language learning programme, on the degree of comprehensibility, fluency and 'accentedness' of a group of grade 3 and 4 ESL learners. No differences were found between both programmes at the end of year 1. However, students following the regular programme obtained better ratings of comprehensibility and fluency at the end of year 2. More recently, Dlaska and Krekeler (in press) assessed the short-term effects of individual corrective feedback on degree of speech comprehensibility. Participants, 169 learners of German, received a treatment consisting of listening-only activities or of listening activities plus individual corrective feedback. Results showed that individual corrective feedback was more effective in improving learners' comprehensibility scores.

In light of the brief overview provided above, there seems to be a majority of studies pointing to the benefits of formal instruction in terms of production/perception of L2 sounds and general improvements in intelligibility, comprehensibility and degree of foreign accent. It is, however, difficult to establish any comparisons among different studies due to their methodological diversity.

It is also important to point out that the aforementioned lack of interest in L2 intelligibility studies and L2 pronunciation instruction becomes even more evident in the area of Spanish as a Second Language. This is partly due to the usual assumptions concerning Spanish pronunciation [10]. Nevertheless, our own experience in the classroom reveals that L2 learners of Spanish do have some pronunciation issues. In fact, the experimental part of this dissertation will demonstrate that, at a secondary school level, those pronunciation problems could prevent effective communication and therefore affect speech intelligibility.

In an attempt to fill some of the research gaps mentioned above, this study offers an assessment of the non-native speech intelligibility of a group of English learners of Spanish both at word level and in connected speech. Specifically, we will aim at analysing the impact of certain categories of phonemic errors, as well as three suprasegmental variables (pause frequency, pause duration and speech rate) on intelligibility scores. In addition, the possible correlation between the degree of speech intelligibility and learners' individual differences (gender, level of proficiency, motivation, aptitude and L1) will also be studied. Based on these results, some insights regarding Spanish pronunciation instruction will be proposed. This study will attempt to fill an existing research gap by focusing on:

1. Issues related to L2 speech intelligibility (most of the research has been undertaken in the area of speech pathology).

[10] As Poch (1992: 2) rightly points out: 'Este tipo de observaciones se fundamentan implícitamente en el hecho de que, en español, la distancia entre la ortografía y la pronunciación no es la misma que la que existe en inglés o en francés. Y, yendo un poco más lejos en el razonamiento, este punto de vista trasluce también una concepción de la fonética basada en la ortografía que, erróneamente, hace que las letras se conviertan en el referente de la pronunciación. Ello conduce a que algunos autores hablen del español como de una "lengua fonética" y que crean que se trata de una lengua "fácil"…… Por otra parte, el hecho de que la distancia entre ortografía y pronunciación sea pequeña no tiene nada que ver con la afirmación de que los sonidos del español no plantean problemas porque son casi idénticos a los sonidos del inglés'.

2. English learners of Spanish (the overwhelming majority of research has focused on learners of English from different L1 backgrounds).
3. English learners of Spanish at a secondary school level (researchers usually focus on university students or even subjects acquiring the language in a naturalistic manner).
4. Possible impact of certain categories of phonemic errors (segmental level) and suprasegmental variables on intelligibility scores both at word level and in connected speech (the majority of studies usually centre on either segmentals or suprasegmentals and on very specific levels of analysis).
5. Potential influence of learners' individual differences on both patterns of errors and L2 learners' intelligibility scores (it is unusual to find studies that offer an assessment of the possible influence of individual factors in combination with segmental and suprasegmental variables).
6. Pedagogical insights for pronunciation instruction (there seems to be a clear divide between empirical research and immediate applicability in the foreign language classroom).

As Rogers (1997) notes, intelligibility studies can benefit our general understanding of L2 speech perception and production. Furthermore, they can be used as a tool to bridge the gap between SLA research and actual instructional practices. From a pedagogical point of view, intelligibility studies can help avoid the 'risk of teaching things that are salient, but which will not result in actual improvement in communication for the speaker' (Derwing and Munro 2009: 482).

1.2 Pedagogical rationale

As pointed out in our previous section, one of the goals of this dissertation centres on offering some general pedagogical insights for the teaching of L2 pronunciation. It has been noted on numerous occasions that there is a gap between SLA research and teaching practice (e.g. Cook 2008). This problem affects the teaching of foreign languages in general and the teaching of pronunciation in particular. It is logical to think that the most fruitful approach to solve this issue is one of collaboration between theorists, researchers and teachers. This cooperation should provide information about

the most efficient conditions and the most common difficulties experienced by students during their learning process. It is also true that one needs to be cautious when attempting to generalise the findings of empirical studies, given that many of them are characterised by a small number of participants and a certain number of methodological shortcomings. We are currently far from having a comprehensive picture of all the variables involved in SLA. Studies correlating pedagogical interventions and success in SLA do not seem to yield any conclusive results (Larsen-Freeman and Long 1991). Moreover, the majority of studies have focused on morphology and syntax, while the empirical research related to L2 pronunciation is more limited. Accordingly, we cannot affirm with certainty what kind of classroom intervention may ensure progress in the pronunciation area. Against this background of uncertainty, we find that foreign language pedagogy cannot wait for researchers to reach a consensus on general learning principles. Teachers have to make immediate decisions to solve day-to-day challenges in the most efficient manner.

With regard to the teaching of pronunciation, it would seem that the fields of SLA and language pedagogy agree on some general principles. Both concur on the inherent complexity involved in the acquisition of the phonological component. One can also say that pronunciation is more resistant to progress and more prone to fossilisation[11]. Moreover, learners' progress seems to be conditioned, among other factors, by the amount of L2 input received, the opportunities they have to use the L2 significantly, the age at which the learning process begins or their degree of motivation. In addition, there are L2 learners who are able to acquire a higher level of phonological competence due to their ability for oral mimicry, as well as their aptitude in perceiving new phonetic elements. Transfer plays a major role in the phonological domain. In fact, we can usually predict the type of errors committed by L2 learners of the same linguistic background.

We have already highlighted the importance of the *intelligibility principle* as a driving force in the teaching of L2 pronunciation (see 1.1). In the following sections, some general pedagogical considerations will be put forward with regard to pronunciation instruction. We will briefly review the most important methodological

[11] SLA researchers often mention the phonological component, i.e. the difficulty in eliminating a foreign accent, as an argument in favour of the Critical Period Hypothesis. Furthermore, pronunciation, unlike grammar or vocabulary, possesses a physical dimension and has a direct impact on an individual's sense of identity. These factors make pronunciation more difficult to teach and more resistant to progress than other areas of language acquisition.

trends in the teaching of L2 pronunciation. Furthermore, we will examine the important notion of spoken fluency, as a dimension that is intrinsically related to speech intelligibility. This section will therefore provide a general framework in which to understand the specific pedagogical considerations described in Chapter 5.

1.2.1 Some general pedagogical considerations

First of all, it is necessary to clarify the notion of *teaching method*, given that its use has been rather inconsistent in the literature. Throughout history, teachers have striven to find an ideal method for foreign language teaching. In spite of the variety of proposals, methodological differences have always centred on two issues: first, the conceptualisation of 'language', and, second, the nature of the learning process. While some seemed to be more concerned with correction and linguistic structure, others have focused on fluency and transmission of meaning. Moreover, even though language instructors have exhibited a tendency towards an eclectic position in their daily practice, methods have always sought universal validity and have attempted to exclude any possible alternatives.

It must be pointed out that the use and conceptualisation of the term *method* itself has varied depending on the author in question. There seems to be an array of alternatives, e.g. method, approach, methodology, teaching style etc. that, on some occasions, are used as synonymous and on others, are considered as different notions altogether. On this issue, Anthony (1963) identified three levels of organisation regarding language teaching/learning:

1. *Approach* refers to a theory related to the nature of language and its learning process[12].
2. *Method* refers to a set of procedures, a system that is able to explain how to teach a language, i.e. what skills and content must be taught[13].
3. *Technique* alludes to specific classroom activities consistent both with a method and an approach[14].

[12] According to Anthony (1963: 64), an approach is 'axiomatic. It describes the nature of the subject matter to be taught. It states a point of view, a philosophy, an article of faith-something which one believes but cannot necessarily prove'.
[13] Anthony (1963: 65) points out that while an approach is 'axiomatic', a method is 'procedural' and should be based upon the selected method.

Richards and Rodgers (1998) drew on Anthony's distinction and used the notions of *approach* to refer to a theory of both the nature of language and the language learning process, *design* to refer to the general and specific goals of a method and *procedure* to allude to specific techniques utilised in the classroom. *Method* was used by these authors as an 'umbrella term' that included the notions of approach, design and technique. More recently, Cook (2008) uses *teaching style* as a broad notion to enable flexible discussions on language teaching. As Cook puts it (2008: 235), 'a teaching style is a loosely connected set of teaching techniques believed to share the same goals of language teaching and the same views of language and of L2 learning'. The conceptualisation of *teaching style* used by Cook (2008) will be adopted here, as it will provide us with a wide and flexible framework for our analysis of the different issues related to pronunciation instruction.

Before examining the different methodological trends in relation to L2 pronunciation, it becomes necessary to establish an initial distinction between pronunciation instruction, phonetic teaching and phonetic correction (Llisterri 2003a). Phonetic teaching involves a detailed analysis of segmental and suprasegmental elements on a perceptual, acoustic and articulatory level. This analysis should be the goal of very specialist courses. Pronunciation instruction, on the other hand, is concerned with the skills that need to be mastered by L2 learners and, therefore, should be incorporated into the foreign language curriculum. Phonetic correction puts forward a series of techniques and strategies aimed at reducing problems related to what is generally known as a foreign accent. This distinction reinforces the idea that it is not required for the instructor to be a specialist in phonetics to carry out phonetic correction or pronunciation teaching.

It is also important to address the issue of choosing an appropriate linguistic variety for pronunciation instruction. As far as variability and standard of language are concerned, it must be pointed out that languages are not uniform. They are indeed conditioned by speakers' geographical and social class. As Llisterri (2003a) notes, the choice of a linguistic variety for pronunciation instruction is not an easy task. In choosing a social variety, it is usually agreed that an educated standard variety will bring students a higher level of prestige. With regard to register, the choice does not usually pose many problems. In general, language courses tend to teach a colloquial

[14] 'A technique is implementational-that which usually takes place in a classroom. It is a particular trick, stratagem, or contrivance used to accomplish an immediate objective' (Anthony 1963: 66).

variety, as this is the one most students will need more often than not. Indeed, most learners aim at acquiring a new language in order to communicate in everyday situations. This does not mean that a more formal register is completely absent; it is simply far less predominant. By contrast, in courses for specific purposes, emphasis may be placed on more formal registers. On the other hand, the choice of a geographical variety is much more problematic. To understand the complexity of this problem, it is necessary to keep in mind the learner's needs, the teacher's own geographical variety and the variety of the place where the educational establishment is located. Teachers need to be aware of the variety of Spanish they are using in class and consider whether it is appropriate to their learners' interests.

The issue of pronunciation assessment is also central to any pedagogical proposal. L2 pronunciation assessment is always a difficult task. First, judgements on someone's pronunciation are unavoidably tainted by a certain degree of subjectivity and, second, it is difficult to separate completely the assessment of pronunciation from other dimensions of L2 speech such as fluency, accuracy, etc. In the general field of second language learning and teaching, numerous works have been published in the past decades focusing on the treatment of language testing/assessment (e.g. Bachman 1990; Hughes 2003). However, few of these works have specifically treated the issue of assessing L2 pronunciation. Llisterri (2003b: 552)[15], after reviewing the assessment criteria used by the *Instituto Cervantes* in their DELE exams, notes:

[15] Llisterri (2003b: 553-554) also notes that pronunciation testing should include an assessment of a subject's capability of production and perception of the following elements:

Elementos segmentales

- Elementos segmentales que configuran el sistema fonológico del español.
- Alófonos – variantes contextuales – que configuran el sistema fonético del español.
- Combinaciones de elementos segmentales - contacto entre vocales, entre consonantes y grupos consonánticos - tanto en el interior de la palabra como en el enunciado (fonética sintáctica).

Elementos suprasegmentales

- Patrones acentuales de la palabra en español, considerada aisladamente (acento léxico) y en el enunciado (acento de frase).
- Patrones melódicos del español ligados a la modalidad oracional.
- Patrones melódicos del español ligados a modalidades expresivas.
- Grupos acentuales, grupos fónicos, pausas y velocidad de elocución como elementos que configuran el ritmo en español.

> La presencia de importantes carencias y contradicciones en el terreno de la definición de objetivos y, en consecuencia, de la evaluación de la pronunciación en E/L2. Ello no hace más que corroborar la poca atención prestada a este campo.

In spite of the paucity of proposals, it is possible to find in the literature examples of authors that have offered some suggestions concerning L2 pronunciation assessment. Kenworthy (1996), referring to the various procedures that can be used to measure speech intelligibility, distinguishes between subjective judgements and objective assessments. Furthermore, this author, after weighing the advantages and disadvantages of tasks involving the reading of texts aloud and the collection of spontaneous speech samples, favours the inclusion of both types of tests in the evaluation of L2 pronunciation. Gimson (1989) recognises the difficulty of finding objective criteria for the evaluation of learners' oral productions. This author proposes to include minimal pairs to assess phonemic oppositions, word lists to assess stress patterns, and grammatical sentences to evaluate the production of sentence stress by L2 learners. Firth (1992) presents a model of explicit and objective pronunciation assessment in her student diagnostic profile. In this test, contents related to pronunciation assessment are divided into five areas: general habits of speech, intonation, rhythm and intonation, consonants, and vowels. Other authors, such as Celce-Murcia *et al*. (1996), favour a combination of subjective and objective assessment techniques. These authors believe that any form of assessment should include not just sections on oral production but also on speech perception. Celce-Murcia *et al.* (1996) include in their diagnostic test of oral perception exercises focusing on both segmental and suprasegmental aspects of speech. Furthermore, their diagnostic test includes two types of exercises: a diagnostic passage, which involves the reading of a carefully designed passage, and a spontaneous speech sample, which requires the L2 learner to talk on a previously agreed topic. Tench (1997) suggests that pronunciation assessment should be able to determine whether a learner's pronunciation is intelligible to the common native speaker in a simulated communicative situation. Moreover, this author suggests that all aspects of pronunciation should be evaluated: consonants, vowels and intonation.

In summary, a considerable number of authors (e.g. Kenworthy 1996; Celce-Murcia *et al.* 1996; Tench 1997) seem to agree on the idea that the assessment of pronunciation should be integrated within a comprehensive evaluation of oral production. This assessment should include a subjective test on a sample of

spontaneous speech extracted from a meaningful and communicative context, as well as an objective test focusing on a comprehensive analysis of all segmental and suprasegmental aspects. Similarly, there is also consensus among different authors on the fact that pronunciation tests should be based on activities and exercises practiced in class. In addition, self-assessment should be encouraged at all times, since this type of evaluation is most effective in providing L2 learners the required feedback about their pronunciation problems (e.g. Celce-Murcia *et al.* 1996; Firth 1992).

We must also address here the issue concerning the absence of appropriate teaching materials for pronunciation instruction. In the Hispanic tradition, there is a prominent theoretical literature on this area (Iruela 2004). However, it is striking to see how this abundance of theoretical perspectives has not yet fully translated into practical treatments of the teaching of pronunciation and phonetic correction. If we examine the teaching materials currently available in the market, we will be able to appreciate an overwhelming predominance of classical structural activities that focus largely on segmental elements (vowel and consonant sounds) to the detriment of the prosodic dimension (intonation, stress, rhythm, tempo and pauses). The treatment of prosody is usually marginalised to a mere few pages at the end of the different manuals. Furthermore, there seems to be a preference for activities and resources that are based on the formula 'listen and repeat' and that do not differentiate between phonetic correction and pronunciation instruction. This lack of specialised materials can also be seen if we examine the treatment of the pronunciation component in general Spanish as a Second Language textbooks. Pronunciation activities usually hold a marginal position and are unrelated to the rest of the curriculum unit. Moreover, they usually focus on isolated sounds through the use of minimal pairs or lists of words and sentences taken out of context. As indicated by Carbó *et al*. (2003), activities normally focus on the relationship between spelling and sound, the role of lexical stress and the relationship between sentence modality and intonational patterns. It is also noteworthy that several authors have put forward different typologies of activities aimed at working on the pronunciation component. Iruela (2004: 264), for example, presents a thorough typology of activities, exercises and teaching techniques by distinguishing between an instruction centred on pronunciation, a teaching centred on meaning and a teaching centred on the learning process. This author further distinguishes between activities that involve information processing, listening and those based on listening and repeating. Cook (2008: 82), on

the other hand, differentiates between activities focusing on imitation, discrimination of sounds, consciousness raising and communication. A comprehensive analysis of the different typologies of activities, exercises and teaching techniques clearly exceeds the scope of this dissertation (see Celce-Murcia *et al*. 1996; Gil 2007 or Iruela 2004 for comprehensive treatments of this issue). Nevertheless, it is important to highlight the widely acknowledged difficulty in finding a sufficient number of appropriate teaching materials that are integrated within the communicative approach and cover a progression of segmental and suprasegmental elements (Torres 2006).

1.2.2 Methodological trends in pronunciation instruction

This section will put forward a brief overview of the main methodological trends that have placed special emphasis on pronunciation instruction. An analysis of their contributions to the teaching of pronunciation will be presented. It must be pointed out that no attempt will be made to provide a chronological review and in-depth treatment of each methodological trend[16]. It is also important to note that computer-assisted pronunciation instruction[17] will be excluded from our analysis.

According to Celce-Murcia *et al*. (1996: 2), the teaching of pronunciation can be framed within two general approaches: the 'intuitive-imitative' and the 'linguistic-analytic' approach. The first one, as its name suggests, rests on the 'learner's ability to listen to and imitate the rhythms and sounds of the target language without the intervention of any explicit information'. The 'linguistic-analytic' approach, on the other hand, resorts to the use of certain tools, such as the phonetic alphabet or articulatory descriptions, in order to supplement the purely imitative nature of the intuitive-imitative approach. Gil (2007: 156-160) comments on this issue that a teacher needs to decide initially between a 'bottom-up' approach, i.e. a methodological perspective that first involves the analysis and correction of individual segments and only later focuses on larger units of analysis (prosodic elements), and a 'top-down' perspective that proceeds in the opposite manner. While the first approach

[16] For a chronological account see, among others, Larsen-Freeman and Long (1991) and Richards and Rodgers (1998). For general accounts of the methodological evolution within the area of pronunciation, see, for example, Celce-Murcia *et al*. (1996); Iruela (2004).

[17] See Levis (2007) and Hardison (2010) for recent developments in the field.

is associated with a structural view of foreign language teaching, the second one is more coherent with communicative methodologies. This second approach has been advocated by Rivers (1981) and Pennington and Richards (1986), among others.

In the history of foreign language teaching, the importance given to pronunciation instruction has varied greatly. In the grammar-translation method, for example, the teaching of pronunciation was virtually nonexistent, given that oral communication was not a primary goal. Audiolingual methodologies, on the other hand, emerged in the 50s in the United States and attached great importance to the explicit teaching of pronunciation. The audiolingual method adopted some of the assumptions of both structuralism and behaviourism and proposed a conceptualisation of language centred on the notion of habit. Based on the idea that children first acquire the spoken form of their L1, it was considered that adult learners would more easily acquire an L2 by focusing first on the spoken and not on the written language. In a typical audiolingual class, only the target language was used, translation was not allowed and both teachers and students played an active role in the learning process. As far as the treatment of errors is concerned, these were corrected immediately, so as to avoid any possible consolidation of incorrect forms. Linguistic structures were first presented and practiced orally and then in written form. It is, however, important to note that the linguistic structures used in this method were repeated mechanically by the student in a decontextualised fashion, thus preventing the creative use of meaningful language and putting into question the learner's ability to apply the target language to real-life situations[18]. Another criticism arises from the total ban on using students' L1 in the classroom (see, for example, Cook 2001, for a review of some of the possible benefits of using learners' L1 in the foreign language classroom).

Another important methodological trend in pronunciation instruction is the direct method of language teaching (Asher 1977; Krashen and Terrell 1983). This method is based on the idea that pronunciation is acquired in an intuitive way, similarly to children acquiring L1 pronunciation. Special attention is devoted, at the start of the learning process, to the listening phase. The goal is to offer L2 learners the chance of becoming familiar with the phonetic traits of the L2. This silent period enables them to be exposed to a new language without any communicative pressure.

[18] A typical example of an activity frequently used in the audiolingual classroom centred on the presentation of minimal pairs. Students were asked to discriminate the different phonemes and to subsequently repeat the different pairs as often as it was deemed necessary.

Moreover, it allows learners to access the target language through music and sounds without prioritising the transmission of meaning. Learners are therefore supposed to use this period to create an acoustic and articulatory image of the L2. The principle of a silent period at the beginning of the learning process is generally accepted and sometimes applied in classroom settings. This silent period is usually followed by a phase of imitation. It is important to note that researchers consider that one's ability for oral mimicry could be correlated with pronunciation accuracy (Purcell and Suter 1980). This aptitude for oral mimicry can certainly be the object of numerous classroom activities. A distinction between imitation and repetition must also be noted. 'Listen and repeat' is probably the instruction most often heard by students when undertaking work on pronunciation. However, repetition does not imply imitation. Through imitation, one attempts to reproduce the entire set of acoustic characteristics, while repetition only focuses on aspects that are perceived as linguistically or phonetically relevant. In spite of this, repetition tasks are the most widely used in pronunciation instruction. The teacher usually aims at proposing successive 'deformations' of different linguistic items that students then have to repeat with the goal of approaching the L2 target.

The verbo-tonal method emerged in the 50s as a treatment designed to help people with hearing problems. This system of phonetic correction centres on the idea that errors in pronunciation are caused by perceptual difficulties. In other words, an individual cannot produce L2 sounds that s/he is not able to discriminate perceptually. According to this method, L2 learners suffer from a 'phonological deafness' that prevents them from correctly perceiving L2 sounds. Perception is therefore influenced by L1 auditory habits. Moreover, similarities are established between L2 learners and individuals suffering from hearing deficiencies. In fact, phonetic correction is considered by this method as a type of perceptual rehabilitation. It is noteworthy that, since its inception, this method has emphasised the importance of considering L2 learners as communicative beings by taking into account the emotional and physical aspects inherent to any communicative exchange. As Gil (2007: 145) comments:

> El método verbo-tonal se anticipó en diversos sentidos a los postulados considerados posteriormente, en los últimos años del siglo XX, como auténticas innovaciones pedagógicas: la prioridad absoluta de la comunicación oral sobre la escrita, la importancia atribuida a la percepción, la integración obligada del contexto real, el recurso al componente no verbal de todo acto comunicativo, la consideración en el

> proceso docente de la afectividad y la subjetividad, la relevancia que le es reconocida al factor humano etc.

There are some similarities between the audiolingual and the verbo-tonal method of phonetic correction. However, in light of some Gil's observations reproduced above, we can also see that the differences between both methods are clear. The 'affective' dimension of the L2 learner as a communicative being is completely absent in audiolingual methodologies. It is also important to point out that in the design of instructional materials, this approach places great emphasis on prosodic elements and resorts to the use of techniques such as the '*prononciation nuancée*' and '*phonétique combinatoire*' (see Renard 1979 for a detailed analysis of these principles).

The communicative approach emerged in the late 60s coinciding with advances in sociolinguistics, pragmatics and developmental psychology. One could say that 'communication' has always been the unspoken goal of foreign language teaching. However, it was the communicative approach that first introduced everyday language into the foreign language classroom through a strong theoretical support from linguistics and psycholinguistics. In addition, it must be pointed out that some of its principles arose, in part, as a reaction to audiolingual methodologies. As far as pronunciation instruction is concerned, the dissatisfying results obtained by the audilingual method of language teaching soon became apparent. The goal of achieving a native accent was seen as extremely difficult and unnecessary for most learners. In fact, from a functional and communicative perspective, an intensive focus on pronunciation practice was no longer necessary as part of L2 teaching, given that foreign-accented speech only, on occasions, seems to disrupt communication. Emphasis was therefore placed not on phonetic correction but on fluency. The communicative approach to language teaching assumed that errors are part of the learning process and disappear as phonological competence develops. Moreover, this approach promoted the use of real language by students, as well as the adoption of real texts as teaching materials. Emphasis was also placed on the importance of intonation, given that there is a high degree of communicative value associated with this feature. In addition to intonation, the importance of rhythm and stress was also highlighted. It is noteworthy that the communicative approach did not provide a systematic treatment of pronunciation issues. In fact, the presence of teaching materials specifically designed for pronunciation instruction was minimised. Thus, it

was not immediately clear how to integrate pronunciation instruction within this framework.

Regardless of the differences between the teaching methods that we have briefly reviewed above, the success of any instructional programme will eventually depend on the teacher's own capabilities and linguistic knowledge. Furthermore, any instructional programme should be adapted to the individual characteristics of the targeted population of L2 learners and the circumstances of the specific institutional establishment. In spite of the difficulties, the existence of a framework provided by a coherent and well-developed teaching approach is certainly desirable so as to tackle, in a systematic manner, the everyday problems that instructors face in the L2 classroom.

Numerous authors (e.g. Gil 2007; Morley 1991) have advocated, without subscribing to any specific teaching method, a three-stage approach in the acquisition of pronunciation. First, an initial 'auditory retuning', which teaches language learners to listen to L2 speech samples in the same way that they listen to their L1. As MacCarthy (1978) points out, students' attention needs to focus initially on specific phonetic elements. Secondly, it is necessary to give learners sufficient time to listen repeatedly to the presented stimuli. Finally, after verifying the students' interpretation of the stimuli, the correct answers are presented. The idea is, therefore, to emphasise perception over production, especially the role of certain perceptual sub-skills that may contribute to effective listening comprehension such as, for example, the identification of features related to the articulatory setting, the recognition of phonological contrasts, the use of stress to indicate boundaries between units etc. (see Bohn and Flege 1996; Cortés Pomacóndor 1999; Kluge *et al*. 2007; Llisterri 1995 and Neufeld 1988, among others, for research exploring the relationship between speech perception and production). This first stage of 'auditory retuning' is then followed by a 'mimicking phase', which, in turn, leads to a 'production stage'. It is argued here that students must go through a stage of imitation of proposed models before jumping into a free production phase. The 'mimicking phase' does not require the endless mechanical repetition of previously recorded models presented out of context. On the contrary, teachers need to manipulate the different speech samples so as to include real materials that reflect contextual variations. Free production must be stimulated by meaningful exercises conceived and designed within a communicative framework.

These types of activities reflect the fact that pronunciation cannot be presented in isolation.

1.2.3 Putting oral fluency at the heart of speech intelligibility

It is assumed here that the mere definition of spoken fluency, as we will see below, is intrinsically related to the notion of speech intelligibility. This section will address the difficult task of offering a definition of this dimension. We will draw on both objective accounts of spoken fluency and on more interactive/social perspectives of this phenomenon. We will also place the concept of fluency within the framework provided by the notion of communicative competence. Moreover, the importance of pronunciation within communicative competence will also be presented.

1.2.3.1 Defining fluency

Levelt's (1989) model of L1 speech production posits that the process of putting ideas into words must proceed through three different stages: conceptualization, formulation and articulation. Production processes at each stage run, as Kormos (2006: 154) notes, 'automatically and in parallel, without the speakers' conscious supervision'. When it comes to L2 speech, it is generally accepted that the degree of automaticity is significantly lower, depending on the speaker's level of proficiency. This usually results in a slower speed of delivery.

Fluency seems to be a rather complex notion that encompasses various linguistic, psycholinguistic and sociolinguistic features. Authors have linked fluency to smoothness in continuity in discourse (Crystal and Davy 1975) or to natural language use (Brumfit 1984: 56). Kormos (2006: 155) notes that the notion of fluency can be understood in a broad sense, as related to the idea of 'global oral proficiency', or in a more narrow sense, by equating it to only 'one component of oral proficiency'. Lennon (1990: 91) defines fluency as 'an impression on the listener's part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently'. Schmidt (1992: 358) defines this notion as an 'automatic procedural skill'. Freed, Segalowitz and Dewey (2004: 279) use quantifiable variables in order to define the notion of fluency: number of words per minute, silent pauses,

filled pauses, repetition, duration of discourse, etc. The Common European Framework of Reference for Languages (2001: 128) defines fluency 'as the ability to articulate, keep going and to cope when one lands in dead end'. Fillmore (1979: 93) argues that fluency relates to the 'ability to fill time with talk, the ability to talk in coherent, reasoned and "semantically dense" sentences, the ability to have appropriate things to say in a wide range of contexts and the ability to be imaginative and creative in language use'. It seems possible, based on the above definitions, to group all the variables related to the notion of fluency into two major categories: those that refer to the degree of accuracy in the production of language and those that point to temporal/sequential aspects such as repetition or pauses. The three suprasegmental variables analysed in the empirical section of this dissertation relate to temporal aspects of speech production and can be directly linked to the notion of fluency. Nevertheless, this series of sequential/temporal aspects of speech are not the only ones that bear an impact on fluency. Following Guillot (1999), it is worth noting that the notion of fluency is highly impressionistic, difficult to define and difficult to assess in objective terms. As this author points out (1999: 40):

> Whatever the actual characteristics of a speaker's verbal production, the degree to which he is in practice regarded as fluent is highly relative. And if the interlocutor's expectations, his processing capacities and tolerance do have a role to play in the setting of fluency parameters, what does it mean for the speaker as a speaker and for the speaker as learner?

More recently, McCarthy (2009: 13) in his literature review on the notion of spoken fluency, points out that the surge of research in this field has centred on three major themes: the speed of delivery, the idea of automaticity and the assessment of fluency by professional and non-professional practitioners. This same author goes on to put forward the term 'confluence' instead of fluency in an attempt to encapsulate the highly interactive nature of this notion. According to the literature available, speech rate is one of the best predictors of spoken fluency (e.g. Lennon 1990). However, measurements have also focused on the number of disfluencies and pauses per unit of time, the length of pauses or the number of stressed words. It is noteworthy that Kormos (2006) accounts for the notion of fluency in L2 speech by referring to two different types of processes: use of formulaic language and automatisation of encoding processes. The automatisation of encoding processes may take place when conscious rules of language become automatic; when clauses assembled thanks to

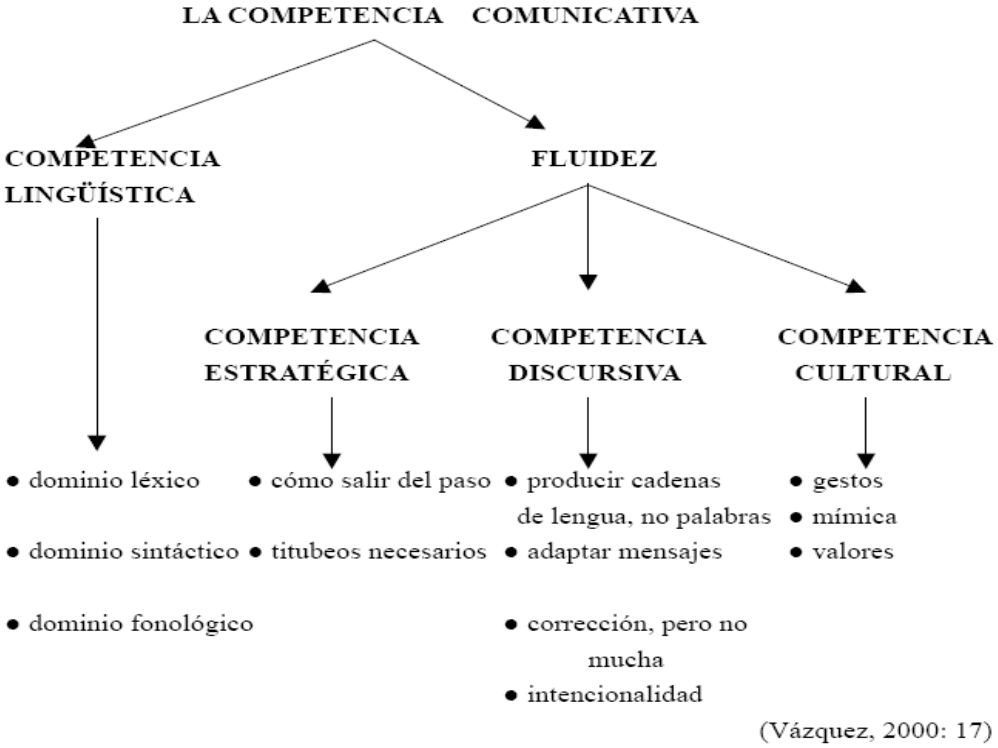
phonological and syntactic rules are stored as one unit; or when stored chunks of language allow for a subsequent deduction of rules during the learning process (see Kormos 2006 for an overview of theories that could potentially account for these three processes).

1.2.3.2 Fluency, communicative competence and pronunciation

This section will examine the role of fluency and pronunciation as part of an individual's communicative competence. We will first review the notion of communicative competence from several perspectives. In addition, we will attempt to clarify the role of spoken fluency and pronunciation within the different subcomponents of communicative competence.

Spoken fluency has, in fact, been linked to the strategic, discourse and cultural subcomponents of communicative competence (see Figure 1 below, taken from Vázquez 2000: 17).

Figure 1: Fluency and communicative competence (Vázquez 2000: 17)



It must be noted that the first theoretical applications of the notion of communicative competence to the teaching of foreign languages emerged in the 80s (Canale and Swain 1980; Canale 1983). Canale and Swain laid the foundations for the communicative approach and had a profound impact on subsequent pedagogical developments. Their model of communicative competence advocates a broader understanding of this term through the inclusion of an interaction between grammatical and sociolinguistic competence. Grammatical competence, labelled by subsequent authors as linguistic competence, is the component that refers to the linguistic code and 'includes the knowledge of lexical items and of rules of morphology, syntax, sentence-grammar semantics, and phonology' (Canale and Swain 1980: 29). Sociolinguistic competence 'will specify the ways in which utterances are produced and understood appropriately' (Canale and Swain 1980: 30). Appropriate use will, in turn, depend on different contextual factors, such as the type and purpose of the communicative interaction or the rules and conventions of language use. A third component included in their initial model is strategic competence (see e.g. Dörnyei and Thurrell 1991 for an analysis of this concept within the field of Foreign Language teaching). As pointed out by Bagarić and Djigunović (2007: 98), strategic competence is qualitatively different from the other two, since it does not refer to 'a type of stored knowledge and includes non-cognitive aspects such as self confidence, readiness to take risks, etc'. This model was subsequently expanded by Canale (1983) with the addition of a discourse component. It must be noted that this component was viewed, in the original model, as part of a sociolinguistic competence that was made up of rules of use and rules of discourse (Canale and Swain 1980: 30). Canale's model sees it as an independent component that alludes to 'the mastery of how to combine grammatical forms and meanings to achieve a unified spoken or written text in different genres' (Canale 1983: 9). As pointed out by Celce-Murcia (2007) or Bagarić and Djigunović (2007), Canale and Swain's model of communicative competence has been widely used in the field of language teaching and remains 'a key source for discussions of communicative competence and related applications in applied linguistics and language pedagogy'(Celce-Murcia 2007: 41).

We have indicated above that the notion of fluency can be placed within the more general concept of communicative competence. It is also noteworthy that the importance of pronunciation, from a pedagogical perspective, can also be justified by

referring to its sheer relevance for each of the subcomponents of communicative competence. Grammatical competence includes knowledge of semantic, syntactic, orthographic and phonological rules. Phonological competence includes, therefore, the skills of perception and production of speech sounds. According to Canale and Swain (1980), sociolinguistic competence is divided into two subcomponents: an organisational competence, which enables individuals to follow schemata of interaction for each specific situation, and a functional competence, which refers to the use individuals make of different linguistic forms. Sociolinguistic competence encompasses the use of different registers depending on the context of communication. Moreover, it also includes differences based on geographical, social or professional varieties. Pronunciation intervenes in a relevant manner in the characterisation of each variety of language. Furthermore, pronunciation permits the distinction of different geographical and social variants. Discourse competence, on the other hand, refers to the rules of cohesion and coherence. Elements that provide cohesion to an oral or written text are numerous. For example, sentences could be connected to each other through the use of connectors, vocabulary from the same semantic field, words with an anaphoric or categoric value, etc. In written texts, punctuation is a resource that provides cohesion. In oral discourse, cohesion is, for example, provided by intonation patterns that serve to distinguish and link phrases and sentences. Strategic competence includes those verbal and non-verbal strategies that are utilised to compensate for breakdowns in communication. Pronunciation may play an important part in cases of register and style shifting. Moreover, intonation plays a relevant role in instances of repetition, modification of messages or paraphrasing.

As seen above, communicative competence is a complex puzzle of different subcomponents. Pronunciation plays an important role within linguistic competence and intervenes in a significant and meaningful way on a grammatical and semantic level. As far as pragmatic competence is concerned, there are many communicative functions associated with certain suprasegmental elements, such as intonation and intensity, which serve to convey a speaker's attitude or communicative intention. Regarding discourse competence, intonation is one of the resources that give cohesion to an oral text. Finally, pronunciation plays a decisive role in sociolinguistic competence, given that it is one of the elements characterising geographical and social varieties of speech.

1.3 Outline of this dissertation

This dissertation will provide an assessment of the non-native speech intelligibility of a group of 20 Key Stage 4 learners of Spanish. Sixty evaluators, native speakers of Peninsular Spanish, will transcribe different speech samples belonging to our group of participants. The elicitation of these speech samples will serve to assess intelligibility at word level (through the use of a single word intelligibility test), at sentence and passage levels (using the Spanish version of the Harvard Psychoacoustic Sentences and the reading of a phonetically balanced text) and in semi-spontaneous production. In terms of individual differences, aptitude will be measured through an oral mimicry task and a working memory capacity test. A Likert scale questionnaire will serve to measure students' degree of motivation. Level of proficiency will be determined through the use of the *Diploma de Español como Lengua Extranjera* (DELE) exam. Correlation and multiple-regression analyses will serve to examine the relationship between different categories of phonemic errors, three suprasegmental[19] variables and students' intelligibility scores. Correlation analyses will also be used in assessing the potential influence of certain individual differences on learners' intelligibility scores. The following research questions will therefore guide the present study:

1. Is there an intelligibility loss when assessing the non-native speech intelligibility of our group of Key Stage 4 learners of Spanish?
2. Is there any correlation between intelligibility scores at word level and in connected speech?
3. At word level, what categories of phonemic errors seem to impact the most on speech intelligibility, as determined by our single word intelligibility test?
4. Do any of the suprasegmental features examined in this study (speech rate, pause frequency and pause duration) have an impact on intelligibility scores in connected speech?
5. Do the appropriate statistical analyses highlight the importance of segmental over suprasegmental deviances or vice versa?

[19] The terms 'suprasegmentals' and 'prosodic features' will be used here as synonymous even though, as Clark and Yallop (1990: 276) note, 'the implication that suprasegmentals are somewhat superimposed on a basic message of consonants and vowels is decidedly misleading, given that prosody is an integral part of speech production and often a fully meaningful contribution to the message itself'.

6. Can we find any correlation between students' intelligibility scores and certain individual variables, such as gender, motivation, aptitude and level of proficiency?

In terms of the structure of this dissertation, Chapter 2 will centre on the notion of speech intelligibility by drawing on studies from L1 and L2 speech. Moreover, our review of the literature will try to clarify some terminological issues and explore the relation between intelligibility and other dimensions of L2 speech such as comprehensibility and 'accentedness'. This chapter will provide an analysis of those speaker, listener, task and environment-related factors that may have an impact on speech intelligibility. In addition, some methodological concerns on the difficulties of measuring speech intelligibility will be identified.

Chapter 3 will introduce the empirical part of this dissertation by presenting the methodological framework adopted for this study. This chapter will outline our research questions, a description of participants and evaluators, some ethical considerations regarding educational research, a description of the materials used in this study, as well as a detailed analysis of the recording procedures and the assessment criteria followed for each test.

Chapter 4 will present our analysis of results in three parts. First, an analysis of students' scores on motivation, aptitude and level of proficiency will be put forward. Specifically, our analysis will centre on participants' performance in the motivation questionnaire, each of the sections of the DELE exam, as well as their results in the oral mimicry task and the four tests measuring working memory capacity. Second, an analysis of learners' results in the intelligibility tests (word, sentence, passage and semi-spontaneous production) will be presented. Third, correlation and multiple-regression analyses will examine the possible relation between intelligibility scores and the segmental and suprasegmental variables under study. Furthermore, correlation analyses will also serve to explore the relation between intelligibility scores and certain individual differences, namely, gender, L1, aptitude, motivation and level of proficiency. Finally, based on the aforementioned analyses, an attempt will be made to answer the research questions originally put forward at the beginning of this study. Additionally, an account of some potential limitations will also be offered.

Finally, Chapter 5 will offer a summary of this dissertation and will highlight the major contributions yielded by this study. Moreover, we will provide an analysis of the pedagogical implications of placing the notion of fluency at the cornerstone of speech intelligibility. Specifically, some pedagogical implications in terms of individual differences and curriculum sequencing will be presented. Some avenues for further research will also be put forward.

Chapter 2: Review of the literature

This chapter provides an overview of the notion of speech intelligibility by drawing on areas such as speech disorders, English as Lingua Franca, World Englishes and L2 speech. In addition, we will examine the relation between intelligibility and other relevant dimensions of L2 speech, such as 'accentedness' and comprehensibility. Given the complexity of variables affecting speech intelligibility, we will provide an analysis of speaker, task, environment and listener-related factors, as well as an assessment of the potential impact of segmental and suprasegmental deviations. An overview of the different methodological proposals regarding the difficult task of measuring speech intelligibility will also be put forward.

2.1 The notion of speech intelligibility

As pointed out in Chapter 1, it has proven difficult to offer a straightforward definition of speech intelligibility given that the terminology is often used inconsistently across disciplines and authors. In the area of communication disorders, the term intelligibility is widely used (Rogers 1997), while in the field of L2 phonological acquisition, terms such as comprehensibility, communicative effectiveness, interpretability or 'accentedness' are usually mentioned in connection with the notion of speech intelligibility[20]. Kent *et al*. (1989: 489) define intelligibility as 'the degree to which a speaker's intended message is recovered by the listener'. Weismer (2008: 2) offers a more thorough definition when noting that intelligibility is a 'relative measure of the degree to which a speaker's speech signal is understood, the relativity depending at a minimum on the identities of the speaker and listener, what is spoken and where it is spoken'. Weismer's definition gives an idea of the complexity of this notion, while underlining the wide array of factors that may have an impact on speech intelligibility. In the area of L2 speech, Rogers (1997: 2-3) distinguishes between intelligibility, comprehensibility and communicative effectiveness.

[20] 'Acceptability' and 'intelligibility' are also frequently interchanged. Hodge and Whitehill (2010: 101) note ' that definitions of acceptability range from subjective impressions of the pleasingness of speech to the potential for a person to experiment social, educational, or vocational problems because of speech'.

Intelligibility refers to the effective 'production, transmission or perception by a listener of the speech sounds of a language'. Comprehensibility alludes to an individual's success, or lack thereof, in conveying a specific message and includes not only variables related to the speech sounds but also to additional linguistic features, i.e. syntactic, lexical etc. Communicative effectiveness, on the other hand, seems to encompass those linguistic and non-linguistic variables that may determine communicative success. Other authors, such as Gass and Varonis (1984), use the term comprehensibility to refer to communicative effectiveness without specifying the possible impact of linguistic and non-linguistic variables.

Catford (1950) distinguishes between intelligibility and communicative effectiveness. The former refers to the understanding of the linguistic form, while the latter points to the hearer's response to the speaker's intended message. Catford (1950: 9) further notes that, from the perspective of the language learner, speech 'should be not only intelligible, in the narrow sense, but also effective'. This author's notion of intelligibility includes therefore the idea of communicative effectiveness. Similarly, Kenworthy (1996: 117) equates intelligibility to 'understandability' and defines this dimension as 'being understood by a listener at a given time in a given situation'. This broad proposal does not distinguish between those linguistic and non-linguistic variables that may have an influence on communicative success.

In the field of World Englishes, Smith and Nelson's (1985) oft-cited proposal establishes a distinction between intelligibility, comprehensibility and interpretability. These dimensions refer respectively to utterance recognition as well as understanding of locutionary and illocutionary force. Jenkins (2000) has also devoted a great deal of attention to the notion of intelligibility and the importance of phonological deviations for Interlanguage Talk (ILT). Intelligibility is seen by this author as a prerequisite for comprehensibility and interpretability. As Jenkins (2000: 78) points out, intelligibility 'concerns the recognition of the formal properties of words and utterances and, in particular, the ability to produce and receive phonological form but regards the latter as a prerequisite (though not a guarantee) of ILT success at the locutionary and illocutionary level'. Differences between Jenkins and Smith and Nelson arise from the emphasis they place respectively on either the importance of phonological deviations or the relevance of pragmatic issues. Furthermore, it must be noted that while Jenkins' use of the terminology stems from the field of linguistics, Smith and Nelson's is

related to the area of cross-cultural communication (see Sewell 2010 for a recent review on the notion of intelligibility within the field of World Englishes).

In the general field of L2 speech, Derwing and Munro (2009: 478-479) distinguish between intelligibility, comprehensibility and 'accentedness'. Intelligibility is defined as 'the degree of a listener's actual comprehension of an utterance'. Comprehensibility is operationalised in terms of 'how difficult or easy it is to understand a given speech sample', while 'accentedness' is understood 'as how different a pattern of speech sounds compared to the local variety'. The definitions of these three dimensions clearly entail a departure from Smith and Nelson's classical paradigm. Intelligibility includes both recognition of words and meaning, while the notion of speech comprehensibility is radically altered (Nelson 2008). Derwing and Munro (2009) go on to note that intelligibility can be measured using a wide range of techniques, none of which is fully satisfying: dictation, transcription of words, comprehension questions etc. Furthermore, as we pointed out in Chapter 1, the same authors (1997, 2009) have also put forward the idea that intelligibility is not necessarily correlated with 'accentedness', i.e. a heavy accent may or may not reduce speech intelligibility. Conversely, intelligibility and comprehensibility seem to be more closely related, while comprehensibility and 'accentedness' can be considered as two different dimensions (Derwing and Munro 2009: 479).

As seen from the sample of authors cited above, the definition of speech intelligibility has not remained consistent across disciplines. However, a general distinction can be established between the notion of recognition of linguistic form and meaning, and the communicative effectiveness of a speech act at the pragmatic level. Researchers seem to focus on different aspects depending on their field of study and methodological decisions. The adoption of one perspective over another will obviously depend on the research goals, the type of instruments of data collection and the contextual factors surrounding a specific research project. In this study, intelligibility will be used to refer to the recognition and production of phonological form, similarly to Jenkins's use of the term, but with the difference that production will point to our group of English learners of Spanish and recognition to a group of evaluators that consists only of native Spanish speakers (a full description of participants and evaluators will be provided in Chapter 3). Pragmatic issues regarding the idea of communicative effectiveness or appropriateness will be excluded from our analysis.

2.2 Intelligibility and other dimensions of L2 speech

Having examined the notion of speech intelligibility from different perspectives, we now turn to the relation between intelligibility and other dimensions of L2 speech. First, some general considerations on the nature of foreign-accented speech and speech comprehensibility will be put forward. This will be followed by an analysis of the relationship between intelligibility, comprehensibility and 'accentedness' through a careful consideration of some of the research carried out in the past two decades.

As pointed out above, 'accentedness' refers to how much a pattern of speech sounds differs from the local variety (Derwing and Munro 2009: 479). In the field of language learning, foreign-accented speech alludes to the non-pathological deviations that usually characterise L2 speech when compared to that of a native speaker. These deviations affect both the segmental and suprasegmental levels. Learners who share the same L1 usually present similar types of deviations due to the phonetic interference exercised by the L1 over the L2. Most authors who have addressed the nature of foreign-accented speech have focused on the segmental level. Trubetzkoy (1939) or Polivanov (1931) concluded that foreign-accented speech is mainly due to the inaccurate production of speech segments. However, some empirical studies seem to indicate that prosodic deviations are essential in determining an individual's degree of foreign accent. Munro (1995), after removing the segmental information from a series of utterances produced by native English speakers and Mandarin learners of English, concluded that untrained evaluators consistently identified foreign-accented speech on the basis of the suprasegmental information available. In addition, it must be noted that the impact of the segmental and suprasegmental levels on foreign-accented speech varies depending on an individual's L1 and L2. Anderson-Hsieh, Johnson and Koehler (1992) analysed the pronunciation of 60 English learners of different L1 backgrounds and concluded that deviations in segmentals, prosody and syllable structure all had a significant impact on pronunciation ratings. The prosodic variable had the strongest impact regardless of the language combination under study. In addition, the perception of a foreign accent may also be linked to the influence of the L1 articulatory setting (see Gil 2007 for a review of the notion of phonetic setting) or even to differences in voice quality (Laver 1980).

Empirical research has also underlined the high degree of sensitivity experienced by listeners with regard to foreign-accented speech. Flege (1984)

examined the ability of a group of trained and naïve native speakers of American English to detect French-accented speech even when the speech samples presented to them were extremely short (the first 30 ms). Both groups were able to detect French-accented speech consistently and accurately. Moreover, as pointed out by Flege and Fletcher (1992), methodological decisions regarding the presentation of the speech samples can influence evaluators' judgements on foreign-accented speech. In Flege and Fletcher's study, familiarity with the speech material resulted in higher ratings of perceived foreign accent. Familiarity also proved to affect the comprehension of foreign-accented speech in Gass and Varonis (1984). Additionally, a study pointing to a higher reliability of experienced versus inexperienced listeners can be found in Thompson (1991). Major (2007), on the other hand, claims that the ability to perceive foreign-accented speech is innate and independent of L1, L2 or an individual's degree of linguistic experience. Similarly, Munro, Derwing and Morton (2006) examined the intelligibility, comprehensibility and 'accentedness' scores of a group of evaluators from four different L1 backgrounds. Foreign-accented speech samples belonged to a group of non-native English speakers with four different L1s. Results did not show significant differences between native and non-native English evaluators. Nor did they show any significant differences between the four groups of evaluators with regard to their intelligibility, comprehensibility and 'accentedness' scores.

It is worth noting that the ability to perceive foreign-accented speech is highly effective even in cases of backwards speech. Munro, Derwing and Burgess (2010) showed that this ability remained intact across different accents and speech samples, even after the removal of different linguistic cues. The detection of foreign-accented speech was impaired only in cases of severe alteration of speaking rates and elimination of the fundamental frequency (F0).

The study of foreign-accented speech is also relevant from a social standpoint. Some studies have shown the benefits of exhibiting certain L2 accents (Varonis and Gass 1982), while other researchers have underlined the negative impact of such phenomenon, i.e. cases of social discrimination (Munro 2003; Munro, Derwing and Sato 2006). Additionally, interesting research has arisen regarding the nature of the relation between identity and accent. Some seem to think that an accent is part of an individual's identity and no attempt should be made to change it (Porter and Garvin 1989). Moreover, the detection of a foreign accent allows the listener to share some responsibility and make the necessary modifications during the communicative

exchange. Others argue that students' desire to acquire a native-like accent needs to be taken into consideration (Timmis 2002). In addition, it is noteworthy that certain learners, although willing to approach the yardstick of a native-like pronunciation to avoid the stigma of a foreign accent, are, at the same time, reluctant to lose their native accent for fear of being rejected by their own countrymen (Stevick 1978). Pennington (1994) suggests that the goal of instruction is not only to improve learners' performance, but also to encourage a change in the psychological and emotional dimensions involved in their acquisition of pronunciation. It seems, therefore, imperative to find more effective ways of integrating psychological and sociological aspects in language teaching and to promote a more positive attitude of L2 learners towards the L2 culture.

Hansen Edwards (2008), in a recent literature review, has examined the impact of sociolinguistic and sociocultural theory on L2 phonological acquisition. Approaches used in this area have ranged from social network theory and speech accommodation to variable rule analysis of social and linguistic factors. Moyer (2004), for example, has focused on the constructivist nature of creating an identity in the L2. This author points out that confidence in using the L2 is essential in creating a sense of self in the L2 community. On the other hand, Thompson (1991) and Lybeck (2002) have examined how maintaining strong links to the L1 community can have an impact on L2 phonological performance. An interesting study that highlights the dynamic nature of the relation between identity and accent can be found in Piller (2002). Piller's results suggest that motivation to achieve a native-like pronunciation is not only linked to those who for professional reasons need to speak without a foreign accent, but can also be extended to ordinary individuals depending on their communicative context, personal and social circumstances. This author carried out a qualitative study on highly proficient L2 learners by describing their ability to pass for native speakers in certain contexts. Data suggest that 'age is not the critical factor in reaching high levels of L2 proficiency it is often assumed to be. Rather, personal motivation, choice and agency seem to be more crucial factors in ultimate attainment' (Piller 2002: 23).

As far as the notion of speech comprehensibility is concerned, it must be noted that definitions of this term vary across authors and are very much dependent on the notion of intelligibility put forward by each individual researcher. In the realm of speech disorders, intelligibility generally refers to the understanding of the acoustic

signal, while comprehensibility points to the conveyance of communicative intent through the use of the acoustic signal plus any additional information that may be available in the environment (Yorkston, Strand and Kennedy 1996). Comprehensibility can be affected not only by factors related to the acoustic signal or the speaker, e.g. speech impairment and compensatory strategies, but also by signal-independent factors such as the semantic and syntactic context, the situational cues and certain aspects of non-verbal communication (see Visser 2004 for a review). In the field of clinical disorders, several studies have explored the relation between intelligibility and comprehensibility. Hustad and Beukelman (2002) examined the relationship between intelligibility and comprehensibility in a group of speakers with severe dysarthria and found no significant correlation. Similar results were also found in Hustad (2008). It would seem, therefore, that intelligibility and comprehensibility are tapping into different cognitive processes. While intelligibility 'can be considered a form of surface code because the focus of measurement is phonetic and lexical identification accuracy, […] comprehensibility refers to propositional content or higher level situation models' (Hustad 2008: 563). In other words, intelligibility is equated with successful matching of linguistic forms and comprehensibility with the successful decoding of the message.

In the area of L2 speech, Derwing and Munro (2009) have operationalised comprehensibility as the degree of difficulty in understanding a speech sample. Furthermore, they have carried out abundant work on the relation between intelligibility, comprehensibility and 'accentedness'. Initially, some considerations regarding the instruments used by these authors to measure the aforementioned dimensions must be put forward. Comprehensibility and 'accentedness' have been assessed through the use of a 9-point Likert scale. In the case of comprehensibility, the scale ranges from 'very easy to very difficult to understand', while the scale used to measure 'accentedness' ranges from 'no accent to extremely strong accent'. Intelligibility, on the other hand, has been assessed using a variety of instruments such as transcriptions, comprehension questions, summaries, etc (see Munro 2008 and Section 2.4 for a detailed analysis).

Derwing and Munro (1997) assessed the degree of intelligibility, comprehensibility and 'accentedness' of 48 ESL students with different L1s: Cantonese, Japanese, Polish and Spanish. Speech samples, elicited through a picture description task, were evaluated by 26 native English speakers. Comprehensibility

and 'accentedness' were assessed using a 9-point Likert scale, while intelligibility was measured using an orthographic transcription of each speech sample. In addition, the number of phonemic and grammatical errors, as well as the prosodic deviations and speaking rate were determined for each utterance. Results suggest that intelligibility, comprehensibility and 'accentedness' are partially independent dimensions. As Derwing and Munro (1997: 11) point out:

> The results provide additional evidence in support of the claim that, although some features of accent may be highly salient, they do not necessarily interfere with intelligibility. A clear implication of this finding is the need to disassociate accent ratings and intelligibility in language assessment instruments, which often confound the two dimensions. The discrepancy between perceived comprehensibility ratings and transcription success suggests that some accented but fully intelligible utterances may require additional effort or processing time (cf. Munro & Derwing, 1995b), which leads listeners to rate them as difficult to understand.

Other studies confirming the relation but also the partial independence between intelligibility, comprehensibility and 'accentedness' can be found in Munro and Derwing (1995) and Derwing, Munro and Wiebe (1998). The differences between these three dimensions have also been explained in terms of processing difficulty. It seems that 'the amount of processing time' has been linked to comprehensibility but not to 'accentedness' (Munro 2008: 204). In other words, evaluators may take longer to rate a specific speech sample not because it has a higher or lower degree of foreign accent but because it is more or less difficult to understand.

These perceptual studies give empirical support to the fact that it is not uncommon to find L2 speakers with a very heavy accent but whose non-native speech is easy to understand. There are, however, a number of methodological limitations to this type of research, some of which have been rightly pointed out by their authors: speech samples are usually very short and comparisons across participants are rather difficult, especially when samples are obtained from spontaneous interactions.

In Section 2.3 below, we will examine some of the variables that, according to the existing literature, can potentially affect the intelligibility of non-native speech.

2.3 Factors affecting speech intelligibility

In the general scheme of communication, where information flows from the speaker to the listener, intelligibility can decrease for reasons related to the speaker, the listener or the transmission channel. When the difficulty arises from the speaker, intelligibility is affected by a deterioration of the speech signal, as in cases of certain speech disorders or non-native speech. The difficulty could also derive from the listener in the event of, for example, deafness, hearing loss or a lack of familiarity with a specific foreign accent. Additionally, the transmission channel may play an important role in intelligibility measurements (e.g. noisy environment, overlapped conversations, etc.). In fact, early research on speech intelligibility focused on the transmission channel by assessing the quality of telephone communications (Weismer 2008). Rogers (1997) offers an idea of the complexity of factors affecting speech intelligibility by classifying these variables in speaker-related (age, time-related factors, and segmental or suprasegmental deviations), listener-related (hearing level, experience, training and age of the listener), task-related (context, complexity and cognitive load) and medium-related variables[21].

In the field of L2 speech, Munro (2008: 205-206) extends Gass and Varonis' (1984) model and mentions the contribution of the following factors in the assessment of intelligibility, comprehensibility or 'accentedness': stimulus properties (segmental, prosodic, grammatical and fluency difficulties), listener-related variables (novelty of the topic, speaker or type of accent) and contextual factors[22].

---

[21] A very similar classification of potential variables can also be found in Hodge and Whitehill (2010).

[22] Munro (2008: 212) provides the following visual representation to illustrate his model of speech intelligibility :

SCORE = SP + LF +CF + . . . + error

SCORE refers to one of 'accentedness' (A), comprehensibility (C), or intelligibility (I)

SP (Stimulus properties) = αSeg + βPros + γGram + δFlue + . . .
LF (Listener factors) = εFTop + ζFSpkr + νFAcc+ . . .
CF (Contextual factors) = θCtxt

SCORE refers to the results obtained by a subject when assessing his or her degree of intelligibility, comprehensibility or 'accentedness'. The use of Greek letters serves to weigh the importance of each specific variable. Furthermore, the values for each coefficient change depending on whether we are assessing intelligibility, comprehensibility or 'accentedness'. In addition, Munro emphasises speaker-related factors over listener-related variables in accordance with the existing literature.

Our review of the literature will follow a very a basic scheme of the communicative process, which assumes the existence of a speaker willing to convey a message to a listener through a specific medium, and will examine those variables related to the listener, the environment, the stimuli and the speaker that could potentially affect the degree of speech intelligibility.

2.3.1 Listener-related variables

As far as the study of intelligibility deficits is concerned, research has focussed mainly on the relative influence of speaker-related variables. However, in recent years, empirical research has turned to listener-related factors in an attempt to provide a more accurate picture of intelligibility loss. In the area of speech pathology, researchers have assessed the impact of listener-related variables on the speech intelligibility of individuals with dysarthria. Studies have centred, among other issues, on the different listener strategies used to understand this type of speech (e.g. Klasner and Yorkston 2005), the effects of familiarisation (e.g. Liss *et al*. 2002) or the different levels of comprehension of deviant speech (e.g. Hustad and Beukelman 2001).

With regard to L2 speech, fewer studies have examined the impact of listener-related variables on intelligibility, comprehensibility or 'accentedness'. Gass and Varonis (1984) concluded that topic, accent and speaker familiarity have a positive impact on comprehensibility. Munro and Derwing (1994) determined that familiarity resulted in listeners perceiving a higher degree of 'accentedness' when rating L2 speech samples. More recently, Zielinski (2008) analysed the difficulties experienced by a group of English listeners when orthographically transcribing speech samples from L2 speakers of English. Incorrect syllabic stress patterns resulted in listeners' inaccurate transcription of the target words.

It is worth noting that the linguistic background of both listeners and speakers seems to play a relevant role in the intelligibility of non-native speech. In general, we can say that 'native listeners find native speech more intelligible than non-native' (Hayes-Harb *et al.* 2008). In the past few years, some empirical studies have attempted to examine the validity of the so-called 'interlanguage speech intelligibility benefit (ISIB)' as it was originally set forth by Bent and Bradlow (2003). Bent and

Bradlow (2003: 1607) found a 'matched interlanguage speech intelligibility benefit' that accounted for the fact that 'for non-native listeners, the intelligibility of a high proficiency non-native talker, and in one case a low proficiency talker from the same language background, was equal to the intelligibility of the native talker'. In other words, the shared phonological knowledge between non-native talkers and non-native listeners from the same linguistic background seems to represent a benefit in terms of degree of speech intelligibility[23]. Empirical support for the ISIB can be found, for example, in Major *et al.* (2002) and Munro, Derwing and Morton (2006). It is also important to note that support for the ISIB does not seem to be consistent across all language combinations and all levels of proficiency. In Major *et al.* (2002), an ISIB was found for Spanish listeners and talkers but not for Chinese listeners. In Munro, Derwing and Morton (2006), there was an ISIB for Japanese but not for Cantonese evaluators when listening to Cantonese-accented speech. On the other hand, evidence against the ISIB can be found in Stibbard and Lee (2006). In Stibbard and Lee's study, the intelligibility of 5 Korean and Saudi Arabian speakers plus a native English speaker, with high and low levels of proficiency, was assessed by 50 evaluators from different L1 backgrounds (Korean, Saudi Arabian, native English speakers and other mixed linguistic backgrounds). Evaluators used a rating scale ranging from 1-'heavy foreign accent; very difficult to understand'- to 5-'no foreign accent; very easy to understand'. Results revealed the existence of an ISI detriment and not a benefit for speakers with a low level of proficiency. As noted by Hayes-Harb *et al.* (2008), this type of discrepancy suggests that the ISIB is mediated by more factors than the native language backgrounds of talkers and listeners.

Regarding the importance of age as a listener-related factor, empirical research shows that older adults exhibit more difficulties than younger individuals in perceiving speech in difficult environments. Harris *et al*. (2010) found an age-related difference in gap detection that resulted in differences in processing speed. Furthermore, declines in auditory processing have been underlined by several studies

[23] Bent and Bradlow (2003: 1607) reported in their study the following findings: '(1) For native listeners, intelligibility of the native talker was greater than the intelligibility of any of the nonnative talkers. (2) For non-native listeners, intelligibility of a high proficiency non-native talker (and in one case a lowproficiency talker) from the same native language background was equal to the intelligibility of the native talker. This is the ''matched interlanguage speech intelligibility benefit.''(3) For non-native listeners, intelligibility of a highproficiency non-native talker from a different native language background was greater than or equal to the intelligibility of the native talker. This is the ''mismatched interlanguage speech intelligibility benefit.''

(Dubno, Horwitz and Ahlstrom 2003; Schneider and Hamstra 1999). Empirical research has also revealed the existence of age-related declines in areas such as frequency or gap discrimination (Harris *et al*. 2008; Philips *et al.* 1994).

Even though we acknowledge the importance of listener-related factors when assessing non-native speech intelligibility, the empirical work of this dissertation will mainly focus on the role of the speaker. Nevertheless, several steps have been taken to minimise the impact of some of the aforementioned listener-related variables (a detailed account can be found in Chapter 3).

2.3.2 Environment-related variables

The acoustic properties of the environment may have an influence on the degree of speech intelligibility. Factors such as reverberation, background noise or even recording conditions have been proven to influence intelligibility measurements.

The effects of noise have been documented in both L1 and L2 speech, as well as across a wide variety of different populations. In speech pathology, negative effects of noise for segmental perception have been reported in multiple studies (e.g. Dubno, Dirks and Morgan 1984). In L2 speech, van Wijngaarden, Steeneken and Houtgast (2002) assessed the impact of speech-to-noise ratio on Dutch listeners of German and English speech with different levels of proficiency. This study concluded that non-native listeners need a better speech-to-noise ratio (1 to 7 dB) than native listeners in order to obtain at least 50% intelligibility at the sentence level. The negative effects of noise on intelligibility for foreign-accented speech were also analysed by Rogers, Dalby and Nishi (2004). Native speech was found to be more robust to the effects of applying noise to the speech signal. Depending on the noise condition, intelligibility differences were estimated at 12% to 33%. In any case, it was concluded that intelligibility of even high proficiency speakers suffered more than that of native speakers'. Munro (1998), when comparing the intelligibility of foreign-accented speech in quiet versus noise conditions, concluded that Mandarin-accented speech was less intelligible under noise conditions than native English speech. Similarly, Wilson and Spaulding (2010) examined the relationship between intelligibility, comprehensibility and noise (signal-to-noise ratio) for a group of native English speakers and a group of Korean speakers with different degrees of intelligibility.

Results showed that noise affected comprehensibility of non-native speech more than that of native English speakers'. Processing time did not, however, seem to be affected.

It is noteworthy that intelligibility studies have made use, on certain occasions, of background noise when presenting the different speech samples. This is due to both methodological, i.e. to avoid possible ceiling effects in the resulting scores, and ecological reasons, i.e. to replicate the perceptual conditions listeners may face in real life (e.g. Tajima, Port and Dalby 1997; Crocker 2010).

It is also important to point out that the provision of visual cues, along with the speaker's acoustic signal, seems to result in significantly higher intelligibility scores (see e.g. Keintz, Bunton and Hoit 2007).

2.3.3 Task and stimuli-related factors

Researchers' use of different types of stimuli and elicitation tasks has proven to influence intelligibility measurements. A perusal of the literature available reveals a wide array of elicitation tasks, which include the reading or repetition of a series of stimuli or the use of more ecologically valid speech samples involving spontaneous or semi-spontaneous speech. Stimuli can range from the presentation of isolated words to the use of semantically anomalous sentences, meaningful sentences or narrative passages.

Kempler and Van Lancker (2002) assessed the impact of five different production tasks (reading, repetition, spontaneous conversation, spontaneous singing and repeated singing) on the intelligibility measurements of a patient with Parkinson's disease. Intelligibility results were very similar for all tasks except for the spontaneous conversation one. Percentages ranged from 78%-88% for the first four tasks to a mere 29% for spontaneous conversation. The authors point to the possible role of conscious processing in an attempt to explain the inter-task differences. It is important to note that Kempler and Van Lancker's findings must be looked at rather cautiously, given that this is a case study that needs further empirical validation. Monsen (1983) assessed 10 hearing impaired subjects through the reading of a set of sentences varying in terms of phonological and syntactic complexity. Results revealed significant differences between experienced versus inexperienced listeners, between

audio-only versus audiovisual presentation of stimuli, and in terms of both phonological and syntactic complexity of sentences. McGarr (1981) evaluated the speech intelligibility of a group of deaf and normal hearing children through the reading of isolated words and words inserted in sentences that varied in terms of the redundant information available. Results indicated that intelligibility scores were higher for words presented in sentences with high redundancy as opposed to those with low redundancy. Moreover, scores were higher for words presented in sentences than for those presented in isolated form. García and Cannito (1996) assessed the intelligibility of a single speaker suffering from flaccid dysarthria. Specifically, the effects of context, familiarisation and method of stimuli presentation were analysed. Results showed that gestures and context do have an impact on intelligibility measurements. These authors also point to the existence of complex interactions between the different contextual factors under study and the different ways of stimuli presentation. Beukelman and Yorkston (1979) studied the speech intelligibility of a group of speakers with dysarthria by comparing their intelligibility scores on a single word and a paragraph transcription task. In addition, they assessed the effect of what they called 'information transfer', i.e. comprehensibility as determined by the answers given by a group of listeners on a series of questions regarding passage content. Intelligibility scores at word level were lower than those at paragraph level. More recently, Santos Barreto and Zazo Ortiz (2010) assessed the effects of different stimuli on intelligibility scores for a group of participants without any speech-related disorders. Subjects were presented with stimuli that included sentences, words and non-words. Results showed a higher degree of intelligibility for sentences as opposed to words and for words as opposed to non-words. This confirms that contextual and semantic information bears an important impact on intelligibility measurements. Furthermore, the morphological and phonological complexity of the different stimuli has also been shown to affect intelligibility scores as attested, among others, by Tikofsky and Tikofsky (1964). It is also important to note that there is a clear preference for tasks that involve the reading of words, sentences or passages. Speech samples elicited in spontaneous or semi-spontaneous production are more ecologically valid but they represent a real challenge in terms of the level of control we can have over their lexical content (see Section 2.4 for a detailed analysis of this issue).

The aforementioned studies suggest that intelligibility scores are affected by the type of stimuli, their linguistic complexity and their presentation in an isolated

manner or within higher units of discourse. In general terms, research indicates that intelligibility scores are higher at sentence than at word level. From a methodological point of view, it seems advisable to use different types of stimuli at different levels of analysis when undertaking the measurement of speech intelligibility. The empirical section of this dissertation will assess students' non-native speech intelligibility at word, sentence and passage levels, as well as through the use of speech samples elicited in a semi-spontaneous production task (a thorough description of stimuli and the reasons behind our methodological choice can be found in Chapter 3).

2.3.4 Speaker-related variables

Speaker-related variables have traditionally received more attention than listener-related factors. Given the nature of the experimental work presented in this dissertation, this section will focus on the impact of segmental and suprasegmental deviations on intelligibility scores. Our analysis assumes that pronunciation accuracy, as indicated by potential segmental and suprasegmental deviations, may have an influence on one's degree of speech intelligibility. We will therefore examine here those variables that, according to the general literature of L2 phonology, have been proven to possess an effect on the acquisition of L2 pronunciation.

The issue of L2 learners' segmental and suprasegmental deviations was at the centre of error analysis studies within SLA during the 60s and 70s. In fact, different taxonomies regarding pronunciation errors were put forward at that time. Moulton (1962) established a distinction between phonemic, phonetic, allophonic and distributional errors for English learners of German. MacCarthy (1978) presented a more 'communicative' taxonomy when distinguishing between (i) pronunciation errors that may prevent, (ii) errors that may render difficult, or (iii) errors that do not have any impact on the effectiveness of the communicative process. In any case, studies within SLA that have centred on pronunciation and intelligibility loss are rather scarce. In this section, the impact of segmental and suprasegmental deviations will be examined by drawing on some of the experimental work undertaken in the past decades both within the field of clinical speech disorders and SLA. As Rogers (1997) points out, several methods have been used to assess the potential impact of segmental and suprasegmental deviations on intelligibility scores: the construction of

a speaker's error profile by simply counting the number and type of phonetic deviations and assessing their potential impact on intelligibility scores, the use of minimal pairs tests (e.g. Kent *et al.* 1989) which seek to probe those phonetic contrasts that are likely to cause intelligibility loss, the use of acoustic measurements in an attempt to study potential deviations from 'normal' speech, as well as the use of correction studies that focus on the alteration of certain acoustic properties of the selected stimuli. In all of the above methods, intelligibility scores are usually determined by orthographically transcribing speech samples, through the use of rating scales or by choosing between several options in a multiple-choice test.

It is also noteworthy that anatomical, e.g. hypotonia, and neurological factors, e.g. difficulties in muscle coordination, may also have an impact on an individual's degree of speech intelligibility. Nevertheless, the assessment of these variables will be excluded from this study and will not be addressed in this review of the literature.

2.3.4.1 Segmental and suprasegmental deviations

Several experiments have explored the relationship between intelligibility loss and segmental deviations. This relation is not as straightforward as it may appear. In the field of speech disorders, research has often revealed contradictory results. In an oft-cited study, Smith (1975) explored segmental deviations in deaf children and their relation with speech intelligibility. A correlation between both dimensions was established at -0.8, even though a certain amount of 'dispersion' or disagreement between intelligibility and segmental deviations was observed in certain speakers. According to Smith, the impact of prosodic variables could account for these results. Whitehill (1997) assessed the speech intelligibility of Cantonese speakers suffering from dysarthria. In her study, a strong correlation was found between segmental deviations and intelligibility. Multiple-regression analyses showed that the problematic phonetic contrasts identified through a multiple-choice minimal pair test accounted for 90% of intelligibility loss. Very different conclusions were reached by Subtelny, Vanhattum and Myers (1972) when they examined the correlation between articulatory deviations and speech intelligibility of a group of 50 speakers with cleft palate. Results showed that there was a mismatch between unintelligible sentences (4%) and sentences with a serious degree of segmental deviations (34%). These

authors (1972: 24) concluded that 'speech characterized by many articulation errors is not inevitably unintelligible'. Similarly, Weston and Shriberg's (1992) literature review on the subject revealed inconsistent patterns of correlation between articulatory deviations and speech intelligibility. As Weston and Shriberg (1992: 1317) point out:

> Correlational designs using phonological process terms to describe speech sound error patterns have not resulted in a comprehensive perspective on the multiple sources of unintelligibility. In addition, methodological issues, including small sample sizes, nonsystematic definitions of sound changes, and assorted correlational assumptions, limit both internal and external validity.

All of the above has been recently confirmed by Ertmer (2010). His study on speech intelligibility for children with hearing loss revealed a weak correlation between intelligibility at word and sentence level. Furthermore, multiple-regression analyses showed that no individual variables were able to predict more than 25% of intelligibility scores. On the other hand, studies showing a strong correlation between segmental deviation and intelligibility loss can be found in Kent *et al*. (1989), where two contrasts accounted for 95.35% of score variance, or Monsen (1978), where three acoustic contrasts accounted for 73% of the variance. Weismer (2008: 575) attempts to explain these contradictory results by pointing out that:

> A whole host of phonetic and linguistic variables, including at a minimum position-induced allophonic variation, context, voice quality, prosody, *speaker* voice level, and speaker rate, cannot be controlled when a speaker with disorder produces single words, sentences, or connected speech.

In other words, methodological challenges make it very difficult to account for the vast array of variables that seem to influence the degree of speech intelligibility. Weismer further mentions that those studies that have shown the predictive power of a reduced amount of acoustic or phonetic contrast may be the result of distorted statistical analyses where variables overlap with each other and seem to exhibit a 'mutual correlation with severity of speaker involvement'.

In the area of L2 speech, research on segmental deviations and speech intelligibility has not been as prolific as within the field of speech disorders. Schairer (1992) studied the speech comprehensibility of a group of American learners of Spanish using a scale ranging from 0 to 6. Speech samples were orthographically

transcribed and Spearman rho correlations were calculated between comprehensibility and speech rate, vowel, diphthong, consonant, as well as consonant linkage production. Results showed high correlations between comprehensibility and overall vowel production (r .92), consonant production (r .62) and consonant linkage (r .84). Speech rate did not show any significant correlation.

Munro and Derwing (1995) assessed the impact of phonetic, phonemic, intonation and grammatical errors on the degree of intelligibility, comprehensibility and 'accentedness' of a group of ESL learners. Likert scales served to measure comprehensibility and 'accentedness', while a transcription task served to measure the degree of speech intelligibility. Results revealed significant correlations between 'accentedness' and phonemic, phonetic, intonation and grammatical errors. Over 70% of listeners showed significant correlations between phonemic, phonetic, grammatical and prosodic deviations and their scores on degree of foreign accent. The impact of phonemic and phonetic deviations on speech comprehensibility was much lower (44% and 11%). With regard to intelligibility, only 28% of listeners showed significant correlations for phonemic errors, 0% for phonetic and 22% for deviances in intonation. Munro and Derwing's study highlights the significant impact of segmental and suprasegmental deviations on speakers' degree of foreign accent but not on their intelligibility scores. In a follow-up study, Derwing and Munro (1997) extended the number of speakers by including talkers from different L1s. Their analysis focused on the impact of phonemic and grammatical errors, as well as speech rate and prosodic scores. Results seem to differ from their 1995 study. Even though the partial independence of the three dimensions of L2 speech under examination was confirmed, only speaking rate was correlated with scores of comprehensibility for a significant amount of listeners. Phonemic, grammatical and prosodic deviances were correlated with 'accentedness', comprehensibility and intelligibility for only a minority of listeners. We must, however, be cautious when analysing these results. As indicated in the limitations section of Derwing and Munro's study, participants' different level of proficiency when compared to the speakers from 1995 may account for the difference in results.

Rogers (1997) assessed the speech intelligibility of Chinese-accented English at word, sentence and passage level. A minimal pair multiple-choice test served to probe those phonemic contrasts that could potentially affect speech intelligibility. Results showed that sentence and word intelligibility were strongly correlated.

Multiple-regression analyses revealed that vowel deviations and not consonant deviances were correlated with intelligibility at the sentence and passage level. The phonemic contrasts probed through the minimal-pair test accounted for 76% of variance in intelligibility scores at the sentence level and only 43% at passage level.

More recently, Bent, Bradlow and Smith (2007)[24] examined the impact of segmental deviations on speech intelligibility of Mandarin-accented English. Intelligibility was calculated as the correct number of key words orthographically transcribed by native English speaker evaluators. Results showed a correlation between accurate vowel production and intelligibility. No correlation was found between consonant production and intelligibility. Nevertheless, deviations in word-initial consonants did prove to correlate with intelligibility scores.

Hardman (2010) studied the impact of a number of variables, such as speaker and listener's L1, segmental deviations and listener's word familiarity, on the intelligibility of Chinese-accented English. Stimuli were presented embedded in white noise. Results revealed that segmental deviations did not have an impact on intelligibility loss, while speaker and listener's L1, as well as word familiarity, were the variables that most affected speech intelligibility.

On a suprasegmental level, in the field of speech disorders, Weismer and Martin (1992) carried out a literature review on the effects of suprasegmental deviations on the speech intelligibility of hearing impaired individuals. These authors (1992: 81) suggest that while correlation studies have pinpointed the importance of a wide variety of suprasegmental factors, research based on the manipulation of acoustic features does not seem to confirm the key importance of the aforementioned variables for intelligibility scores. If we focus on L2 speech, several studies have centred on the impact of suprasegmentals on degree of foreign accent. In Munro and Derwing (2001), a slower speaking rate increased degree of foreign accent. Speaking rate was also found to be correlated with 'accentedness' in Gut (2007). Munro and Derwing (1995) concluded that phonemic, phonetic, grammatical deviations and intonation scores were correlated with ratings on 'accentedness'. In Flege (1988), removal of pauses in non-native speech did not have any effect on degree of foreign

[24] In their literature review, these authors suggest that the type of stimuli and the different methodological choices seem to have an influence on the potential correlation between intelligibility scores and segmental deviations. It would seem more feasible to find significant correlations between segmental deviations and intelligibility at word level and with the use of controlled elicitation tasks, such as the reading of sentences or passages, than if we were to use conversational speech samples.

accent. However, opposing results were found by Trofimovich and Baker (2006). In their study, speaking rate, peak alignment, pause length, pause frequency and stress timing were all correlated with degree of foreign accent, especially speaking rate and pause duration. Boula de Mareüil and Vieru-Dimulescu (2006) conducted two experiements in order to assess the role of segmentals and prosody on the recognition of foreign-accented speech. Using diphone speech synthesis and speech manipulation techniques, these authors concluded that, even though segmentals help, prosody seems to be a more reliable cue when identifying a foreign accent. As far as the effects of pronunciation instruction are concerned, Derwing and Rossiter (2003) linked pronunciation centred on suprasegmentals with higher gains in comprehensibility and fluency. In a similar study, Derwing, Munro and Wiebe (1998) reported higher gains in fluency for students that had received prosodic training, while no gains were reported in terms of 'accentedness' and comprehensibility.

The amount of empirical work focusing explicitly on suprasegmentals and their impact on non-native speech intelligibility is unfortunately much scarcer. Several prosodic features have been examined in an attempt to determine their impact on speech intelligibility. Tajima, Port and Dalby (1997) assessed the impact of temporal patterns of speech on the intelligibility of one native Chinese speaker. Eleven short sentences served to elicit the speech samples. These samples were presented with different levels of S/N ratio to thirty-six native English speakers who assessed their degree of intelligibility. Non-native samples were manipulated so that each segment's duration matched that of a native English speaker. Results suggest that duration has an effect on speech intelligibility, since modified non-native samples that had been matched with native English duration reached higher intelligibility scores. Using the same methodological approach, Crocker (2010) has examined the impact of duration and intonation of Hindi-accented English. Results indicate that both dimensions have an influence on non-native speech intelligibility; however, manipulation of F0 yielded a higher level of intelligibility loss than that of segmental duration. Similarly, Holm (2008) assessed the effects of intonation and duration on the degree of intelligibility of foreign-accented Norwegian. Results were inconclusive in his study due to the high degree of inter and intra speaker variability.

From the above review of empirical research, no definitive conclusions can be reached with regard to the primacy of segmental over suprasegmental elements for non-native speech intelligibility. Comparisons across studies are extremely difficult to

undertake, considering the wide array of populations under examination, the language combinations, the different techniques for measuring speech intelligibility or even the segmental and suprasegmental features analysed in each study. Furthermore, the amount of variables affecting speech intelligibility are numerous and very difficult to control under experimental conditions. Moreover, it seems rather problematic to try to generalise results outside of the specific population under study. In fact, if one of the goals is to bridge the gap between experimental research and actual teaching practice, it would seem logical to focus on specific populations of L2 learners immersed in specific educational settings. One can say rather cautiously, in light of some of the contradictory results previously examined, that both segmental and suprasegmental deviations have an effect on speech intelligibility, even though very few studies have attempted to determine their simultaneous influence on intelligibility scores. Nevertheless, further experimental work in the field will have to be undertaken before we can make any conclusive statements.

2.3.4.2 General factors affecting L2 phonological acquisition

Having seen that pronunciation accuracy may potentially affect speech intelligibility, we turn now to those general factors that, according to the literature available, seem to affect the phonological acquisition of a second language[25]. One can say that success in SLA varies a great deal from individual to individual. While all children achieve full mastery of their native language, except in severe cases of input deprivation, L2 learners experience variation in terms of rate of acquisition and ultimate attainment. In general, L2 learners reach lower levels of proficiency when compared to those exhibited by native speakers of the L2 (see Cook 2008; Muñoz and Singleton 2011 and Ortega 2009, for criticisms on using a monolingual native speaker as the yardstick for success in SLA). There is, however, some evidence of learners who manage to achieve a near-native level of proficiency in cases of short typological distance between L1 and L2 or in multilingual societies where the foreign language can be considered as an L2. There are also cases of exceptional L2 learners attaining a near-native level of proficiency in different linguistic areas, such the ones described in

[25] See e.g. Gass and Selinker (2008), Larsen-Freeman and Long (1991) and Ortega (2009) for a general review of factors affecting SLA.

Bongaerts *et al*. (1997), Ioup *et al*. (1994), Moyer (1999) and, more recently, in Piller (2002). There exist, however, numerous methodological concerns with some of these studies, as pointed out, for example, by Abrahamsson and Hyltenstam (2009).

The area of study that focuses on individual factors affecting SLA has received a lot of attention in the past decades and has yielded numerous inconclusive and controversial results. The number of factors playing a role in SLA seems to be an ongoing debate. Different variables account for different amounts of variance depending on the researcher. This is not surprising considering that methodological shortcomings prevent, in most cases, a direct comparison between studies. In the area of pronunciation, Moyer (1999) cites formal instruction, L2 exposure, motivation, attitudes towards learning, personal success in linguistic and cultural assimilation, learning style, aptitude, attitude, use of input and feedback as factors affecting L2 phonological acquisition. Purcell and Suter (1980) stress the importance of aptitude for oral mimicry, length of stay in the target-language country or concern for pronunciation accuracy. More recently, Derwing (2008) mentions several factors including age, aptitude, motivation, phonological distance between L1 and L2, and degree of L2 exposure. Similarly, Moyer (2007: 112) points to the effects of a wide array of potential variables:

> Results point to a balance of socio-psychological and exposure-type variables for predicting accent, including contact with native speakers, length of residence and age of onset, as well as intention to reside in the TL-speaking environment permanently or long-term, comfort with assimilation to the TL culture, desire to improve accent, and sense of overall fluency.

Furthermore, this author mentions (2007: 113) the importance of socio-psychological factors such as motivation, aptitude for oral mimicry, concern for pronunciation accuracy, attitudes towards the TL culture and, in general, 'the significance of learner affective factors and learner orientation to the target language'.

In the following sections, we will examine some of the aforementioned variables in an attempt to get a comprehensive picture of the factors that seem to play a major role in L2 phonological acquisition.

2.3.4.2.1 L1 effects

Interference, transfer or cross-linguistic influence, are terms that have been used to describe the role played by one's L1 during the acquisition process of an L2 (Ortega 2009). Contrastive Analysis (CA), as a theoretical framework, conforms to Behaviourist Psychology and Structural Linguistics and builds on ideas that can be traced back to authors such as Polivanov, Trubetzkoy or Weinreich. Polivanov (1931) posited the idea that we interpret L2 sounds through L1 phonetic categories. Similarly, Trubetzkoy (1939) put forward the concept of a 'phonological grid', which acts as a filter when perceiving L2 sounds. Weinreich (1953), on the other hand, formulated and defined the notion of interference as those instances of deviation from the norm of either language that occur in the speech of bilinguals. These proposals, along with Fries' (1945) contributions to the design of effective pedagogical materials in language teaching, inspired Lado (1957) to lay the foundations of the Contrastive Analysis Hypothesis (CAH). The CAH assumes that one can predict learners' errors based on a previous contrastive analysis between the L1 and L2[26]. The CAH also led to a weaker version of the theory, which aimed at analysing learners' errors *a posteriori* through a systematic comparison between L1 and L2 (Wardhaugh 1970).

It must be pointed out that, even though CA as a theoretical framework in L2 phonology has been abandoned[27], studies focusing on transfer or crosslinguistic influence have yielded interesting results in the past two decades on a segmental level (e.g. Marghany 2002; Wang and Geva 2003; Zampini 1996), a suprasegmental level (e.g. Archibald 1992; Seubsunk 2000; Zsiga 2003) and on the dichotomy between surface and abstract categories as the object of L2 transfer (see Major 2008, for a review of this issue).

[26] In terms of level of difficulty, Lado (1957: 2) predicts that 'those elements that are similar to his native language will be simple for him, and those that are different will be difficult. The teacher who has made a comparison of the foreign language with the native language of the student will know better what the real learning problem is'. As Larsen-Freeman and Long (1991) point out, the CA framework yielded not only simple binary predictions but also more sophisticated proposals involving hierarchies of difficulty. An example can be found in the oft-cited hierarchy of phonological difficulty (Spanish/English) put forward by Stockwell and Bowen (1965).

[27] Chomsky's cognitive revolution, as well as the introduction of the notion of *interlanguage* (Selinker 1972), discredited the CAH. In fact, the framework was too simplistic when attempting to associate all L2 learners' errors to instances of interference between L1 and L2 (see e.g. Selinker 1992 for a detailed discussion). The existence of errors arising as a consequence of universal influences was not acknowledged. An oft-cited example of such type of errors is the systematic devoicing of final consonants (e.g. Broselow, Chen and Wang 1998; Flege and Davidian 1984).

It is widely agreed that the L1 seems to play a more important role in phonology than in other linguistic domains (Ellis 1994; Ioup 1984). Nevertheless, the effects of the L1 on L2 phonological acquisition can only be understood when studied in combination with, among others, the following factors:

1. 'Markedness': many definitions of the term 'markedness' have been proposed over the years. Battistella (1990: 1) points out that the notion of markedness 'refers to the relationship between the two poles of an opposition; the terms marked and unmarked refer to the evaluation of the poles; the simpler, more general pole is the unmarked term of the opposition while the more complex and focused pole is the marked term'. Based on this distinction between marked and unmarked elements, Eckman (1977) put forward his Markedness Differential Hypothesis (MDH)[28] aimed at predicting the areas of difficulty that L2 learners would encounter when acquiring the L2. The MDH (Eckman 1977: 321) predicts that:

   (a) Those areas of the target language which differ from the native language and are more marked than the native language will be difficult;
   (b) The relative degree of difficulty of the areas of difference of target language which are more marked than the native language will correspond to the relative degree of markedness;
   (c) Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.

   More recently, Eckman (1991) has formulated the Structural Conformity Hypothesis (SCH) to complement his original proposal and to address some of the criticisms directed at the MDH. The SCH (Eckman 1991: 24) posits that 'the universal generalizations that hold for primary languages hold also for interlanguages'. The MDH attempts to account for those learners' errors that follow the markedness principle but do not arise in areas where a difference between the L1 and L2 can be determined. As Eckman (2008: 102) notes:

[28] Empirical support for the MDH can be found, for example, in Broselow (1983), Carlisle (1991) and Osburne (1996). See Sato (1984) for counterevidence.

'essentially, then, the SCH is the result of stripping NL-TL differences from the statement of the MDH' (see Eckman 2008 for further discussion).

2. Universals and level of proficiency: It is important to note that universal influences that operate in all natural languages seem to exercise a powerful effect on L2 phonological acquisition. On this issue, Major (2001, 2008), through his Ontogeny Phylogeny Model, stresses the role of three different forces in SLA: L1, L2 and universal and developmental processes. The presence and specific importance of these three forces will vary depending on the learner's stage of acquisition. According to this model, transfer phenomena play a crucial role during the first stages of L2 acquisition. Their number, however, seems to decrease in later stages of the learning process. On the contrary, the influence of developmental and universal factors is minimal at the beginning, increases in importance with time and finally diminishes when learners attain a higher level of proficiency[29].
3. L1 activation level: this variable seems to be related to degree of foreign accent in naturalistic contexts (Flege, Munro and MacKay 1995; Flege, Bohn and Jang 1997). Flege *et al.* (1997) assessed the oral productions of two groups of Italian speakers that differed in the amount of L1 use but had immigrated to Canada at a very similar age. Results revealed that those who reported more L1 use had, in fact, a stronger foreign accent. Similar conclusions can be found in Guion *et al*. (2000).

As seen above, transfer phenomena are especially present in L2 phonology. On occasions, however, it is difficult to identify instances of cross-linguistic influence. Moreover, even when we are able to identify those instances, it is not always easy to give an explanation of the conditions that trigger this phenomenon in the first place. The notion of 'similarity'[30] seems to play a crucial role in this regard. Wode's (1976)

[29] As Major (2008: 79) notes, the Ontogeny Phylogeny Model also attempts to explain the possible relationship between transfer, universals and similarity by claiming that 'for marked phenomena L2 acquisition proceeds slowly, transfer decreases and then decreases more slowly, universals increase quickly and decrease slowly'. Empirical support for this model can be found in Elliott (1997).

[30] The notion of 'similarity' is also at the heart of certain models of speech perception such as Flege's SLM (see our Chapter 1 for an overview of this model). It is also noteworthy that, in general, similar phenomena seem to be more difficult to learn than dissimilar (Oller and Ziahosseini 1970). Empirical studies supporting this claim can be found, for example, in James (1983) and Young-Scholten (1985).

'Crucial Similarity Measure' and Andersen's (1983) 'Transfer to Somewhere Principle' suggest that 'not only the L1 but also the L2 must have some feature that invites the (mis)perception of a similarity' (Ortega 2009: 33). Opposing this notion, Kellerman (1995) formulated his 'Transfer to Nowhere Principle' to emphasise the idea that there may be instances of transfer that are not conditioned by similarities between the L1 and L2[31]. In any case, as we have seen above, transfer is mediated by numerous factors, such as level of proficiency, L2 complexity, universal constraints etc. Future studies focusing on the interaction among these variables should be able to offer a clearer picture on the role of the L1 in L2 phonological acquisition.

2.3.4.2.2 Age-related effects on the acquisition of pronunciation

Research on age-related effects or maturational constraints on SLA is extremely abundant. This area of study has spurred vigorous debates over the years. Literature reviews in the form of state-of-the-art articles or book chapters are published on a regular basis (see Muñoz 2008; Muñoz and Singleton 2011 for some recent examples)[32]. Given that the empirical part of this dissertation focuses on contexts of formal instruction, this section will specifically examine age-related effects in that type of environment.

It is widely accepted that a critical period applies to first language acquisition (Johnson and Newport 1989; Lenneberg 1967, among others)[33]. Several studies involving children conclude that, when contact with L1 is delayed, individuals are only able to attain a very limited and irregular competence in their native language (Gleitman and Newport 1995). However, when we turn to SLA, this issue seems far from settled. In fact, not only do we find a lack of consensus among researchers, but

---

Major (2008) cites Bohn and Flege (1992) and Major (1987) as two studies that seem to contradict this claim.

[31] Kellerman (1995: 137) points out that 'there can be transfer which is not licensed by similarity to the L2 and where the way the L2 works may very largely go unheeded'. See Odlin (2003) for a criticism of the analysis of the empirical data put forward by Kellerman to support his claim.

[32] See Birdsong (2006) and Scovel (2000) for general overviews of this area within SLA. Book-length treatments can be found, among many others, in Birdsong (1999) and Scovel (1988).

[33] Lenneberg (1967: 196) formulated the basis of the CPH by arguing that 'most individuals of average intelligence are able to learn a second language after the beginning of their second decade, although the incidence of "language-learning-blocks" rapidly increases after puberty. Also automatic acquisition from mere exposure to a given language seems to disappear after this age, and foreign languages have to be taught and learned through a conscious and labored effort. Foreign accents cannot be overcome easily after puberty'. The CPH has been subject to different interpretations over the years. Consequently, a great deal of confusion has arisen in the field (see Hyltenstam and Abrahamsson 2003 for discussion).

also the mere existence of this phenomenon is minimised by some or even refuted altogether by others.

Different explanations have been put forward in an attempt to clarify the reasons behind the existence of a possible critical period in SLA. Some researchers seem to favour biological explanations that focus on brain lateralization, loss of brain plasticity or myelination processes (Lenneberg 1967; Long 1990; Pujol *et al*. 2006; Singleton 1989). Some favour the existence of a selective tuning to L1 sounds upon the establishment of the L1 phonetic system (Best 1995; Flege 1995). Cenoz and Perales (2000) and Marinova-Todd, Marshall and Snow (2000) emphasise the importance of sociocultural and motivational variables. In any case, different critical periods seem to exist for different linguistic aspects (Walsh and Diller 1981). In fact, those who favour the existence of a critical period generally base their claims on the idea that phonology is the only component directly affected by age. Even though one may conclude that the dominant opinion within the research community supports the existence of a Critical Period Hypothesis (Scovel 2000), we certainly do not lack examples of researchers that question the CPH (e.g. Aoyama *et al*. 2008) or even those who seem to accept it only partially (Flege 1995, 1999; Thompson 1991).

García Lecumberri and Gallardo Del Puerto (2003) note the close relationship between age-related effects and L2 exposure. These authors conceptualise exposure in terms of quantity of L2 input, ranging from exposure to mainly written input to cases of total immersion in naturalistic settings with no L1 contact. Quality includes cases of non-native input with a heavy foreign accent to instances of varied native input. In L2 phonological acquisition, exposure to L2 input seems to play a fundamental role. Some researchers claim that a combination of high quality and high quantity input with an early age of onset in the learning process is a good predictor of native-like attainment in the L2 (Flege *et al*. 1995; Thompson 1991). In cases of late learners, however, quality and quantity of L2 exposure do not seem to predict pronunciation accuracy (Johnson and Newport 1989; Moyer 1999).

In contexts of formal instruction, different arguments have been levelled against the Critical Period Hypothesis. Native-like attainment seems to be possible only in cases of L2 learners who start the learning process early and receive massive and native L2 exposure. It is true that, in naturalistic settings, adults show an initial advantage over children. However, in the long run, children manage to catch up and surpass adults in terms of ultimate attainment. In instructional contexts, the quality of

L2 exposure required to reveal the advantages of an early start has been equalled to 18 years of formal instruction (Singleton 1989).

García Lecumberi and Gallardo Del Puerto (2003) studied the phonological acquisition of English as a foreign language in three groups of bilingual Basque/Spanish children who started their learning process at age 4, 8 and 11. All three groups were tested after receiving six years of exposure in the L2 through formal instruction. In terms of perception, intelligibility and degree of foreign accent, older learners performed significantly better than younger ones. These authors (2003: 130) conclude that 'early starting age is not a factor which facilitates FL sound acquisition in the case of formal non-natural exposure to the FL in the medium term'.

In Catalonia, thanks to recent changes in the Spanish educational system and within the framework provided by the Barcelona Age Factor project, several studies have analysed the effects of maturational constraints in contexts of formal instruction. Fullana (2005) examined the phonological acquisition of English as a Foreign Language in several groups of Spanish/Catalan bilingual children that differed in their starting age of foreign language study and amount of L2 exposure. Differences in production were not significant and, as far as perception is concerned, early starters performed better in the long term. However, in the first stages, older learners seemed to perform significantly better.

Further research focusing on age-related effects in classroom settings is needed in order to reach more solid conclusions. Thus far, empirical results suggest that assumptions arising from SLA research in naturalistic settings cannot be automatically translated into contexts of formal instruction. As seen above, this may be due to issues relating to quality and quantity of L2 input, as well as to methodological challenges regarding the type of cognitive abilities required to perform some of the tasks used by researchers when eliciting learners' utterances (see Muñoz 2008 for a detailed analysis on the asymmetries related to maturational constraints in naturalistic and classroom settings).

2.3.4.2.3 L2 input

The amount of input in natural settings has been operationalised as length of residence (LOR) in the target language country. Empirical studies seem to provide rather varied results. Some of them have analysed the relationship between LOR and improvement

in pronunciation. Tahta, Wood and Loewenthal (1981) studied the pronunciation of 115 immigrants of different nationalities with at least two years of residence in the United Kingdom, and noted that LOR did not exert any influence on pronunciation accuracy. Piske, MacKay and Flege (2001) carried out an experiment with 72 Italian immigrants to the United States. Results resembled those of Tahta *et al.* (1981): after overcoming the initial learning period, LOR seems inconsequential to improvements in pronunciation accuracy. Moyer (1999) analysed the pronunciation of 24 Ph.D. students at the University of Austin. These students had lived for some time (an average of 2.7 years) in Germany. This author notes that the determining factor for pronunciation accuracy is the age at which L2 exposure occurred and not LOR. In contrast, other studies have established a positive correlation between LOR and accurate pronunciation. Asher and García (1969) was one of the first studies to show a correlation between both variables. Similarly, in the last two decades, Flege and his team have conducted several experiments that have yielded similar results (an exception can be found in Flege 1988). Flege *et al.* (1995) conducted a study with 240 Italian immigrants to Canada, whose average LOR was 32 years. Results indicated that LOR exerted a minor influence on pronunciation improvement. Moreover, Flege, Yeni-Komshian and Liu (1999) assessed 240 Korean immigrants living in the U.S. (15 years average LOR) and reached the same conclusion as in their previous study. Additional research has attempted to quantify L2 exposure and L2 use in order to assess their influence on pronunciation accuracy. Results are again contradictory. Flege and Fletcher (1992) maintain that, in the case of a group of native Spanish speakers living in the US, the percentage of daily use of English did not affect the acquisition of pronunciation. Thompson (1991), who examined the English pronunciation of 36 Russian immigrants to the United States, concluded that there is a direct relation between L2 use and improvement in pronunciation. Purcell and Suter (1980), however, found that there was a close relationship between input and accuracy in pronunciation for 61 respondents of different nationalities. On this occasion, input was calculated by taking into account LOR, as well as the amount of contact with native speakers at school or work. Results revealed that L2 input, after age and motivation, is one of the most important factors in the acquisition of pronunciation. Several reasons seem to account for the different results in studies focusing on the effects of L2 input. First, in terms of input measurement, the majority of studies only consider the time participants have spent in the host country. This parameter is highly

misleading. Many immigrants live, in fact, in areas with an important presence of speakers of their own language and thus, have no real need to use the L2. Furthermore, once an individual's basic communicative needs are covered, they may not experience further desire to improve their linguistic competence (as in the famous case of Alberto, studied by Schumann in 1978). Consequently, time spent in the L2 country is a relative index, considering that other factors seem to be at least as significant when quantifying L2 input.

Quality of L2 input is another factor that must be taken into account in the acquisition of pronunciation. The most beneficial type of input for L2 speech comes from interactions with native speakers (Flege and Liu 2001). In this regard, input from different native speakers seems to help L2 learners in developing the appropriate phonetic categories. It is also noteworthy that interaction with native speakers is not the only source that facilitates L2 phonological acquisition. Input from other learners (Interlanguage Talk) may also be useful in promoting linguistic interaction. This opinion is not unanimous and authors such as Flege and Efting (1987) have shown that non-native input may hinder the formation of phonetic categories. Their study examined how a group of Puerto Rican children, aged five to six, produced English voiceless stops with shorter voice-onset-time (VOT) values than those of native English speakers, due to the type of input they had received. This has led Flege (1991) to propose the 'accented L2 input Hypothesis', which posits that only access to native speech leads to the creation of the phonetic categories that will allow the learner to accurately perceive and produce L2 sounds. Finally, an aspect that must be considered when studying the influence of input is learners' level of proficiency. It has been shown that learners who are in the early stages of their learning process benefit more from an abundant L2 input than learners at a later stage (Tahta *et al.* 1981; Flege 1988). Some authors do, in fact, consider that adult L2 learners, in spite of having an appropriate L2 input in terms of quantity and quality, exhibit strong tendencies to fossilization. Moyer (1999: 88), for example, estimates that 'fossilization or stagnant variability may become apparent at intermediate levels and persist in spite of consistent and plentiful input'. It is also true that there are cases where adult learners, whether intermediate or advanced, highly motivated and with access to plentiful L2 input, have shown progress in terms of pronunciation accuracy. Proof of this is found in Dechert and Lennon's study (1989), which shows that quantity and quality of input

can lead learners to improve their pronunciation even at an advanced level of proficiency.

2.3.4.2.4 Aptitude and pronunciation

In general, aptitude refers to one's talent to acquire a new skill. Specifically, when talking about language aptitude, it alludes to an individual's potential for learning a foreign language. It is, along with motivation, one of the most influential factors on language achievement (Ellis 2005). There are different viewpoints regarding the existence of a special aptitude for language learning. Some researchers defend the existence of a special aptitude in Second Language learning and maintain that it can be analysed by means of certain instruments, such as the Modern Language Aptitude test (MLAT) or the Pimsleur's Battery of Linguistic Aptitude. Others (e.g. Krashen 1981) consider that aptitude tests possess intrinsic limitations because they do not take into account the communicative aspects of language learning. On this issue, Skehan (1989), Reves (1983) and DeKeyser (2000) provide empirical research suggesting that aptitude refers to both learning and acquisition and affects both communicative and formal contents. A third group of researchers advocates the lack of relevance of linguistic aptitude altogether (Neufeld 1979). In fact, there exists nowadays a widespread perception that the notion of language aptitude is 'anti-egalitarian' and could hinder some learners' chances of overcoming their possible low aptitude. Perhaps the most encouraging evidence offered by contemporary SLA research is the finding that aptitude, far from being a one-dimensional skill, is a multidimensional construct. This is the theoretical proposal advanced by Skehan (2002) and Robinson (2005). Their proposal can be summarised in two key points: (a) aptitude is a complex dimension that includes at least two kinds of skills: memory capacity and analytical ability, and (b) learners present complex and varied profiles of language aptitude that need to be addressed with curricular adjustments and individual attention.

As far as L2 phonological acquisition is concerned, the number of empirical studies focusing on the possible existence of a special ability in L2 pronunciation is rather scarce. As a result, Leather and James (1991) mention that we do not have enough studies to support the idea that learners' progress could be constrained by 'biological differences in awareness of, and control over, the changing configurations of the articulators, or by differences in auditory sensitivities'. Additionally, Leather

believes that the concept of a "good ear" for learning languages has not yet been incorporated as a variable in L2 speech (Leather 1999).

Some researchers have focused on the relationship between degree of foreign accent, "aptitude for oral mimicry" and musical ability. Authors such as Purcell and Suter (1980), Thompson (1991) or Flege, Yeni-Komshian and Liu (1999) seem to have identified the ability of mimicking unfamiliar sounds as a significant predictor of the degree of foreign accent. Regarding the issue of musical ability, some researchers (Nakata 2002; Tanaka and Nakamura 2004) have reported a positive relation between pronunciation and musical ability. Slevc and Miyake (2006: 679), who claimed to have carried out 'the first study that rigorously tested the musical-ability hypothesis and provided clear evidence for it', assessed 50 native speakers of Japanese who arrived in the US after the age of 11 and concluded that musical ability is a significant predictor of L2 phonological ability (receptive and productive) even when controlling for other variables. However, it did not predict variance in syntax or lexical knowledge.

Working memory capacity, as an important component of linguistic aptitude, has been widely studied over the past three decades. Early studies of individuals with exceptional memory skills were conducted by neuropsychologists during the 70s and 80s. As pointed out by Obler and Hannigan (1996), these studies showed that memory skills were correlated with L1 verbal ability.

It is useful to establish a basic distinction between long-term memory, short-term memory and working memory. Long-term memory is unlimited in terms of capacity and is made up of two memory types: explicit and implicit. Much of the knowledge that is encoded in our long-term memory is explicit, i.e. it is knowledge that we can verbalise and recall consciously. Implicit knowledge, on the other hand, refers purely to skills and habits. Tulving (2002) proposes a further distinction between episodic and semantic memory. Semantic memory points to decontextualized knowledge related to facts that we all know, while episodic memory includes knowledge of events in which the individual has been personally involved. In contrast to long-term memory, working memory is limited and appears to be intrinsically related to the notion of 'access'. Robinson (1995: 304) defines working memory as 'the workspace where skill development begins and where knowledge is encoded into and retrieved from long term memory'. In other words, working memory is essential in information storage, while playing an important role in long-term memory retrieval.

Working memory includes both controlled and automatic processes and is characterised by having a limited capacity. Information, under normal conditions, remains available in our working memory for about two seconds. After this period, information is forgotten, unless it is repeated in what is known as *phonological loop*. This repetition enables information to be integrated in our long-term memory. Given that memory is involved in information processing, it is assumed that those individuals with higher working memory capacity learn a foreign language in a more efficient manner. It is therefore widely acknowledged that working memory may help in the prediction of both rate of acquisition and ultimate attainment in SLA.

Over the past decades, Baddeley and his colleagues (2007) have studied the concept of working memory and have conceptualised it as a temporary space that allows for the storage and manipulation of information. It is a space that combines the storage function of short-term memory with the ability needed to use information effectively. Baddeley and Hitch's model (1974) consisted originally of a *central executive* and two storage components: the *phonological loop*, responsible for the storage of verbal information, and the *visual-spatial buffer*, responsible for the storage of spatial and visual information. The central executive is the component that performs most of the tasks within working memory. It coordinates the flow of information between the phonological loop and the visual-spatial buffer and directs the recovery of information stored in our long-term memory. Baddeley (2000) has recently added a new component to his model: the *episodic buffer*. The episodic buffer serves to integrate information from the phonological loop and visual-spatial buffer into our long-term memory. The emergence of this new component responds to experimental evidence showing that individuals with poor short-term memory capacity were able to store and manipulate information too complex to be stored in the limited capacity of the two slave systems. Since the central executive has no storage capacity, Baddeley was forced to put forward the existence of an additional storage component. These changes are significant, given that they have somehow weakened the separation between working and long-term memory.

As far as the measurement of the working memory capacity is concerned, researchers have used both passive tasks, which simply measure storage ability, as well as active tests that focus on processing skills. Examples of passive tasks include the repetition of digits or words. Some researchers prefer the repetition of nonsense words because they eliminate the recall strategies provided by our long-term memory.

Others favour a sentence repetition task, as a possible alternative to the mere repetition of isolated words.

As we have pointed out, working memory is used not only as a storage but also as a processing device. For this reason, researchers have devised tests to measure working memory capacity from an active perspective. Daneman and Merikle (1996) designed an instrument in which subjects had to read a series of sentences. After reading the sentences, they were asked to recall the last word of each phrase or some of the words that were underlined. This type of instrument reflects an individual's ability to keep information in the short-term memory, while performing a processing task. In the field of SLA, it seems important to determine which instruments are more effective in measuring working memory capacity. Harrington and Sawyer (1992) note that active tasks possess a greater predictive validity. In any case, it seems clear that a passive measurement of working memory capacity seems insufficient if we want to capture the contribution of this construct to the SLA process.

Much of the research on working memory in the context of SLA has been used to examine its potential association with certain language skills. Daneman and Carpenter (1980) introduced the *Reading Span Test* to examine the working memory capacity of 20 college students. Participants were required to read aloud a series of phrases and try to remember the last word of each sentence. The purpose of the experiment was to correlate working memory results with students' scores on reading comprehension. Results demonstrated the existence of a significant correlation between working memory and reading ability. Daneman and Carpenter's experiment was seminal in the sense that it established a standard in the measurement of working memory capacity.

Some researchers have examined the role of the phonological loop in the acquisition of different linguistic aspects. Ellis and Schmidt (1997) stress the importance of the phonological loop not just in vocabulary but also in syntactic acquisition. Ellis and Sinclair (1996) showed that, what is known as "subvocal rehearsal", enabled participants to achieve a better performance not only in terms of knowledge of vocabulary and pronunciation, but also in terms of understanding of syntactic and morphosyntactic rules. According to these authors, the phonological loop influences grammatical acquisition, given that it enables individuals to acquire sequences of several words in the correct order. Although most of the research on working memory capacity has been carried out in connection with reading ability,

some experiments have been devoted to speech production. Specifically, Daneman and Green (1986) concluded that level of fluency in the L1 was correlated with participants' scores in their *Speaking Span test*. In this test, subjects read the words that appeared on a computer screen. Each word was displayed for one second. After the last word, participants were required to produce a meaningful sentence using each of the words they had previously seen on the screen.

The effects of training on working memory capacity is something that has not received much attention from an empirical standpoint. McNab *et al*. (2009) concluded that adult males between 20 and 28 years were able to increase their working memory capacity after specific working memory training that lasted 5 weeks. French (2009) evaluated a group of French and Arabic children who were following an ESL course. Findings revealed that an intensive in-class practice of English structures at the phonological, lexical, and prosodic levels had a positive effect on students' phonological skills, as measured by a test involving the repetition of words. This study also showed that, even though phonological capacity remained unchanged, the efficiency with which individuals can process new information may certainly improve with training. As mentioned above, very few studies to date have examined the effects of training on working memory capacity. However, results seem to warrant further attention from researchers and teachers alike.

2.3.4.2.5 Phonological awareness

Phonological awareness refers to an individual's ability to reflect on and manipulate the sounds of a language independently of their function and meaning (Bruck and Genesee 1995). L1 phonological awareness is part of the L1 acquisition process. L2 phonological awareness is strongly influenced by L1 awareness and by its spelling system. Phonological awareness has been related to higher levels of motivation and, thus, to general success in SLA. It is also noteworthy that the notions of metalinguistic awareness (Doughty and Williams 1998) and noticing (Schmidt 1990) are at the cornerstone of recent 'focus on form' teaching methods. More specifically, research findings have been able to link phonological awareness to reading ability (see e.g. Saiegh-Haddad and Geva 2008). Venkatagiri and Levis (2007) examined the relation between phonological awareness and speech comprehensibility in a group of 17 EFL students with different L1 backgrounds. In their study, the notion of

phonological awareness was operationalised as learners' performance in a set of tasks involving phonological blending, phonological manipulation, phonological segmentation, phonological sequencing, rhyming and alliteration abilities as well as non-word reading. Two tasks to measure phonological short-term memory were also included in this study. Results showed a significant correlation between speech comprehensibility and phonological awareness. However, no significant correlation was found between speech comprehensibility and phonological short-term memory. A regression analysis determined that phonological awareness accounted for 19% of the variance in speech comprehensibility.

2.3.4.2.6 Affective variables: motivation

Affective variables refer to a wide range of factors related to emotional and even cognitive aspects of an individual's personality. Factors such as personality type, extraversion, foreign language anxiety, willingness to communicate, cognitive and learning style, to name a few, have yielded a great deal of empirical research within the general field of SLA (see Ortega 2009 for a review). Their study in connection with pronunciation accuracy or L2 phonological acquisition is much scarcer. In this section, we will look at one factor that seems to play an important role in the phonological acquisition of a second language: motivation.

The importance of motivation in language learning is recognised by teachers and students alike. At an intuitive level, motivation is usually considered a matter of quantity rather than quality. Motivation is also considered a psychological construct, something that comes "from within". It is important to point out that traditional views on this variable (e.g. Gardner and Lambert 1959) have given way to more contemporary perspectives that are based on three important principles: (a) motivation is a matter of quality over quantity (Ortega 2009); (b) motivation is primarily social, given that the social context and social relationships have a profound influence on the type of motivation that each learner has (McGroarty 2001), and (c) motivation is dynamic, because it changes at different stages of the learning process (Ushioda 2001).

With regard to L2 phonological acquisition, Piske, MacKay and Flege (2001) concluded in their extensive literature review that motivation exerts some influence on degree of foreign accent. However, they also pointed out (2001: 12) that 'factors such

as professional motivation, integrative motivation or strength of concern for L2 pronunciation accuracy does not automatically lead to accent-free L2 speech'. Furthermore, they did not consider motivational variables in their own experimental study by arguing that motivation has little effect when participants are immigrants in the L2 country. They do, however, concede that it may possess a stronger influence on those individuals who need to speak the target language without a foreign accent for professional reasons. From our perspective, the fact that motivational variables are difficult to define and to properly quantify does not necessarily mean that they do not possess a heavy influence on L2 phonological acquisition. In fact, Piller (2002: 23) suggests that 'age is not the critical factor in reaching high levels of L2 proficiency it is often assumed to be. Rather, personal motivation, choice and agency seem to be more crucial factors in ultimate attainment'. Piller's analysis offers interesting insights when linking the concept of passing for a native speaker to actual performance and not to a 'quality of being'.

Oyama (1976) and Thompson (1991) are examples of studies that found no correlation between motivation and accuracy in pronunciation. On the other hand, there exist several studies that show a correlation between motivational variables and degrees of foreign accent. Suter (1976) tested, among other factors, the relation between motivation and performance in pronunciation. Sixty-one subjects from diverse linguistic backgrounds were required to fill out a questionnaire and undergo an interview. Participants had to mimic new sounds and produce free speech on a holiday experience. Additionally, they had to complete personality and aptitude tests. Results showed that subjects' concern towards pronunciation was one of the significant factors that could account for accuracy. Purcell and Suter (1980) carried out a more in-depth statistical analysis based on data from Suter's study. Results revealed that concern for pronunciation accuracy accounted for 7% of the variance. Bongaerts *et al*. (1995, 1997) focused on highly advanced Dutch learners of English who had been exposed to extensive L2 input and received explicit training in English pronunciation. Results showed that, out of the 11 highly motivated participants, only 5 received similar scores to those obtained by native speakers. Similarly, Flege *et al.* (1995) did find that concern for pronunciation or integrative motivation were variables affecting degree of perceived foreign accent. However, these factors only accounted for less than 3% of the variance. Moyer (1999) focused on highly motivated advanced English learners of German. After several tests that involved the

reading of a list of words, sentences and a paragraph and being rated by four different judges, only one of the participants managed to pass as a native speaker. Professional motivation accounted for 41% of the variance. Moyer (2007) conducted an additional study with 50 immigrant learners of English in the US and concluded that accent is strongly linked to attitudes towards the target language such as reasons for learning English, perceived ability to improve in English or desire to improve one's accent.

In summary, results suggest that motivation does have an impact on L2 phonological acquisition. However, it seems difficult to quantify the exact nature of such influence. Additionally, there is an overwhelming preference for quantitative methods of analysis. The inclusion of qualitative methods could offer new insights concerning the interaction of motivational variables with other affective factors. Furthermore, it could also provide a clearer definition of what the concept of motivation really entails.

2.3.4.2.7 Gender[34] differences

Empirical research in L1 seems to suggest that there exist, in fact, gender differences with regard to pronunciation, i.e. women tend to exhibit better pronunciation and show a preference for more formal and prestige forms than their male counterparts[35] (e.g. Silva Corvalán 2001). When we turn to the field of L2 phonological acquisition, research does not seem to confirm the existence of gender differences regarding pronunciation accuracy (e.g. Asher and García 1969; Elliott 1995; Purcell and Suter 1980). It is also noteworthy that, in the past decade, some SLA researchers (e.g. Hansen Edwards 2006; Ohara 2001), following developments in areas such as sociolinguistics, have departed from the notion of gender as a biological construct, and have put forward a more dynamic conceptualisation based on the idea that gender is 'something we do and not something we are' (Ehrlich 1997). On this issue, with

[34] In the field of sociolinguistics, the term *sex* has been mostly replaced by the term *gender* over the past two decades. While *sex* usually refers to a mere biological or psychological distinction between males and females, *gender* is used to allude to a constructed social identity. Gender is a category that is therefore acquired through an individual's relationships with others (see e.g. Meyerhoff 2011 for a detailed analysis of both terms).

[35] In the area of language variation, three generalizations have been put forward regarding the possible impact of the gender category: (i) when it comes to stable linguistic variables, women tend to use the standard more than men, (ii) when there is a change in progress above the level of awareness, i.e. 'change from above', women tend to use the innovative and more positively judged variant more than men do, (iii) women also use the incoming variant more often than men in instances of changes in progress below the level of awareness, i.e. 'changes from below' (see e.g. Labov 1990, 2001 or Meyerhoff 2011, for an in-depth analysis and linguistic examples illustrating each principle).

regard to the potential impact of gender differences on pronunciation accuracy, Hansen Edwards (2008: 255) notes the following:

> When gender is framed and investigated as a biological construct, it does not seem to be a significant factor in L2 pronunciation accuracy. However, when gender is framed and investigated as a social construct, it does appear to impact the level of access learners have to L2 use opportunities and therefore the ability to get L2 input and negotiate meaning, which appear to affect L2 development.

It would seem, therefore, that when gender is conceptualised as a social construct, biologically-based factors interact with other variables that may arise as a result of different contextual situations. Given the limitations, constraints and research goals of this study, the notion of gender has been conceptualised here as a mere biological distinction between males and females. It is clear that the choice between a biological construct and a more socially-dynamic notion of gender is determined by the type of study, the contextual factors surrounding the research and the different methodological choices.

2.3.4.2.8 General conclusions on factors affecting L2 phonological acquisition

The field of individual differences in second language learning is exceedingly complex. Ellis (2005) points out that we do not possess a comprehensive theory to explain the influence of individual factors in SLA. Interesting proposals from Dörnyei and Skehan (2003) in the field of aptitude and motivation offer the possibility of integrating the study of individual differences within an overall framework for SLA. These new models allow for the integration not only of traditional static perspectives but also of a more dynamic dimension that reflects the changing nature of the learning situation. As far as L2 phonological acquisition is concerned, the amount of empirical research has been significantly scarce. More research is required in order to properly understand and more accurately assess the impact of constructs such as phonemic coding ability or musical ability on L2 phonological acquisition. Regarding motivational factors, Dörnyei's model and Piller's (2002) suggestion of linking the concept of passing for a native speaker to actual performance and not to a 'quality of being' could open up the possibility of exploring motivational variables from a dynamic perspective. Motivation would therefore be seen as a construct that changes depending on the nature of the social or communicative situation.

As we have seen above, there exists a wide array of variables that could potentially influence L2 phonological acquisition, as well as the degree of speech intelligibility. The choice of including the assessment of certain variables in an experimental design will depend on the nature of the study, the research goals and the feasibility of obtaining reliable measurements of the selected variables. Given that aptitude and motivation have yielded fairly strong correlations with some degree of success in the general field of SLA, we decided to include in this study the assessment of these two factors as potential individual differences that could be correlated with L2 speech intelligibility. Furthermore, the possible impact of gender, level of proficiency and L1 background will also be assessed due to two main reasons: (i) it is relatively feasible to obtain a fairly objective measurement of these dimensions and (ii) these variables have not been studied in combination with degree of speech intelligibility (see Chapter 3 for additional details).

2.4 Measuring intelligibility: some methodological considerations

Considering that the mere definition of speech intelligibility varies depending on the field of study and even on the different conceptualisations put forward by different researchers, it is not surprising that we do not have a widely agreed upon procedure to measure speech intelligibility. In this section, we will draw on the area of speech and communication disorders to examine the different instruments and approaches used in the measurement of this dimension. We will then examine some of the studies undertaken specifically in L2 speech to assess how researchers in the field of SLA have faced the difficult challenge of offering an objective measure of speech intelligibility.

2.4.1 Measuring intelligibility: insights from the area of speech disorders

As far as the area of communication disorders is concerned, attempts to measure speech intelligibility have relied primarily on two methods: (i) subjective assessments

using rating scales and (ii) identification of verbal stimuli[36]. Procedures based on rating scales are a common method in the study of human behaviour. In speech therapy, this method has been used to assess various dimensions of verbal behaviour such as voice quality, verbal fluency or degree of dysphonia. The most widely used technique of this kind is the assessment based on scales at equal intervals, whereby a listener must assign a number to a speech sample so as to evaluate its position within a continuum of the assessed dimension. Scales often use odd numbers[37] in order to have a central figure in addition to the two extremes. Some scales use qualitative descriptors associated with each score (e.g. from 1-completely unintelligible- to 5-completely intelligible), while others include only the scores. The speech samples usually consist of previously recorded sentences or paragraph fragments that are presented to the evaluator in a random order. Another less common approach is the direct magnitude estimation (DME) of intelligibility. In this case, listeners are not restricted to a pre-established number of intervals and can judge each stimulus through a number that they consider proportional to the ratio of intelligibility between samples. This procedure can be performed with or without the use of a standard or module, which would correspond to a stimulus that serves as a reference when comparing each speech sample. The choice between interval scales and direct estimation techniques has proven controversial, given that researchers have repeatedly shown that the results of an assessment through perceptual scales varies depending on the magnitude of the assessed dimension (Stevens 1975; Schiavetti 1992). Stevens (1975: 13) distinguishes between two types of perceptual continua. The first type forms an auditory continuum, known as a metathetic continuum, in which the various scale values can be described as 'degrees of magnitude or quantity'. A typical example is the perception of sound loudness. The second type is based on a prothetic continuum where each value is not perceived as an accumulation of the former, but as something qualitatively different. A prototypical example is the perception of sound pitch, since a high pitch is not perceived as an accumulation of previous tones but as something qualitatively different. Psychophysical studies show that, while a metathetic continuum can be evaluated using both types of procedures (interval and

[36] 'Amount of effort' experienced by listeners has been recently used, along with rating scales and identification tasks, in the judgement of speech intelligibility. This technique is based on the premise that a decrease in speech intelligibility results in an increase in the amount of time and mental effort required to understand speech (see e.g. Whitehill and Wong 2006).
[37] See, for example, the *Speech Intelligibility Rating Scale* (Allen *et al.* 2001).

direct estimation), a prothetic continuum cannot be assessed through interval scales, given that observers are unable to divide it into equal intervals (Stevens 1975). Since researchers have shown that intelligibility does not represent a linear continuum, a direct estimation technique would seem preferable to interval scales (Schiavetti 1992).

It is also noteworthy that the technique of identification or recognition of verbal stimuli, usually words or phrases uttered by a speaker, can be traced back to early studies on speech intelligibility. The beginnings of telephony stimulated extensive research on those physical conditions that best transmitted speech through this type of channel. It was a human listener who provided the ultimate test about the communicative efficiency of a specific system. In spite of efforts to develop mathematical algorithms, listener judgments still serve to assess the degree of communicative efficiency.

It seems, therefore, that the identification of verbal stimuli is a more appropriate technique to measure speech intelligibility than the use of rating scales. In addition to the difficulty of dividing a prothetic continuum into equal intervals, Schiavetti and other authors have put forward further reasons in favour of using identification tasks (Schiavetti 1992). First, the percentage of words or other stimuli correctly recognised by a listener is a much more manageable measurement that can lead to more reliable interpretations. Additionally, the use of identification tasks facilitates research on those acoustic characteristics of speech that have a bigger impact on intelligibility. It must be pointed out that the proper selection of stimuli is very significant in determining intelligibility values. Lists of words in isolation and sentences are the most common verbal stimuli. In terms of tasks or procedures, there is a repertoire varying in difficulty and ranging from the recognition of elements within a closed set of options, to open recognition with no alternatives.

Weismer (2008), within the field of speech disorders, mentions four types of intelligibility measures: (i) 'feature-analytic' measures aimed at providing an explanation for intelligibility deficits, (ii) transcriptions tasks, (iii) rating scales, and (iv) measurements based on listeners' accurate answers to comprehension questions. Furthermore, Weismer (2008: 570) points out that 'feature-analytic and transcription measures are most often associated with explanations of intelligibility deficits'. In fact, 'feature analytic' tests attempt to determine those factors that may cause a deterioration of the acoustic signal. In addition to their diagnostic value, they have a special clinical interest because they can facilitate the design of specific treatments in

order to target those features that possess the greatest impact on the communicative process. Kent *et al*. (1989) is an oft-cited example of this kind of test. Their study can be regarded as a systematic attempt to create an instrument that could not only offer a global index of intelligibility, but also a detailed analysis of the contributing factors to intelligibility loss. These authors selected a set of stimuli in order to reveal the most common articulatory and phonetic errors for dysarthric patients. Initially, the researchers involved in the project conducted a thorough analysis of the literature available on those English phonetic contrasts that appear to be most frequently affected in dysarthric patients. Resulting from this analysis, a list of 19 contrasts was identified as potentially responsible for intelligibility loss. The identification of these contrasts was made according to two main criteria: (i) vulnerability: they should represent phonetic contrasts affected by different types of dysarthria and (ii) acoustic correlates: these contrasts should be characterised from an acoustic standpoint in order to offer the possibility of an objective measurement. There are two versions of this test: the first version is a general multiple-choice and the second one is aimed at more severely affected patients. Stimuli are arranged so as to isolate the affected phonetic contrasts for a particular patient. A recording of each patient pronouncing each stimulus is carried out and then played to one or more listeners who must choose between the multiple options presented. This line of work has proved very fruitful both for its practical results as well as for its potential to further explain the basis of intelligibility loss.

Table 1 presents some examples of studies from the field of speech disorders and the type of measurements used in assessing speech intelligibility. Single word tests seem to be the preferred type of assessment, while the use of comprehension questions or rating scales is more unusual.

Table 1: Intelligibility measures in the area of speech disorders

Study	**Intelligibility measurement**
Beukelman and Yorkston (1979)	Single word intelligibility test (feature analytic)
Yorkston and Beukelman (1981)	Orthographic transcription
Kent *et al*. (1989)	Single word intelligibility test (feature analytic)
Schiavetti (1992)	Numerical scale
Whitehill (1997)	Single word intelligibility test
Sell *et al*. (2001)	Phonetic transcription
Whitehill and Chau (2004)	Single word intelligibility test
Jeng, Weismer and Kent (2006)	Single word intelligibility test
Klasner and Yorkston (2005)	Comprehension questions
Ertmer (2010)	Phonetic transcription (list of words)

It is noteworthy that the use of single word feature analytic intelligibility tests is not exempt from methodological problems. Intelligibility at word level may not necessarily correlate with intelligibility in connected speech. Furthermore, there are a number of problematic issues arising from the selection of the phonetic contrasts included in the stimuli, such as the presence of asymmetry biases in the test structure (see Weismer 2008 for a detailed review of this issue).

There is also a growing effort to characterise objectively and precisely those acoustic features that are primarily responsible for the deterioration of speech intelligibility. This type of research has been facilitated by the increasing affordability of the equipment required for acoustic analysis. In the field of dysarthric speech, studies comparing dysarthric and normal speech have shed some light on certain acoustic correlates that may help explain instances of intelligibility loss. Efforts have been made to relate these acoustic correlates to the information provided by perceptual judgments and identification tasks. It is, therefore, a dual approach arising from an acoustic and perceptive perspective (see Kent *et al*. 1989 for further details on this type of study).

2.4.2 Measuring intelligibility in L2 speech

It must be pointed out that intelligibility studies in the general field of communication disorders are more abundant than in the area of L2 speech. Munro (2008), in a recent literature review on foreign-accented speech, highlights the diversity of instruments used in measuring speech intelligibility, as well as the difficulty in eliciting appropriate speech samples. As far as elicitation techniques are concerned, researchers are presented with the dilemma of using controlled elicitation tasks (reading of words, sentences or passages) or using tasks aimed at obtaining more spontaneous speech samples (e.g. picture-description task). Both possess intrinsic advantages and disadvantages. Spontaneous or semi-spontaneous speech is more ecologically valid. However, it is also true that the elicitation of this type of speech sample makes any inter-participant comparisons extremely difficult. On the other hand, tasks consisting in the reading of lists of words, sentences or passages provide a higher degree of control over the lexical content. Furthermore, as noted by Munro (2008: 202), 'if intelligibility is defined as the amount of a message that is actually understood, a comparison of the intended message with the received message is essential'. This is extremely difficult to accomplish in the case of spontaneous or semi-spontaneous speech samples. In order to achieve a balanced approach, it seems logical to advocate the use of a diversity of elicitation techniques so as to obtain a clearer picture of the potential instances of intelligibility loss. As far as intelligibility measurements are concerned, Table 2 presents a sample of approaches used by L2 studies when evaluating speech intelligibility.

Table 2: Intelligibility measures in L2 speech

Study	**Type of intelligibility measurement**
Smith and Rafiqzad (1979)	Cloze test based on reading passage
Smith and Bisazza (1982)	Multiple-choice questions based on reading passage
Gass and Varonis (1984)	Orthographic transcription of sentences: scores were based on deviations
Fayer and Krasinski (1987)	Five-point scale
Perlmutter (1989)	Summary of main idea

Anderson-Hsieh and Koehler (1988)	Comprehension questions
Anderson-Hsieh, Johnson and Koehler (1992)	Seven-point scale
Munro and Derwing (1995)	Orthographic transcription task: scores based on deviations between transcript and intended message
Derwing and Munro (1997)	Orthographic transcription: percentage of exact word matches
Rogers (1997)	Forced choice: minimal pairs task Orthographic transcription of sentences and passages
Munro, Derwing and Morton (2006)	Orthographic transcription task: scores based on deviations between transcript and intended message
Burleson (2007)	Forced choice task: identification of minimal pairs
Zielinski (2008)	Orthographically transcribed sentences from conversational speech
Isaacs (2008)	IPA transcription Impressionistic judgements on intelligible and unintelligible pronunciation
Holm (2008)	Orthographic transcription
Hardman (2010)	Orthographic transcription
Quené and van Delft (2010)	Speech Reception Threshold Method

The majority of studies seem to opt either for rating scales or for orthographic transcriptions in the assessment of non-native speech intelligibility. It is noteworthy that the use of rating scales, a purely impressionistic way of assessment, makes it difficult to establish a distinction between intelligibility and perceived comprehensibility. If we opt for an orthographic transcription of speech samples, a decision regarding how to score evaluators' transcriptions must also be made. The scoring procedures at the sentence level have ranged from the count of correctly transcribed words (e.g. Maasen and Povel 1985), correctly transcribed key words (e.g. Bradlow and Bent 2002), or correctly transcribed sentences as a whole (e.g. Benoît 1990). Rogers (1997), in her review of the literature, distinguishes between rating

scales, phonetic transcription and identification tasks. This author concludes that, while rating scales do not offer information about the types of errors that may cause intelligibility problems, identifications tasks can be used as a diagnostic tool in explaining intelligibility deficits. In her own study on intelligibility of Chinese-accented English, Rogers (1997) first analysed the patterns of error of two Chinese learners of English in order to create a minimal pair test to cover all the error inventories. The minimal-pair test, a series of sentences, as well as a passage were then presented to a group of native English speakers. Correlation and multiple-regression analyses were used to establish the relation between intelligibility scores and the minimal pair phonemic contrasts.

In summary, the chosen type of instrument and measurement procedure depends on the definition of speech intelligibility put forward by each researcher, as well as on the specific goals of each study. In any case, from the overview of the literature presented above, we can conclude that rating scales are of little use when the research goals centre on obtaining information on those segmental or suprasegmental deviations that may be responsible for intelligibility loss. On the other hand, the use of multiple-choice tests, along with the transcription of subjects' productions by a group of native speaker evaluators, seems to be more appropriate when the research goal centres on obtaining diagnostic information. It is also true that, while multiple-choice tests can be administered and scored in a somewhat straightforward manner, transcription studies represent a time-consuming undertaking. Nevertheless, transcription studies seem an appropriate option when the assessment centres on connected speech, given that the use of single word intelligibility tests does not seem to be feasible for this level of analysis.

At word level, researchers have generally followed a three-step approach when their intention is to construct a multiple-choice test aimed at identifying those deviations that may be responsible for intelligibility loss:

1. Study of patterns of error for the specific population.
2. Creation of a test based on the error inventory. This type of instrument should be, according to Kent *et al.* (1989), reliable, efficient, tailor-made to a specific population, easy to administer, objective and must provide an overall measurement of intelligibility (see Rogers 1997 for a detailed explanation of these features).

3. Correlation between intelligibility scores and the different phonetic contrasts identified in step 2. Multiple-regression analysis can also be used to generate models to explain the interaction between different factors.

Our study will follow these steps when attempting to offer a measure of intelligibility loss at word level. This will be combined with further assessments of the degree of non-native speech intelligibility for the population under study at sentence and passage levels, as well as in a semi-spontaneous production task. In connected speech, the elicitation of speech samples is often achieved through the use of phonetically-balanced sentences and phonetically-balanced reading passages. The elicitation of samples in spontaneous or semi-spontaneous productions has traditionally represented a bigger challenge. In our study, speech samples will be elicited through the use of a series of phonetically-balanced sentences and reading passages. Furthermore, a tailor-made production task will be designed in order to achieve the elicitation of speech samples in semi-spontaneous production. The procedures for the evaluation of intelligibility loss in connected speech will also include the transcription of our participants' productions by a group of 60 native speakers of Peninsular Spanish (a detailed description and further reasons behind our methodological choice will be presented in Chapter 3).

2.5 Summary and conclusions

This chapter has offered a theoretical framework for the experimental part of this dissertation. The concept of speech intelligibility has been examined by drawing on multiple disciplines and research fields. The intrinsic difficulty of providing a definition of speech intelligibility has become clear. In fact, definitions seem to vary depending on the discipline, the author and the type of measurement used in assessing intelligibility. Additionally, the relation of speech intelligibility with other dimensions of L2 speech, namely comprehensibility and 'accentedness', has also been put forward. The relation between these three dimensions is not as straightforward as previously thought, and the experimental work carried out by Derwing and Munro in the past two decades has revealed that intelligibility and 'accentedness' are not necessarily correlated. One of the difficulties in undertaking any research project in

the area of speech intelligibility stems from the number of factors that seem to have an impact on this dimension. We have examined speaker-related variables, including general factors that affect L2 phonological acquisition, listener-related factors, variables related to the environment, as well as those arising from the task or stimuli used when measuring speech intelligibility. Given the nature of the empirical section of this dissertation, special attention has been devoted to studies, both within the fields of speech pathology and L2 speech, that have focused on the impact of segmental and suprasegmental deviations on speech intelligibility. Moreover, problems related to the type of intelligibility measurement further complicate the issue of providing an accurate assessment of speech intelligibility. Researchers' choices have varied between subjective measurements using rating scales, orthographic transcriptions and the use of multiple-choice tests using minimal pairs in an attempt to identify those phonetic contrasts responsible for intelligibility loss. From our analysis of the literature, we can conclude that more studies focusing on specific populations of L2 learners are necessary. Furthermore, the assessment of intelligibility must be carried out at different levels of analysis, and not only at word level, if one intends to offer useful empirical data that could yield interesting insights immediately applicable in the foreign language classroom.

In Chapter 1, we presented the implications of the *intelligibility principle* for the field of pronunciation instruction. The unrealistic goal of attaining a native-like pronunciation has been abandoned in favour of a more realistic objective: the achievement of comfortable intelligibility. Moreover, it has become apparent that intelligibility studies have largely focused on English as a second language. This interest on speech intelligibility within the context of English as an international language has generated proposals such as Jenkins' Lingua Franca Core[38] (Jenkins 2000) that, in turn, have yielded a great deal of empirical research. Even within L2 speech, some interesting principles such as the 'Functional Load'[39] (Munro and

[38] Jenkins (2000) put forward the Lingua Franca Core to emphasise the importance of certain segmental elements based on results arising from studies on intelligibility and communicative interactions between non-native English speakers. The relevance of the non-native perspective serves to acknowledge that there are more than 300 million non-native speakers using the English language to communicate among themselves. The Lingua Franca Core includes consonant sounds (except for dark /l/, dental and fricatives), vowel length contrasts, consonant clusters (initial and medial but not final) and nuclear stress.

[39] King (1967: 831), as cited by Munro and Derwing (2006), defines Functional Load (FL) as 'a measure of the work two phonemes do in keeping utterances apart'. Different approaches have been used in an attempt to rank segmental contrasts based on their importance for English pronunciation (see

Derwing 2006) have attempted to provide an explanation for those features affecting the degree of speech intelligibility. However, when one looks outside the realms of English as second language, it becomes obvious that very little is known about the specific L2 features that may hinder an L2 speaker's degree of speech intelligibility. Our next chapter will present the methodology adopted in this study. Our primary goal will centre on offering a comprehensive assessment of the non-native speech intelligibility of a specific population of foreign language learners immersed in a specific educational context: a group of English learners of Spanish at Key Stage 4.

Brown 1991; Catford 1987). Munro and Derwing (2006) consider that FL can be measured by the number of initial and final minimal pairs between two sounds and their importance for different linguistic varieties. These two authors carried out a study in which a group of 13 native English listeners assessed 23 sentences produced by Cantonese learners of English. Sentences contained both high and low functional load errors. Native speakers rated the sentences for 'accentedness' and comprehensibility using a Likert scale ranging from 1 to 9. Results showed that high frequency load errors did have an impact on comprehensibility and 'accentedness', while sentences with low functional load errors only had a minimal impact on comprehensibility.

Chapter 3: Methodology

3.1 Aims and research questions

As mentioned in Chapter 2, the 'intelligibility principle' has been widely adopted as the goal for pronunciation instruction. Chapter 2 has also highlighted that our knowledge of the contribution of specific segmental and suprasegmental elements to intelligibility loss is rather limited. In an attempt to fill some of the existing research gaps, this dissertation centres on the assessment of the speech intelligibility of a group of 20 Key Stage 4 English learners of Spanish at word level (through the use of a single word intelligibility test), at sentence and passage levels (using the Spanish version of the Harvard Psychoacoustic Sentences and the reading of a phonetically-balanced passage) and in a semi-spontaneous production task. Furthermore, this dissertation seeks to determine the impact of segmental and suprasegmental deviations on intelligibility loss, as well as to examine the potential correlation between certain individual variables (L1, level of proficiency, gender, aptitude and motivation) and learners' intelligibility scores. It is noteworthy that only the impact of segmental and certain suprasegmental deviances, as well as the aforementioned speaker-related variables will be assessed. Other possible linguistic and non-linguistic contributing factors, such as grammatical or lexical accuracy, will be excluded from this study. Additionally, no attempt will be made to analyse those listener-related variables that could have an effect on students' intelligibility scores.

This chapter will describe the methodology adopted in this study. Some methodological issues related to research in educational contexts, as well as a description of participants, evaluators, materials and procedures will be presented below. The following research questions were initially formulated:

1. Is there an intelligibility loss when assessing the non-native speech intelligibility of our group of Key Stage 4 learners of Spanish?
2. Is there any correlation between the intelligibility ratings at word level and in connected speech?

3. At word level, what type of phonemic contrasts seem to impact the most on speech intelligibility as determined by our single-word intelligibility test?
4. Do any of the suprasegmental features examined in this study, i.e. speech rate, pause frequency and pause duration, have any impact on students' intelligibility scores in connected speech?
5. Do the appropriate statistical analyses highlight the importance of segmental over suprasegmental deviances or vice versa?
6. Can we find any correlation between students' intelligibility scores and certain individual variables such as L1, gender, motivation, aptitude and level of proficiency?

It can be hypothesised that our group of participants will experience some degree of intelligibility loss. However, given the contradictory results highlighted by the review of the literature undertaken in Chapter 2, it is more difficult to predict the extent of the impact of segmental and suprasegmental deviations on speech intelligibility, as well as the possible correlation between intelligibility scores at different levels of analysis. With regard to the correlation between certain individual differences and degree of speech intelligibility, a positive relation between both sets of variables could *a priori* be predicted. However, we must be extremely cautious with these initial hypotheses. Once again, the empirical research undertaken thus far has been unable to provide any conclusive results, partly due to the wide range of variables involved in these types of studies.

An additional goal of this dissertation is to offer some insights for pronunciation instruction. Based on our empirical findings, we seek to provide some pedagogical insights for the teaching of Spanish pronunciation aimed at the population of L2 learners under study. Thus, an attempt will be made to fill the gap that usually exists between empirical research and actual teaching practice.

3.2 Some notes on research paradigms

From a purely methodological standpoint, research can be divided into exploratory, descriptive, correlational and explanatory[40] (Hernández Sampieri, Fernández Collado and Baptista 1997). In practice, any study can combine elements from these four research paradigms. Exploratory studies are used to prepare the ground for further research. Descriptive studies seek to examine the properties of a specific phenomenon under analysis. On the other hand, correlation studies[41] attempt to determine the relationship between different variables through the so-called correlation coefficient. The magnitude of this coefficient describes the strength of the relationship between two variables. A positive coefficient indicates that the factors under study co-vary in the same direction. A negative coefficient indicates that the variables in question co-vary in opposite directions, i.e. an increase in one of them is accompanied by a decreasing value in the other. The magnitude of the coefficient varies between 0 and 1. The greater the degree of co-variation between variables, the closer the correlation coefficient will be to 1. It is important to point out that correlation does not, in any way, mean causation (see the limitations section of this study). Finally, explanatory studies go beyond the mere description of phenomena or the establishment of relationships between variables and seek to explain the causes of physical or social events. Even though studies may essentially be described as pertaining to one of the aforementioned categories, they may also contain a combination of elements from the others. Moreover, an investigation may start as exploratory or descriptive and later become correlational and even explanatory. In our case, this study attempts to explore

[40] Alternative taxonomies of research paradigms can be found, for example, in Brown and Rodgers (2002). These authors distinguish between primary and secondary research. Primary research, i.e. research based on original data, can be divided into qualitative, survey and statistical research (descriptive, exploratory, quasi-experimental and experimental research). Other authors (e.g. Larsen-Freeman and Long 1991) introduce a distinction between longitudinal and cross-sectional studies. In any case, researchers seem to uphold, in different degrees, the latent and traditional distinction between qualitative and quantitative research.

[41] See Cohen, Manion and Morrison (2005) for a discussion of correlational studies within educational research. It is also noteworthy that these authors (2005: 199-200) distinguish between two types of correlational studies: 'relational' and 'prediction' studies. Relational studies are 'mainly concerned with achieving a fuller understanding of the complexity of the phenomena or, in the matter of behavioural or educational research, behavioural patterns, by studying the relationships between the variables which the researcher hypothesizes as being related'. Prediction studies, on the other hand, assume 'that at least some of the factors that will lead to the behaviour to be predicted are present and measurable at the time the prediction is made'. In our case, the present research falls within the category of 'prediction studies', given that we will be focusing, exclusively, on a series of factors that, *a priori*, we know may potentially have an impact on instances of intelligibility loss.

the possible relation between intelligibility loss, segmental deviances, suprasegmental deviations and certain individual differences. Thus, this project can be described as an example of correlational research and, consequently, correlation and multiple-regression analyses will be used in our data analysis.

It is also noteworthy that Grotjahn (1987) further classifies classroom research in accordance with three criteria:

1. The treatment of data: qualitative or quantitative
2. The method of data analysis: interpretive or statistical
3. The experimental or non-experimental way of obtaining the data.

The combination of these criteria gives rise to eight paradigms: two pure and six 'mixed' (see Table 3 below). As Nunan (1992) points out, the use of some of these paradigms is extremely rare. Furthermore, it is noteworthy that the different research paradigms must be seen as a full continuum of hybrid examples and not in dichotomous terms (Larsen-Freeman and Long 1991).

Table 3: Grotjahn's research paradigms (Nunan 1992: 6)

Mixed paradigms	**Pure paradigms**
Experimental-qualitative-interpretive 1. experimental/quasi-experimental 2. qualitative data 3. interpretive analysis	Exploratory-interpretive 1. non experimental design 2. qualitative data 3. interpretive analysis
Experimental-qualitative-statistical 1. experimental/quasi-experimental 2. qualitative data 3. statistical analysis	Analytical-nomological 1. experimental/quasi-experimental 2. quantitative data 3. statistical analysis
Exploratory-qualitative-statistical 1. non-experimental design 2. qualitative data 3. statistical analysis	
Exploratory-quantitative-statistical 1. non experimental design 2. quantitative data 3. statistical analysis	
Exploratory-quantitative-interpretive 1. non experimental design 2. quantitative data 3. interpretive analysis	
Experimental-quantitative-interpretive 1. experimental/quasi-experimental 2. quantitative data 3. interpretive analysis	

This study uses a quantitative approach in the treatment of data, as well as statistical methods of data analysis. Moreover, data has been obtained under control conditions that involve a careful selection of stimuli and testing procedures. In Grotjahn's terminology, an experimental study requires the selection of an experimental and control group of participants, and some degree of intervention through the manipulation of one or more variables. In our case, there was no selection of experimental and control groups as such, and there was no intervention in terms of variable manipulation. Our study can therefore be framed within a non-experimental, quantitative-statistical research paradigm.

3.3 Ethical considerations

Respecting the privacy and confidentiality of participants' data is at the heart of ethical conduct in research (Folkman 2001). Privacy involves the subject's right to choose what information to share, at what time, and under what circumstances. It also concerns the participant's right not to give any information that he or she does not want to share. Confidentiality agreements refer to what responsible researchers can do with the data. Sharing data with colleagues or funding agencies is part of the research process and, therefore, the conditions under which this should be done need to be part of the agreement. Moreover, the privacy and confidentiality agreement must satisfy the need for participants' privacy, while allowing the research to go on. Researchers should be aware that any violation of this agreement may have serious implications.

An ethical responsibility arises from the recruitment of participants. The possibility of coercion or just the mere appearance of coercion should be carefully avoided when looking for cooperation. Educational researchers, in particular, must be aware of the rights, dignity and welfare of the subjects concerned and should inform them of the type of study in which they are getting involved (Fischman 2001). Informed consent includes a clear explanation of the purposes, procedures, risks and benefits of the research, along with the obligations and commitments of both participating subjects and researchers. As Fischman (2001) notes, the informed consent must be voluntary and given by a competent subject. We must ensure that this consent is obtained prior to the subject's inclusion in the research project.

In this study, we made sure that all participants understood the different steps involved in the research process and why their participation was necessary. It was paramount to obtain the Head of the School's as well as each of the subjects' consent before the research began. All participants, as well as their parents or guardians, were required to read and complete the appropriate information and consent forms (Appendix A). Participants were also informed of their right to withdraw from the research at any time. Additionally, we sought, at all times, to minimise the impact of this study on students' regular school workload. Confidential treatment of subjects' data was maintained throughout this study[42].

3.4 Participants

Any research project that involves L2 learners receiving formal language instruction is very difficult to undertake. The level of difficulty increases even further when those students are immersed in a secondary school setting. A perusal of the relevant literature points out that the majority of researchers focusing on L2 learners in foreign language contexts tend to recruit subjects from university settings. As Derwing and Munro (2009) note, there is a need for more research in secondary school contexts and outside of the ESL field, especially if we intend to bridge the existing gap between research and pedagogical practice.

Twenty Key Stage 4 students (Year 11) from a mixed-ability Spanish language class participated in this study. Participants belonged to an average-size secondary school in Hertfordshire, England. The proportion of students from minority backgrounds was half the national average, while that of students exhibiting learning difficulties was below average. The latest Ofsted report awarded outstanding or good grades in the following areas: overall effectiveness of the school, effectiveness of the sixth form, achievements and standards, students' personal development, teaching and learning areas, leadership and management. The percentage of students with GCSE

[42] Cohen, Manion and Morrison (2005: 70) note the following on the treatment of data: 'of the two most important principles which do concern research data, one states that personal data (i.e., data that uniquely identifies the person supplying it) shall be held only for specified and lawful purposes. The second principle states that appropriate security measures shall be taken against unauthorized access to, or alteration, disclosure, or destruction of personal data and against accidental loss or destruction of personal data'. It is important to point out that this study adhered to both principles at all stages of the research process.

results (including Maths and Science) of A-C had increased steadily in the last five years and had reached over 75% in the last academic year. In terms of foreign language provision, French was the main foreign language in Key Stage 3 and 4. Spanish becomes an option in Year 10 and is taught in mixed-ability classes. There were plenty of opportunities for extra-curricular activities and community-based programmes.

All students completed and returned a preliminary questionnaire (Appendix B), along with the aforementioned consent and information forms. Only English learners of Spanish with no hearing or speech impediments, no prolonged periods of stay in Spanish speaking countries and no exposure to Spanish at home were considered for this study. Twenty-two participants initially met our set of criteria. Two of them did not complete all the recording sessions and were therefore excluded from this study. Our preliminary questionnaire gathered information on the following: gender, age, place of birth, native language, father's L1, mother's L1, languages spoken at home, period of time spent in Spanish speaking countries, known hearing or speech problems, number of years of Spanish study, previous study of any other foreign languages and practice of Spanish outside of school.

Table 4: Summary of participants' data

Student	Age	Gender	L1
1	16	F	English
2	15	F	English
3	15	M	English
4	15	F	English/Polish
5	16	M	English/Polish
6	15	M	English
7	16	F	English
8	15	M	English
9	16	F	English
10	15	F	English/Punjabi
11	15	M	English
12	15	M	English/Polish
13	16	F	English/Punjabi
14	15	F	English/Polish
15	16	F	English/Polish
16	15	M	English
17	15	F	English
18	16	F	English
19	16	M	English/Punjabi
20	16	M	English/Punjabi

It is worth pointing out that Spanish was, for all participants, a second foreign language, since the study of French is compulsory at Key Stage 3. Subjects had been studying Spanish for an average of two years and had not received any extra exposure to Spanish outside of school. None of the participants had spent a significant amount of time in Spanish speaking countries. Fourteen subjects reported on several short stays in Spain for a maximum of two weeks (summer holidays, school trips). All students were born in England and can be considered as native English speakers, even though five of them had Polish-speaking parents and used Polish on a regular basis at home. Additionally, four participants had Punjabi-speaking parents and also used this language at home. In terms of Spanish practice outside of the classroom, scores on the Likert scale ranged from 1-less than 1 hour a week- to 5-more than 4 hours a week- with a mean of 2.4, which indicated that students claimed to spend over two hours per week outside of the classroom involved in activities that were related to Spanish language learning.

Students were also assessed in terms of aptitude, motivation and level of proficiency. The materials used in assessing these variables will be presented in Section 3.7. Results will be provided in Chapter 4.

3.4.1 Some notes on the Spanish curriculum at Key Stage 4

All participants in this study belonged to a Year 11 class following a Spanish language course as part of their preparation for their General Certificate of Secondary Education (GCSE) examinations. Year 11 is a key year in the foreign language curriculum, given that students finalise their preparation for their GCSEs and take the different examinations in the second half of the academic year. It is also noteworthy that preparation for the GCSE examinations usually begins in Year 10. The key role of Year 11 within the national curriculum is one of the reasons that led us to focus on this specific population of L2 learners.

Since September 2004, the study of a foreign language is not a statutory requirement at Key Stage 4. In other words, schools are not required to teach a foreign language to their students but they must make courses available to students who wish to do so. To fulfil the statutory requirement that applies to entitlement areas (Qualifications and Curriculum Authority 2004: 6):

> Schools must provide access for all students at key stage 4 to a minimum of one course in a modern foreign language that leads to a qualification approved under Section 96 of the Learning and Skills Act 2000. Schools must provide access to at least one such course in an official working language of the European Union (EU), but may in addition offer courses in any other modern foreign language that lead to approved qualifications.

At the end of Key Stage 4, students usually sit their GCSE in multiple subjects. GCSE exams, although not compulsory, are the most common qualification awarded at the end of compulsory secondary education in England. GCSE exams are administered by five different boards: Assessment and Qualifications Alliance (AQA), Oxford, Cambridge and RSA Examinations (OCR), Edexcel, Welsh Joint Education Committee (WJEC) and the Council for the Curriculum, Examinations and Assessment (CCEA). All boards are regulated by the Qualifications and Curriculum Authority.

The Qualifications and Curriculum Authority points out (2004: 8-9) that the Modern Foreign Languages entitlement area should enable students to:

- speak and write in a modern foreign language
- understand and respond to speech and written materials in that language
- learn about the culture of countries or communities where that language is spoken
- acquire knowledge and understanding of the language(s) studied
- develop language skills
- develop language-learning skills
- develop cultural awareness

A through description of the programme of study is left to the different examination boards that oversee the GCSE examinations in each language. In the case of our school, the OCR examination board was in charge of overseeing the GCSE examination in Spanish. The OCR specifications (2009: 5) set the following goals for students at this level:

- derive enjoyment and benefit from language learning by following a coherent, satisfying and worthwhile course of study
- develop understanding of the language in a variety of contexts

- develop knowledge of the language and language learning skills
- develop the ability to communicate effectively in the language
- develop awareness and understanding of countries and communities where the language is spoken
- recognise that their linguistic knowledge, understanding and skills help them to take their place in a multilingual global society and also provide them with a suitable basis for further study and practical use of the language

These goals are similar to the ones established by the National Curriculum. Moreover, the OCR board sets eight topic areas, including home, health, leisure, travel and education, in order to give the necessary context for learners to explore the target language. This board also provides lists with the compulsory linguistic structures and lexicon that candidates are required to understand. A detailed marking scheme for each section of the examination is also provided. It is noteworthy that no reference is made to issues related to pronunciation. Accuracy in pronunciation is not mentioned at all in the marking scheme used in assessing the oral section of the exam. This absence of pronunciation/phonological issues lies in sharp contrast with the provisions of other internationally recognised instruments such as the Common European Framework of Reference for Languages (CEFR). The CEFR mentions the phonological competence as part of a general linguistic competence and makes explicit the knowledge and skills involved in the perception and production of the different phonetic features (Council of Europe 2001: 116)[43].

[43] The CEFR (2001: 16-17) points out the following on the issue of phonological competence: 'phonological competence involves a knowledge of, and skill in the perception and production of:
• the sound-units (*phonemes*) of the language and their realisation in particular contexts (*allophones*);
• the phonetic features which distinguish phonemes (*distinctive feature*s, e.g. voicing, rounding, nasality, plosion);
• the phonetic composition of words (*syllable structure*, the sequence of phonemes, word stress, word tones);
• sentence phonetics (*prosody*)
• sentence stress and rhythm
• intonation
• phonetic reduction
• vowel reduction
• strong and weak forms
• assimilation
• elision'

This absence of a systematic treatment of pronunciation instruction both at the National Curriculum and OCR levels was reflected on the Spanish language course our participants were required to follow. Classes respected the general division in topic areas set by the examination board. As far as the teaching provision was concerned, instructors made use of IT resources and implemented what can be regarded as a communicative approach with certain elements from task-based learning methodologies. No specific provision was made for the teaching of pronunciation. Occasional corrections by the instructor arose as a response to segmental deviations in students' productions. There was no treatment whatsoever of deviations at a suprasegmental level. Furthermore, the students' textbook made no reference to pronunciation/phonological issues and did not present any exercises for the treatment of pronunciation errors.

3.5 Evaluators

Sixty evaluators, (33 male and 27 female) native speakers of Peninsular Spanish (Santander) were recruited to assess the speech samples collected in this study. Listener-related variables, such as familiarity with foreign accents, previous training or linguistic background, have been proven to have an effect on intelligibility and comprehensibility ratings (see Derwing and Munro 2009; Kennedy and Trofimovich 2008; Rogers 1997 and our Chapter 2 among others). In order to neutralise the effect of some of these factors, we decided to recruit naïve listeners with no hearing or speech impediments, from different social and educational backgrounds, from a broad age range, with no regular contact with English speakers and with no extended periods of stay in English-speaking countries. All evaluators completed a preliminary questionnaire (Appendix C) which inquired about the following aspects: age, gender, place of birth, native language, father's native language, mother's native language, hearing or speech impediments, foreign languages spoken and level of proficiency, highest level of formal education achieved, profession, contact with English speakers (the scale ranged from 1-no contact at all with English speaker- to 5-regular contact on a daily or weekly basis), periods of stay in English speaking countries (scale ranging from 1-no time spent in English speaking countries- to 5-over six months).

All subjects were native Spanish speakers with no reported hearing or speech impediments. Their mean age was 26.5. In terms of educational background, 35% held university degrees, 15% were studying towards their Bachelor's Degree, 30% held secondary school qualifications (*Bachillerato, Formación Profesional*) and 20% had either diverse continuing education certificates or had completed the E.G.B. (*Enseñanza General Básica*). None of them had regular contact with English speakers and they did not report on any long-term stays in English speaking countries. On occasion, they did report on holiday periods spent in English speaking countries but always for a maximum of up to two weeks. As far as knowledge of foreign languages is concerned, 45% of evaluators did not speak any foreign languages at all, 30% claimed to speak French at a beginner or intermediate level of proficiency, while 25% claimed to possess a beginner level of proficiency in English. It is worth pointing out that any individuals that reported an intermediate or advanced level of proficiency in English were automatically excluded from this study.

3.6 Native Spanish speaking subjects

Two native Spanish speakers (one male and one female of 16 years of age at the time of the study) provided the necessary baseline data and recorded all the stimuli used in assessing speech intelligibility at different levels. Neither subject presented any speech or hearing impediments, and both were students in a Spanish secondary school and used a peninsular variety of Spanish (both were living in Santander at the time of the recordings).

3.7 Materials

This section will present the materials used in this study. First, a description of the instruments used in the assessment of students' level of proficiency, motivation and aptitude will be put forward. Furthermore, an analysis of the materials used in evaluating speech intelligibility at the word, sentence and passage levels, as well as in the semi-spontaneous production task will be presented. With regard to the assessment of intelligibility, a single word intelligibility test will serve to measure intelligibility at word level. A transcription task involving the Spanish version of the

Harvard Psychoacoustic Sentences and a phonetically-balanced text will be used in the assessment of intelligibility at the sentence and passage level. A further transcription of speech samples obtained through a semi-spontaneous production task will serve to assess our participants' degree of speech intelligibility at this level of analysis.

3.7.1 Assessing motivation

Students completed a motivation questionnaire (Appendix D) to assess their level of motivation towards the learning of Spanish pronunciation. The motivation questionnaire was adapted from Wen (2005) which, in turn, was based on Scherer (1984) and Gardner's Attitude/Motivation Test Battery (AMTB). The test consisted of 18 questions, which were rated on a scale of 1 to 10 ('strongly agree to strongly disagree'). Wen's questionnaire was originally designed to study learners' motivation in a Second Language environment, i.e. an environment where students are learning the L2 in the target language country. Thus, only questions applicable to contexts related to foreign language learning were used in this study. Questions 3, 8, 12, 13, 16, 18, 19, 20, 22, 23, 24, 25, 27 and 28 were eliminated altogether from the original questionnaire and questions 11, 14, 17 and 21 were adapted to make them more suitable for the specific population under study. Nevertheless, the same four areas of motivation were examined: novelty, i.e. students' reaction to new material, pleasantness, which alludes to the pleasantness or unpleasantness that arises out of a stimulus, coping material, which refers to a student's ability to adjust to a specific stimulus, and goal significance, i.e. the relevance of a specific stimulus with regards to an individual's goals (see Schumann 1999 for a detailed analysis of models of stimulus appraisals). It is also noteworthy that we reduced the Likert scales from 10-point to 5-points.

This questionnaire can be criticised on a number of reasons, some of which will be pointed out in the limitations section of Chapter 4. It is true that in the past 15 years the study of motivation has veered towards dynamic and more qualitative analysis of this construct (see Ortega 2009 for a review). However, given the number of participants and the multiple constraints that we encountered when carrying out this

research project, it was considered that a quantitative questionnaire of this kind was the most effective way to obtain a measurement of students' level of motivation.

3.7.2 Measuring aptitude

In this study, aptitude will be operationalised as a measurement of students' working memory capacity and ability for oral mimicry. Working memory capacity will be assessed through a battery of tasks recently designed by Lewandowsky *et al.* (2010), while an adaptation of Lord's (2006) mimicry task will be used to examine participants' ability for oral mimicry. A full description of each instrument is presented in the following sections.

3.7.2.2 Assessing working memory capacity

As an instrument for measuring students' working memory, we chose a battery of tests designed by Lewandowsky *et al*. (2010) for MATLAB. Tests included *a sentence-span task, an operation-span task, a spatial short-term memory task and a memory-updating task*. As the authors point out (2010: 571):

> These tasks were chosen in order to provide a heterogeneous set of measures of working memory capacity, thus reducing method variance and tapping into two content domains of working memory (verbal, including numerical, vs. spatial) and two of its functional aspects (storage in the context of processing and relational integration.

Moreover, these tasks offer a measurement of both students' memory capacity and processing ability. This battery of tests is freely available for download on www.cogsciwa.com. It is noteworthy that, in order for the programme to work, a version of MATLAB must be previously installed on each computer, along with the Psychophysics Toolbox. A full description of each task plus detailed scoring procedures can be found in Lewandowsky *et al.* (2010).

The *memory-updating task* presents participants with a series of digits that students attempt to remember while performing a series of arithmetic operations. A total of 15 series of digits are presented. Sequences are generated at random.

The *operation-span task* presents participants with a sequence of arithmetic equations (e.g. 10 - 4 = 6) and a number of consonants for subsequent recollection. First, candidates need to indicate if the equation is correct or incorrect. After each equation, a consonant appears on the computer screen for 1 sec. The goal is not only to judge if the equation is right or wrong, but also to recall all consonants presented at the end of the series. All sequences are generated at random.

The *sentence-span task* is similar to the previously described operation-span task. In this case, the processing element of the test centres on determining whether a series of phrases are meaningful or meaningless (see Lewandowsky *et al*. 2010 for a full description of the type of sentences included as stimuli).

In the *spatial short-term memory task*, participants are presented with a series of dots in a matrix on the computer screen. Dots appear one by one within different cells of the matrix for 900 milliseconds. Participants need to recall their patterns of presentation, i.e. the absolute position of each dot is irrelevant and what becomes important is the recollection of their overall patterns of display.

With regard to the scoring procedures, the indications provided by Lewandowsky *et al*. (2010) were followed at all times. Nevertheless, it is worth pointing out that all results were transformed into percentages in an attempt to facilitate comparisons between tasks and groups of students (see Chapter 4 for a presentation of results).

3.7.2.3 Assessing oral mimicry

We followed the procedures set by Lord (2006) in the design of an oral mimicry task. The main researcher in this study (a native Spanish speaker) read out loud ten sentences (Appendix E). Each sentence contained one invented word that was phonotactically possible in Spanish. Words acted as nouns or adjectives in the carrier sentences and ranged from 2 to 4 syllables in length. Students were recorded repeating each sentence immediately after the researcher. Participants' recordings were subsequently transcribed and a score was calculated based on the number of correctly produced invented words (out of a maximum of 10).

3.7.3 Measuring level of proficiency

The DELE (*Diploma de Español como Lengua Extranjera*) Exam *Nivel Inicial* released in 2009 [44] was used as a reliable instrument to measure the level of proficiency of our students. DELE tests are designed by The *Instituto Cervantes,* a Spanish cultural institution aimed at promoting the teaching of Spanish language and dissemination of the Spanish and Latin America culture throughout the world. A passing grade on the DELE exam *Nivel Inicial* proves learners' ability to address the most common communicative situations of everyday life. It certifies a basic knowledge of Spanish language equivalent to a level B1, in accordance with the description of the Council of Europe. A B1 level of proficiency is, in turn, similar to the level reached by a secondary school student in England with a GCSE in Spanish. The following procedures have been established by the *Instituto Cervantes* to ensure the reliability of the DELE exams:

- The selection and preparation of materials is carried out by the University of Salamanca.
- Piloting is performed to test the materials for each of the four skills with international students enrolled at the University of Salamanca.
- An analysis of results is performed in order to draw conclusions on the different test items.
- The selection of items is based on degree of difficulty and level of discrimination.
- Items are stored for future construction of tests.
- The final design is approved by the *Instituto Cervantes*.

Table 5 below summarises the exam structure and the scoring criteria for each section.

[44] A copy of the exam is available on the following website:
http://diplomas.cervantes.es/informacion/niveles/nivel_b1.html

Table 5: DELE exam structure and scoring criteria

TEST 1: Interpretación de textos escritos	**TEST 2: Producción de textos escritos**	**TEST 3: Interpretación de textos orales**
Maximum score: 20 points. Total items: 20. Value of each item: 1 point. Score calculation: (NA * NM) / NI (Number of correct items multiplied by the maximum possible score and divided by the number of items).	Overall rating for each part (maximum score): 4 points. Maximum score in total: 15 (NA* NM)/ NI (Number of correct items multiplied by the maximum possible score and divided by the number of items).	Maximum score: 15 points Total items: 22 Score calculation: (NA * NM) / NI (Number of correct items multiplied by the maximum possible score and divided by the number of items).
TEST 4: Conciencia comunicativa	**TEST 5: Expresión e interacción oral**	**TOTAL**
Maximum score: 20 points Total items: 30 Score calculation: (NA * NM) / NI (Number of correct items multiplied by the maximum possible score and divided by the number of items)	1. Initial questions: 4 points 2. Communicative interaction: 4 points 3. Dialogue based on visual stimulus: - Description: 4 points - Narrative: 4 points - Tastes and preferences: 4 points Total Points: 20. Score calculation: Multiplication of the candidate's result by the maximum possible score and divide the result by the total of points available)	PART 1: - Test 1 (Interpretation of written texts): 20 points - Test 2 (Production of written texts): 15 points - Total: 35 points PART 2: - Test 3 (Communicative Awareness): 20 points - Total: 20 points PART 3: - Test 4 (Interpretation of oral texts): 15 points - Test 5 (Expression and oral interaction): 30 points - Total: 45 points Overall total: 100

The DELE exam includes an assessment of listening comprehension, reading comprehension, speaking and writing ability. This assessment, therefore, covers the

four traditional skills, as well as a specific section that focuses on purely linguistic form and lexis. A full description of the DELE exams can be found on the *Instituto Cervantes*' website. The 2009 version of the DELE exam *Nivel Inicial* was used in this study, as well as the scoring procedures set by *Guía Diploma Nivel Inicial*. Both documents are freely available on the same website.

3.7.4 Assessing intelligibility

As pointed out in Chapter 2, researchers in the field of speech disorders have used a variety of materials to assess speech intelligibility. Materials have ranged from lists of words in isolation to the use of sentences with or without context (Jeng *et al.* 2006; Samar and Metz 1988; Weismer and Martin 1992;Whitehill and Chau 2004, among others), picture description tasks (e.g. Samar and Metz 1988) or even samples from spontaneous speech (e.g. Flipsen 2006). With regard to the analysis of L2 speech, our review of the existing literature has also underlined the use of a wide variety of methods and elicitation tasks when assessing non-native speech intelligibility (see Chapter 2). The goal of this study is to present a thorough analysis of our group of L2 learners' speech intelligibility and, therefore, it was decided to assess our group of participants at word, sentence and passage levels, as well as in a semi-spontaneous production task. It also noteworthy that there are certain potential drawbacks associated with speech samples obtained in spontaneous conversation versus those obtained as a result of highly controlled tasks[45]. The use of samples taken from participants' spontaneous productions would seem *a priori* more ecologically valid, since one can assume that they would possess a higher degree of naturalness. On the

[45] Systematic variation in learners' interlanguage has been observed depending on the degree of attention to language form. Tarone (1979) points out that this systematic variation ranges from vernacular to careful speech styles. The vernacular speech style exhibits the least degree of variation and the least amount of attention to linguistic form. On the other hand, careful styles present a higher degree of systematic variation. Researchers must therefore be aware of this type of contextual variation during the process of data collection. It is also necessary to point out that experimental support for Tarone's style shifting hypothesis has not always been consistent. In Beebe (1987: 386), learners' interlanguage 'becomes more permeable to a superordinate rule system in formal situations'. However, 'either the NL or TL may act as the superordinate rule system'. Oyama (1976) and Sato (1984) provide counterevidence to Tarone's claims. Ellis (1985), on the other hand, sees variability as a driving force in the learning process, since it seems to be responsible for the introduction of new linguistic forms in the learner's interlanguage. In any case, it would seem that style shifting is only one of the multiple factors that account for the high degree of variability in learners' interlanguage (see Larsen-Freeman and Long 1991 for discussion).

other hand, it has proven difficult to obtain comparable spontaneous utterances in order to analyse the role of certain contributing variables. In fact, several studies that have attempted the assessment of non-native speech intelligibility in spontaneous conversation can be criticised on these grounds (see e.g. Derwing and Munro 1997). In this study, we opted for a semi-spontaneous production task, as the best way to ensure the necessary control over certain lexical items, as well as the possibility of inter-participant comparison of speech samples (see Section 3.7.4.4).

After these initial considerations and taking into account the review of the literature undertaken in Chapter 2, a single word intelligibility test was adopted to assess speech intelligibility at word level. A transcription task involving speech samples elicited through the reading of the Spanish version of the Harvard Psychoacoustic Sentences and a phonetically-balanced text were used to assess intelligibility at the sentence and passage level. Furthermore, a tailor-made elicitation task served to elicit speech samples in semi-spontaneous production. Each instrument is fully described in the following sections.

3.7.4.1 Single word intelligibility test

The creation of a single word intelligibility test involves a phonemic error analysis for the language combination under study and a subsequent selection of stimuli probing those potential phonemic errors. This test usually adopts the format of a multiple-choice identification task (see Kent *et al.* 1989; Rogers 1997 or Whitehill 1997 for a thorough description of this type of procedure). When the language combination in question has been scarcely studied, an error analysis of participants' productions becomes necessary. In our case, the Spanish and English combination has been widely treated in the literature and, thus, the extrapolation of patterns of errors at the phonemic level for English learners of Spanish is fairly straightforward. It is important to point out that the single word intelligibility test will attempt to explore if those problematic areas do actually translate into intelligibility loss.

3.7.4.1.1 Some notes on the Spanish and English phonological systems: potential difficulties for English learners of Spanish

This section will not present a contrastive analysis between the phonological systems of English and Spanish (there is plenty of literature available offering a detailed treatment of this issue[46]). We are, however, going to examine some of the potential difficulties encountered by English speakers when acquiring the Spanish phonological system. These potential difficulties will be the basis for the design and selection of stimuli in the single word intelligibility test. It is also noteworthy that a standard variety of Peninsular Spanish will be adopted in the following analysis, given that all our evaluators use this variety. Moreover, Peninsular Spanish was the model used in the foreign language classes followed by the participants in this study. We will refer to British RP (Received Pronunciation), the 'best described and most studied' (Roca and Johnson 1999: 172) of all English varieties, when analysing the phonological system of the English language.

As far as vowels are concerned, Spanish and English present systems that are subject to very different dynamics. English has a broader range of vowel quality and quantity. The Spanish vowel system is a five-vowel system, while the English system (RP pronunciation) presents up to twenty or twenty-one phonemes (O'Connor 1973). Variation within dialects or accents is far greater in English than in Spanish. Differences between both systems go beyond vowel quality and quantity and affect both segmental and suprasegmental features. Odisho (1992) rightly establishes a direct link between the nature of a vowel system (centripetal versus centrifugal) and the different rhythms of the two languages (syllable-timed versus stress-timed[47]). In English, there is a clear connection between vowel quality, vowel quantity and stress. In syllables with a secondary stress, both quantity and quality are reduced. This reduction, also known as shwaization, is one of the most characteristic features of the English vowel system. In fact, in unstressed syllables, almost all vowels can be reduced to [ə]. This phenomenon of vowel reduction is typical of a centripetal system

[46] See, for example, Bowen (1956); Cárdenas (1960); Graham (1978); Puigvert Ocal (2001) or Stockwell and Bowen (1965).

[47] The initial distinction between syllable-timed and stress-timed languages was established by Pike (1945). While stress-timed languages possess a rhythm that is based on the reoccurrence of stressed syllables at regular intervals, in syllable-timed languages rhythm is based on the notion that each syllable has the same duration. Spanish has traditionally been considered as a syllable-timed language, even though experimental data has not always supported this view (see Section 3.8.4.1 for further discussion).

in which schwaization can pull almost all vowels to the centre of the system, while reducing their tenseness and length to a minimum. Regarding this issue, Navarro Tomás (1932: 46) comments the following:

> Conviene, finalmente, advertir que las vocales españolas no llegan en ningún caso en la pronunciación normal al grado de imprecisión y vaguedad que se manifiesta, por ejemplo, en las vocales relajadas inglesas. La relajación de las vocales a la manera inglesa es uno de los principales escollos que los estudiantes ingleses y angloamericanos necesitan evitar al aprender el español.

Unlike in English, Spanish vowels do not suffer reduction in unstressed syllables and never undergo schwaization. Based on the previous analysis, it must be noted that one of the most difficult problems for English learners of Spanish concerns the accurate production of unstressed vowels (Stockwell and Bowen 1965). Moreover, English speakers should try to avoid their natural tendency of pronouncing the stressed Spanish vowels /e/ and /o/ as diphthongs. This tendency should be avoided by placing a strong articulatory tension while pronouncing the vowel sound. Other general articulatory differences between both systems have been pointed out by Quilis and Fernández (1996: 60):

> a. El comienzo de la vocal inglesa se realiza bruscamente a causa de una entrada en vibración muy rápida de las cuerdas vocales. Esto hace que se perciba un pequeño ruido glotal al principio de una emisión vocálica, conocida por el golpe de glotis. En la vocal española las cuerdas vocales se aproximan lentamente, y entran despacio en vibración.
> b. El final de la vocal inglesa no es tan rápido tampoco como el de la española: en aquella, la glotis va dejando de vibrar paulatinamente. El final de la vocal española es rápido, cortante, seco.

In spite of these apparent differences, it would seem logical to assume that a transition from a more complex to a simpler vowel system would be an easy task. However, this is not the case and Spanish vowels are generally considered more difficult to master than consonants. This is due, as pointed out by our previous analysis, to the differences between both languages in terms of rhythm, stress and vowel reduction. In addition, vowel combinations can potentially be a source of problems for English learners of Spanish. Dalbor (1980: 259) notes the tendency of English speakers to make 'a syllabic division between vowels which should fuse into one'. Even in vowel

combinations, which are present in both languages, differences in terms of the quality of the second vowel could arise.

To summarise, some of the major differences between both vowel systems, as well as potential difficulties for English speakers, stem from the following points:

1. Spanish has a centrifugal type of vowel system subject to a syllable-timed dynamic while English vowels belong to a centripetal system in accordance with a stress-timed dynamic.
2. The most difficult problem among Spanish vowels for English speakers concerns the vowels under weak stress (Stockwell and Bowen 1965).
3. English has a range of vowel quality twice as broad as that of Spanish (Odisho 1992).
4. The English system is far more complex in terms of number of syllabic nuclei.
5. Spanish presents a symmetric and very clear vowel system, while in English, the complexity and variations from accent to accent are far greater.
6. Navarro Tomás (1932) argues that the characteristic feature of Spanish pronunciation is the brevity of its vowels, whatever the degree of stress of these sounds may be and the form in which they appear
7. English is a system that tolerates a wide variety of vowels ranging from very tense to very lax and from very long to very short, while Spanish is a system of tense vowels with no tolerance for lax vowels (Odisho 1992).
8. English speakers should try to avoid their natural tendency of pronouncing the Spanish vowels /e/ and /o/ as diphthongs by adopting a strong articulatory tension while pronouncing the vowel sound.
9. English speakers should also avoid the lengthening of Spanish stressed vowels, particularly /a/, as well as the breaking of vowel combinations that should be pronounced as diphthongs or the creation of artificial divisions to mark word boundaries and prevent instances of *sinalefa* (Dalbor 1980: 259).

As far as the consonant system is concerned, both English and Spanish have nasal, oral and liquid consonants. In terms of number of phonemes, Spanish has less consonants but each language has contrasts that do not exist in the other. English speakers do not generally have any problems in distinguishing between categories

such as stops, fricatives, oral, nasal, voiced and voiceless. The majority of difficulties for English learners of Spanish may arise from the lack of recognition of certain phonemes that do not exist in English such as /r/, /χ/, /ɲ/, /ʎ/. One can assume that the English learner, when trying to articulate these phonemes, will attempt to identify them with those L1 phonemes that are more similar. In addition, there are phonemes that exist in both languages but are produced with articulatory differences. We can predict that the voiceless stops /p/, /t/ and /k/ will be produced with aspiration in those contexts in which they are rendered as aspirated in English, i.e. in stressed syllables, when they appear in word initial positions and do not follow a fricative /s/. It is also predicted that learners will produce /t/ and /d/ as alveolar and not as dental. Furthermore, difficulties could arise from the approximant pronunciation of /b/, /d/ and /g/ in certain contexts. In Spanish, the stop rendition of these phonemes only occurs after pause, nasal or liquid. Considering that English speakers do not have these allophonic variants in their L1, and also given that /b/, /d/, /g/ are rendered as stops in English in all contexts, one could predict that English learners of Spanish will exhibit difficulties in perceiving and producing the aforementioned phonemes and allophonic variants. Furthermore, in certain contextual instances (CVdV), perceptual misunderstandings could appear between /d/ and /r/ (Stockwell and Bowen 1965). Another aspect that must be taken into account is the articulation of the Spanish alveolar consonant /l/ as the English velar [ɬ], in particular when appearing in word-final position (e.g. “mal” or “sal”). Further examples of articulatory transfer from the L1 can be found in the production of /s/ and /n/ as alveolar instead of apico-alveolar, or the articulation of /θ/ as dental instead of interdental. In cases of students with lower levels of proficiency, problems could arise from the pronunciation of “ll” as /l/. Further problematic instances can be found in the English distinction between /s/ and /z/ (in Spanish /z/ is a mere allophonic variant), as well as in the association of the letter “z” with the phoneme /z/ (in Peninsular Spanish “z” is rendered as /θ/). Moreover, the distinction

between /ɾ/ and /r/ is particularly important and is usually the object of focused instruction on courses tailored to English speakers.

To summarise, as far as consonants are concerned, the majority of potential problems for English learners of Spanish will involve the lack of perception of certain phonemic distinctions that do not exist in their L1 and the assimilation of the different distribution of phonemes and allophones which exist in both languages. It is noteworthy that some of the aforementioned differences in articulation and distribution of allophonic variants only operate on a phonetic and not on a phonemic level. In other words, some of the potential problems pointed out above may not necessarily translate into intelligibility loss.

In terms of suprasegmentals, we have already referred to the differences between syllable and stress-timed languages and the consequences in terms of vowel reduction of unstressed vowels. Dalbor (1980: 259-260) mentions also the difficulties for English speakers with regard to the possible misplacement of stress in cognate words, the use of the wrong terminal juncture with yes/no questions or the wrong use of pitch levels. Further observations regarding the importance of additional suprasegmental features such as pause frequency, pause duration and speech rate will be put forward in Section 3.8.4.

3.7.4.1.2 Selection of stimuli and design of the single word intelligibility test

Based on the observations that we have made in the previous section, along with a perusal of the relevant literature regarding the hierarchy of difficulties for English learners of Spanish, contrastive analyses between both languages (e.g. Cárdenas 1960; Stockwell and Bowen 1965; Whitley 2002) and general considerations on the difficulties of the Spanish phonological system for English speakers (e.g. Dalbor 1980; Gil 2007; Hualde 2005; Quilis and Fernández 1996), a series of potential problems for English learners of Spanish on a phonemic level were selected. It must be pointed out that only those phonemic errors listed as capable of causing instances of intelligibility loss were initially considered. In other words, segments whose mispronunciation may result in only foreign-accented speech were excluded[48]. Table

[48] Authors such as Stockwell and Bowen (1965) use additional criteria in the selection and treatment of potential problematic areas: hierarchy of difficulty, functional load, potential mishearing and pattern

6 below offers a list of the eight categories of phonemic errors that were finally selected, as well as the specific minimal pair contrasts that will be probed through the single word intelligibility test. It is important to note that consonant clusters were excluded from this test. First at all, Spanish exhibits a lesser degree of complexity in terms of consonant sequences when compared to English. Furthermore, it was considered that the potential difficulties arising from the pronunciation of consonant clusters were unlikely to result in intelligibility loss, even though their impact on the degree of foreign accent could certainly be significant (see e.g. Stockwell and Bowen 1965 for a thorough review of problematic instances of initial, medial and final consonant clusters). No potential confusions between /ɟ/ and /ʎ/ were considered. Moreover, confusions between /s/ and /θ/, common in certain varieties of Spanish, were disregarded. Other instances that we did not deem significant in terms of intelligibility loss include the use of word-initial aspirated stops /p/, /t/, /k/, the use of English velar [ɬ] instead of Spanish alveolar /l/ and the non assimilation of nasals to the following consonants. A very broad criterion of selection was therefore adopted in the inclusion of potential phonemic errors, given that our main goal was to provide a thorough examination of all possible instances of intelligibility loss in our group of Key Stage 4 learners of Spanish. Table 6 offers a list of all potential phonemic confusions, along with a description of each type of error. As far as vowels are concerned, point and manner of articulation (P/M) were used to describe the potential phonemic errors. Point and manner of articulation are described by Quilis and Fernández (1996) as the only two pertinent vowel features in Spanish from a phonological standpoint. In the case of consonants, manner of articulation (M), point of articulation (P), voicing (V), nasality (N) and deletion (D) were used in the description of potential phonemic errors. Diphthong (Dip), as an additional category, was also selected to include all errors affecting vowel combinations whether they involved deletion (D) or epenthesis (E) of segments within a vowel sequence.

congruity. It is important to point out that the potential for intelligibility loss is the only criterion followed in this study.

Table 6: Categories of phonemic contrasts and minimal pairs

Categories of phonemic errors	Description
Stressed vowels	
/i/–/e/	M
/i/–/a/	M/P
/i/–/o/	M/P
/i/–/u/	M
/e/–/a/	M/P
/e/–/o/	P
/e/–/u/	M/P
/a/–/o/	M/P
/a/–/u/	M/P
/o/–/u/	M
Unstressed vowels	
/i/–/e/	M
/i/–/a/	M/P
/i/–/o/	M/P
/i/–/u/	P
/e/–/a/	M/P
/e/–/o/	P
/e/–/u/	M/P
/a/–/o/	M/P
/a/–/u/	M/P
/o/–/u/	M
Vowel combinations	
/ai/–/a/	Dip/D
/ei/–/e/	Dip/D
/au/–/a/	Dip/D
/eu/–/e/	Dip/D
Stops	
/p/–/b/	V
/p/–/f/	M
/p/–/m/	N
/p/–/t/	P

/p/–/k/	P
/b/–/m/	N
/b/–/f/	V/M
/b/–/d/	P
/b/–/g/	P
/d/–/t/	V
/d/–/θ/	V
/d/–/g/	P
/t/–/k/	P
/t/–/θ/	M
/k/–/g/	V
/k/–/x/	M
/g/–/x/	V
Fricatives	
/f/–/θ/	P
/f/–/s/	P
/f/–/x/	P
/θ/–/x/	P
/s/–/x/	P
/s/–/tʃ/	M
/ʝ/–/tʃ/	M
/ʝ/–/ ɲ/	N
Affricates	
/tʃ/–/p/	M/P
/tʃ/–/t/	M/P
/tʃ/–/k/	M/P
/tʃ/–/ ɲ/	N
Nasals	
/m/–/n/	P
/m/–/ɲ/	P
/n/–/ɲ/	P
Liquids	
/ʝ/–/l/	M/P
/ʝ/–/r/	M/P

/ɟ/–/ɾ/	M/P
/ɟ/–/ɲ/	M
/l/–/ɾ/	M
/l/–/r/	M
/r/–/ɾ/	M
/r/–/d/	M/P
/r/–Ø	D

Sixty-five potential contrasts at a phonemic level were finally identified: 20 contrasts involving vowels, 4 affecting vowel combinations and 41 involving consonants. Based on these 65 potential phonemic contrasts, a single word minimal pair test was designed in an attempt to provide an assessment of our participants' speech intelligibility at word level. Given that variability is one of the intrinsic characteristics of non-native speech, it was decided to include 3 tokens for each contrast. The complete list of items can be found in Appendix F. Following Rogers (1997), each item contains a target word and a foil, which differ only by one phoneme. Further elements of description, such as whether the contrast involves a vowel, consonant or vowel combination or the type of potential confusion reflected in each pair (manner, point of articulation, voicing and deletion), are also provided both in Appendix F and in the sample of minimal pairs listed below (Table 7).

Table 7: Sample of minimal pairs

Target	**Foil**	**V/C**[1]	**Minimal pair**	**Type**[2]
Carro	Caro	C	/r/–/ɾ/	M
Toda	Toga	C	/d/–/g/	P
Tela	Tila	V stressed	/e/–/i/	M
Sala	Sola	V stressed	/a/–/o/	M, P
Pisado	Pesado	V unstressed	/i/–/e/	M

1 Vowel (V) or consonant (C) error

2 Manner of articulation (M), point of articulation (P), voicing (V), deletion (D), diphthong (Di)

After the recording of the target words by the 20 participants in this study and the listening sessions in which evaluators indicated which of the two pairs (target or foil) they had actually perceived, this test yielded information on those categories of

phonemic contrasts that resulted in instances of intelligibility loss at word level (see Chapter 4 for a presentation of results).

3.7.4.2 Assessing intelligibility at sentence level

The Harvard Psychoacoustic Sentences (Egan 1948) are widely used to assess the intelligibility of words in sentence context. They consist of a group of phonetically-balanced sentences with varied and correct syntactic structures. Evaluators do not choose between several options but simply transcribe or repeat the sentence they have heard. It is easy to administer and score. However, it does present some important limitations, such as a restricted number of items, strong learning effect or high percentage of correct answers. This is due to the fact that the evaluators are able to use not only phonetic but also semantic information during the transcription task.

The Spanish version of this test (Valero 1991) consists of 10 groups of 10 sentences each (Appendix G). Only the first 5 series will be used to assess intelligibility at this level of analysis. Sayings, colloquial and idiomatic expressions, as well as interrogative or exclamation sentences have been excluded from the corpus. Sentences are syntactically well formed and include simple and complex sentences, both coordinated and subordinated, as well as all verb tenses. Sentences are phonetically balanced in groups of 10 but not necessarily when considering the 100 sentences as a whole. With regard to phoneme frequency of occurrence, data from Navarro Tomás (1946) and Alarcos (1965) has been used when compiling the list of stimuli. It must also be pointed out that Navarro Tomás does not offer any details on frequency of occurrence of allophones and, thus, the balance of the corpus must be considered in a phonological and not in a phonetic sense.

3.7.4.3 Assessing intelligibility at passage level

The reading of a phonetically-balanced text served to elicit speech samples and measure the speech intelligibility at passage level. After a perusal of the available literature, we only found two phonetically-balanced texts in Spanish: the first one was prepared by Ortega, González and Marrero (2000) as part of the AHUMADA

language corpus. It contains 180 words, 305 syllables and 712 phonemes. The text has a statistical correlation coefficient of 0.99 with regard to the frequency of phonological occurrence in Castilian Spanish. The second passage was initially selected by Bruyninckx, Harmegnies, Llisterri and Poch (1994). It contains 209 vowels and 231 consonants. The text follows the phonetic distribution set by Navarro Tomás (1946). From a thematic point of view, the first passage is a literary text that uses the first person, while the second is a journalistic passage that is likely to have been taken from the events section of a newspaper. After reviewing the vocabulary used in both texts, we selected the first passage (Ortega, González and Marrero 2000) as a tool to elicit the necessary speech samples at this level of analysis. We decided to select the text that, from a lexical point of view, was better suited to the level of proficiency of our group of participants in this study.

3.7.4.4 Assessing intelligibility through a semi-spontaneous production task

Controlled tasks such as the ones used in this study at the word, sentence and passage levels allow for a total control over the lexical items included in the speech samples. However, they lack a degree of naturalness. A speech sample collected from students' spontaneous conversations would *a priori* seem to be more suitable if our goal is to reach a certain degree of naturalness. Nevertheless, two facts make the selection of this type of speech sample an extremely difficult task: first, according to the definition of intelligibility adopted in this study, it becomes necessary to establish a comparison between our group of participants' intended message and the message actually perceived by the group of evaluators. Furthermore, similar speech samples must be collected from each participant if we intend to establish any inter-participant comparisons. In other words, it would seem that control over the lexical items included in the different speech samples is essential if we are to meet the aforementioned requirements. It was decided to create a task by using five series of the Spanish version of the Harvard Psychoacoustic Sentences (series 6 to 10). Students were presented with the different sentences in which both the subject and the verb were underlined (Appendix I). They were asked to transform a singular subject into plural and operate the corresponding subject-verb agreement and vice versa (transformation from plural into singular depending on the specific sentence), as well

as to pronounce out loud the resulting sentence. When the subject was omitted from the sentence, only a verb transformation was required. Moreover, when the subject was a proper noun, students were told to add a second proper noun to the subject of the sentence, therefore, making it plural, and then perform the necessary subject-verb agreement. Participants were not allowed to write anything down while they were performing the task. It was expected that this grammatical task would divert students' attention from focusing exclusively on word pronunciation while ensuring, at the same, a total control over the lexical items.

3.8 Procedures

3.8.1 Recording sessions

All recording sessions of students' oral productions were made on a pc/laptop equipped with an M-Audio MobilePre USB interface. Each session ran individually in a quiet room in the school premises. The microphone was placed 30 cm from each participant's mouth. Subjects were instructed not to speak directly into the microphone. We used ProRec 1.2 (developed by Mark Huckavale University College London and freely available on the UCL website) as audio recording software. Stimuli were presented to students on the computer screen and were recorded directly to disk from the microphone using a 48 kHz sampling rate. Files were saved in WAV format. A translation of stimuli into English was also provided to students in paper format. Given the number of tests and stimuli presented, we decided to run two different recording sessions: one to record the stimuli of the single word intelligibility test along with the Harvard Psychoacoustic Sentences, and a second session allocated to the recording of the reading passage and the semi-spontaneous production task. Recording sessions were carried out as follows:

1. Preliminary explanation of the task
2. Questions and answers
3. Testing of the recording device
4. Recording of practice stimuli
5. Clarification of possible questions

6. Recording of real stimuli on the computer

Stimuli for the single word intelligibility test were recorded in 9 series of 20 words each, plus one series of 5 words. Stimuli for the Harvard Psychoacoustic were presented respectively in 5 series of 10 sentences. Procedures were slightly different for the reading of the phonetically-balanced passage and the semi-spontaneous production task. Enough time was given to participants to read the selected passage in its entirety and to consult the English translation if necessary. In the semi-spontaneous production task, participants had up to 15 seconds to operate the grammatical transformation and utter the resulting sentence.

3.8.2 Listening sessions

Each participant in this study was randomly assigned to a panel of three evaluators (a total of 15 panels). An additional participant who had recorded the single word intelligibility test and the reading of the Harvard Psychoacoustic Sentences served as a control subject. This subject had not completed the recording of the reading passage and the semi-spontaneous production task and had therefore been initially excluded from this study. It was decided subsequently that all evaluators should listen to a series of stimuli from at least a single subject. This was essential in order to assess possible listener-related differences between each panel of evaluators. This assessment was made possible thanks to use of the stimuli previously recorded by the aforementioned control subject. Furthermore, two of the panels had to listen to the stimuli recorded by the two native Spanish speakers described in Section 3.6. The presentation of stimuli was randomised to avoid any possible learning effects.

The evaluation sessions ran at the *Laboratorio de Idiomas-Aula Multimedia (LIAM), Universidad de Cantabria, Santander*. Stimuli were presented through headphones and all transcription tasks were completed on the computer using the software interface provided by E-Prime 2.0. Each evaluation session was carried out as follows:

1. Preliminary reading of the instructions on how to perform the task
2. The researcher orally explains the task and answers any possible questions

3. Clarification of any possible questions
4. Real stimuli are finally presented

The presentation of stimuli consisted of 195 minimal pairs for the single word intelligibility test (evaluators chose between a foil and a target stimulus), 50 sentences from the Spanish version of the Harvard Psychoacoustic Sentences (evaluators had to orthographically transcribe each sentence), a phonetically-balanced passage and 50 sentences from the semi-spontaneous production task. A ten-minute break was scheduled between each of the four tasks. Kendall's Coefficient of Concordance was used to examine inter-evaluators reliability with regards to the assessment carried out on the control subject. Kendall's coefficient was estimated at W= 0.955 which shows a high degree of agreement.

3.8.3 Assessing intelligibility: scoring procedures

The scoring procedure for the single word intelligibility test consisted in counting the number of correctly perceived target words out of the 195 possible stimuli for each participant. Scores were converted into percentages.

As far as the assessment of the Harvard Psychoacoustic Sentences is concerned, two different scoring criteria could be adopted *a priori*: (i) a strict criterion that would penalise those errors that can be explained through contextual information and (ii) a more lenient criterion that would disregard this type of error. For example, if one of the evaluators transcribes “los libro están sobre la mesa” and “libros” is a key word, adopting a strict scoring criterion would consider this instance as an error. Additionally, there is also the option of taking into consideration only those errors committed on key words or those committed on any word. Therefore, the following alternatives were initially examined:

1. To consider only errors in the production of key words without taking into consideration errors that could be explained from the context (as explained above).
2. To consider errors in key words and to penalise any mistake.
3. To consider errors in any word without taking into consideration the mistakes

that could be explained from the context.

4. To consider those errors in any word and to penalise all mistakes.

After examining the test results, we decided to follow the first of the aforementioned criteria. Due to the nature of the stimuli involved (meaningful sentences with correct syntactic structures), it is unlikely that evaluators would write sentences without respecting basic grammatical rules. Evaluators, when transcribing the answers, look for meaningful grammatical patterns and, thus, a strict corrective criterion would offer no substantial variations on intelligibility scores. Results will therefore be presented by determining the number of key words that have been correctly transcribed for each sentence. Key words are those with a clear lexical content. Modal or auxiliary verbs, adverbs, pronouns, prepositions, conjunctions, articles, determiners have been excluded.

The scoring procedure described above was also followed in the scoring of the transcription of the reading passage and the semi-spontaneous production task. All scores were transformed into percentages to facilitate inter-tasks comparisons.

3.8.4 Assessment of suprasegmental features

As pointed out in Chapter 2, many authors (Anderson-Hsieh and Koehler 1992; Fries 1945; Hahn 2004) argue that the prosodic domain is more important in ensuring speech intelligibility than a precise articulation of sounds. In other words, prosodic deficiencies may hamper the understanding of L2 utterances to a greater extent than segmental deficiencies, and may cause more serious misunderstandings than those due to segmental errors (Celce-Murcia *et al*. 1996). A good command of L2 prosody facilitates the overall comprehension by the interlocutor, even when other types of errors (lexical, stylistic, grammatical, etc.) are being committed. Prosody is considered to be difficult to assimilate and difficult to teach (Laroy 1995). In the context of formal instruction, where the importance of prosody in the assessment process is, at best, very small, it is perfectly understandable that learners themselves, whether consciously or unconsciously, focus on other linguistic aspects and disregard accent and intonation patterns. Additionally, research on non-native prosody has been characterised by the scarce number of studies, the variety of methods used, and the

difficulty in filling the gap between theoretical research and its application in the foreign language classroom.

A series of 10 sentences from the Harvard Psychoacoustic Sentences, the phonetically-balanced text, as well as 10 sentences from the semi-spontaneous production task served as stimuli for this part of the study. Acoustic analyses, carried out using Praat (version 5.1.32), were performed on a series of suprasegmental features that, according to the literature consulted, have been known to contribute to intelligibility or comprehensibility issues for L2 learners. The selected suprasegmental elements were: speech rate, pause frequency and pause duration. It is noteworthy that we used the Praat script developed by De Jong and Wempe (2009) in the calculation of speech rate, pause frequency and pause duration.

3.8.4.1 Speech rate

Speech rate is usually measured by the number of syllables or words that are pronounced in a certain unit of time. From a terminological point of view, we must establish here a distinction between *speech rate* and *articulation rate*. Researchers often use the term *speech rate* to include a measure of pauses. *Articulation rate*, on the other hand, provides a measure of the number of words or syllables per unit of time excluding pauses[49]. In this study, we will offer a measurement of articulation rate. However, following Kendall (2009: 130), we will use *speech rate* to refer to *articulation rate*, feeling that 'the term is clearer and more obvious to readers'. Another notion that is frequently used in connection with speech rate is 'tempo', which alludes to the perceptual dimension of speech rate.

Different degrees of speech rate can be associated with different languages or even with individual cognitive characteristics. For example, Grosjean (1980) points out that the average speech rate in German is higher than in French. Gili Gaya (1950) mentions the tendency of the Spanish language towards an *andante* tempo, which corresponds to a range between 90 and 126 metronomic oscillations per minute[50]. It is

[49] This initial conceptualisation of *speech rate* as equivalent to *articulation rate* will have to be altered in order to avoid a multicollinearity problem during the statistical analysis (see Section 4.7). The term *speech rate* will be used during the multiple-regression analyses to refer to a measure that includes the number of syllables per unit of time including pauses.

[50] Roach (1998: 154), after reviewing some of the empirical research carried out in the field, concluded that: 'useful though the above findings are, they do not yet bring an answer to the question of whether

noteworthy that, in general, speakers adapt their speech rate to the perceptual needs of their listeners. For instance, when an individual is in a noisy environment, he/she usually speaks louder and slower. Thus, speech rate may be modified due to reasons related to the communicative message or to external reasons, such as the circumstances in which the communicative exchange takes place.

In the field of psycholinguistics, speech rate has generally received less attention than other temporal aspects of speech such as pause frequency or pause duration. Goldman-Eisler (1961) noted a very small degree of variation in articulation rate at the speaker-level. In her study from 1968, this author pointed out a high degree of inter-speaker variation with regard to the same dimension (see Deese 1984 for opposite results). From a sociolinguistic perspective, authors such as Robb, Maclagan and Chen (2004) have noted differences in speech rate for different varieties of English. Furthermore, researchers seem to agree on the fact that males speak faster than females (Salmons, Jacewicz and Allen Fox 2008).

Speech rate has also been shown to have an influence on comprehensibility, intelligibility and judgements on foreign-accented speech (Derwing and Munro 1997; Derwing *et al*. 2004). It has also been used as an indicator of fluency and general language proficiency. Studies have revealed that non-native speech rate is slower than that of native speakers' (Munro and Derwing 1995). Several methods have been used to measure speech rate: number of words per minute, number of syllables per minute and number of phonemes per time unit (see Cucchiarini, Stirk and Bovis 2002; Towell 2002). In our study, following Munro (1995), the number of uttered syllables was divided by their duration excluding pauses longer than 100 ms (Kendall 2009). An average for each student was calculated and then used in the correlation analyses. A comparison with the values provided by two native speakers will also be presented.

3.8.4.2 Pause frequency and pause duration

From a psycholinguistic perspective, the study of pauses has generated a great deal of research over the past decades. Levelt's (1989) model of speech production posits that the process of putting ideas into words must proceed through three different stages:

---

some languages are spoken more rapidly than others (when situational and personal factors have been taken into account).

conceptualization, formulation and articulation. Speakers could experience problems at any given point of the aforementioned process and be compelled to interrupt the smooth flow of talk. 'Disfluency' is a general term that seems to include any phenomena capable of causing a breakdown in the smooth flow of speech. Different classifications of disfluencies have been put forward over the years. Harley (2001: 401), for example, distinguishes between unfilled pauses, due to micro or macroplanning, filled pauses and other disfluencies, which include false starts, repetition or parenthetical remarks. Others (e.g. Fox Tree 2002) establish a difference between pauses, *ums* and *uhs* and repetitions, replacements and restarts. It is also important to differentiate between articulatory pauses, which appear as a consequence of the articulatory demands of certain sounds such as, for example, the gaps between stop consonants, respiratory pauses, which are related to the process of inhalation or exhalation of air, and hesitation pauses, which are not related to physiological phenomena. Goldman-Eisler (1968) points out that hesitation pauses are responsible for a larger proportion of gaps in communication than respiratory pauses.

Studies of temporal patterns of speech date back at least 50 years. Frieda Goldman-Eisler, a pioneer in this field, noted that pauses are delays in the production of speech that could be associated with language planning tasks. Studies on pauses are based on time measurements during stimulus-response tasks. These experiments assume that the greater the delay between stimulus and response, the greater the number of cognitive operations that are necessary to produce such a response (Butterworth 1980). However, when analysing real speech samples, it is not always feasible to apply the stimulus-response procedure and thus, measurements focus on aspects such as speech rate or duration of silent pauses. The analysis, in this case, follows the same principle: pauses could be an indication of cognitive activity, i.e. they could be indicative of a time when the processing load is high. In this research area, Butterworth (1980) estimated that pauses are associated with two types of processes: (i) microplanning, which is related to the selection and retrieval of difficult words and (ii) macroplanning, which is connected to the semantic and syntactic organisation of larger portions of speech.

It is also important to note that, when analysing non-verbal aspects of speech production, Goldman-Eisler (1968) included the study of phenomena such as 'ums and uhs', as well as the use of false starts and repetitions. This classification has been maintained over time, although slight changes can be found sometimes depending on

the author. Boomer (1965), for example, distinguished between silent and filled pauses such as *uh, ah, a, um*. Garman (1990), on the other hand, distinguished between silent and breathing pauses, and also included in his taxonomy filled pauses. Unlike Goldman-Eisler and Boomer, Garman (1990) established a distinction between repetitions, false starts, fillers and reformulations. In the area of pathological speech, Illes (1989) conducted a comparative study on the temporal patterns of speech in subjects with Parkinson's disease as opposed to normal subjects. Results showed that the number of silent pauses per minute and the number of words between pauses were much higher in subjects suffering from the disease. Blanken *et al*. (1987) conducted a study of spontaneous speech in patients with Alzheimer's disease as opposed to subjects suffering from aphasia. A third group, made up of individuals with no language impairments, was added to this experiment. The authors noted that the first group of participants used strategies to delay the conversation and thus, to facilitate word-retrieval tasks. Bucks *et al*. (2000) conducted a comparative study on temporal patterns of speech in normal subjects and individuals suffering from Alzheimer's disease. To do this, they selected a sample of a thousand words and measured different temporal aspects such as speech rate and pause duration. Results showed that pause duration, as a statistical variable, was able to differentiate between the group of participants with the disease and the group of normal subjects.

It should also be noted that Harley (2001: 403), when evaluating the research undertaken on pauses, refers to the intrinsic difficulty in determining the specific function of pauses within discourse. In fact, pauses may be inserted as a way of assisting the listener during the comprehension process. Moreover, they could also be used to achieve 'interactional goals'. It is also very feasible that pauses may have more than one function or that different types of pauses may play different roles. As Harley (2001: 404) points out, 'although the early work was originally interpreted as showing that pausing reflected semantic planning, this is far from clear'. In fact, pauses may arise as a consequence of difficulties in the retrieval of difficult words or even due to the planning of syntactic or semantic components in a sentence.

Repetitions, restarts and repairs all refer to parts of speech where individuals have stopped and then resumed their speech. Words can be repeated exactly or they can be replaced. The use of terminology is not completely uniform is this area. While 'repair' (Levelt 1989) usually refers to the part of speech that is resumed, it can also allude to repetitions, restarts and other similar phenomena. According to Levelt

(1989), this type of phenomenon arises as a consequence of problems related to the conceptualisation stage, the selection of words or the articulation of utterances. It has also been argued that individuals follow rules in order for their repetitions, restarts and repairs to be well formed (see Levelt 1989 for further discussion). The issue of whether speakers prefer accuracy over fluency when dealing with self-repairs has also received some attention in the specialised literature. Seyfeddinipur, Kita and Indefrey (2008), for example, offer empirical data that shows that speakers seem to prefer fluency over accuracy when managing potential problems in their speech. The repair strategies used by a group of native German speakers were compiled and analysed by the researchers in this study. Subjects did not immediately interrupt their speech once they detected a problem, but rather they waited until they were able to produce the repair, as shown by the different statistical analyses.

When compared to the study of disfluencies in language production, the investigation of this type of phenomenon in language comprehension has been much scarcer. The impact of disfluencies on comprehension has been measured either through tasks that involve the measurement of comprehension while the speech is being heard (online task) or after it has been heard (offline task). Offline tasks have shown that pauses may facilitate the process of syntactic parsing (e.g. Price *et al*. 1991). Furthermore, filled pauses may offer information on turn-taking or turn-ending (e.g. Cook and Lalljee 1970). They can also have an impact on the impression the listener has on a speaker's personality (e.g. Brennan and Williams 1995). Online tasks, on the other hand, have revealed that filled pauses may speed up the recognition of words and sentences. In addition, certain types of restarts have been shown to slow recognition (e.g. Fox Tree 1995).

It is also noteworthy that pause duration seems to vary depending on the language under study. Campione and Veronis (2002) compared pause duration in English, French, Italian and Spanish. Pauses in Spanish were on average 100 ms longer than in the rest of the languages.

If we look at the area of Second Language Acquisition, a pause can be defined as a temporal interruption in the production of sounds. Dalton and Hardcastle (1977) distinguish between pauses that arise as a consequence of the articulation of stop consonants, pauses linked to a speaker's respiration patterns and those intentional pauses that may appear at significant locations in the speech chain. A further distinction can be drawn between filled and silent pauses. Laver (1980: 536) defines a

filled pause as 'any gap in the verbal structure of a speaking-turn filled by non-linguistic material'. Languages and even dialects differ in the way they fill this type of pause. While Spanish uses the nasal resonant [m:] or the vowel [e:], English utilises [a: m] or [ə:]. The use of pauses varies also with age (Kowal and O'Connell 1980). As the child develops, the frequency and duration of silent pauses decreases, while there is an increase in his/her use of filled pauses. These authors have also noted some gender differences, since boys seem to pause more than girls.

Pauses serve, from a linguistic point of view, to articulate the message in breath groups. A 'breath group' can be defined as a fragment of the speech chain between two pauses. In spite of a certain amount of variability, there is normally a correspondence between breath groups and syntactic structures. As languages possess different structures, the location of syntactic pauses varies from one language to another.

From an extralinguistic perspective, pauses can be linked to a speaker's individual characteristics or personality traits. From a paralinguistic point of view, the percentage of interruptions that characterises an individual's speech may be altered by affective or emotional factors.

With regard to reading comprehension, pauses can be classified in two types: those that are orthographically marked and those that are not associated with any particular punctuation mark. As far as Spanish is concerned, the analysis of orthographically marked pauses was addressed in Puigví and Fernández (1992), whereas Puigví and Fernández (1993) studied orthographically unmarked pauses. The goal of the study presented by Puigví and Fernández (1992) was to determine the relationship between orthographically marked pause duration and type of punctuation mark. The material for the analysis consisted in two speakers reading a literary text. Pauses were located and measured and the relationship between pause duration and type of punctuation mark was studied. The results of this analysis were used in speech synthesis to develop a set of rules that would estimate the average pause duration for each punctuation mark. In 1993, Puigví and Fernández focused primarily on two issues: first, pause distributions in the speech, and secondly, their duration. This study provided a list of syntactic boundaries, as well as values for their duration.

Research has shown that pause frequency and duration may affect the degree of intelligibility and 'accentedness' of L2 speech (Trofimovich and Baker 2006). In

spite of a long research tradition on temporal factors affecting different L1s, L2 studies have lacked, until recently, the required methodological rigour (see Riazantseva 2009 for a review).

In our study, pause duration is operationalised as any pause longer than 100 ms (Trofimovich and Baker 2006). As we pointed out above, we will only focus on silent pauses and we will not establish a distinction between hesitation and respiratory pauses. An average of the number of pauses per participant will be calculated to assess subjects' pause frequency. A comparison to native speakers' values will also be put forward. Additionally, an average in terms of pause duration will be estimated for each subject and then compared with the values provided by our two native Spanish speakers.

3.8.5 Correlation and multiple-regression analyses

Correlation analyses will be undertaken in the data analysis section of this dissertation to determine the impact of students' deviances at segmental and suprasegmental level on their intelligibility scores. Pearson's r will serve to determine the correlation between the following variables:

1. Subjects' intelligibility scores in the single word intelligibility test, the Harvard Psychoacoustic Sentences, the reading of the phonetically-balanced text and in the semi-spontaneous production task. In other words, we aim at exploring the correlation between intelligibility at word level and in connected speech.
2. Subjects' patterns of segmental deviations, i.e. different error categories, and intelligibility scores at word level.
3. Further exploration of the validity of the single word intelligibility test by focusing on the possible correlation between the different categories of phonemic errors at word level and intelligibility scores in connected speech.
4. Subjects' patterns of segmental deviations and intelligibility scores in connected speech (sentence, passage and semi-spontaneous production).
5. Subjects' deviations in pause frequency, pause duration, speech rate and intelligibility scores in connected speech.

6. Individual variables (gender, motivation, level of proficiency and aptitude) and intelligibility scores at different levels.

Multiple-regression analyses will serve to identify those predictor variables most closely associated with a loss of speech intelligibility. This statistical procedure will shed light on those error categories that seem to have the most impact on speech intelligibility. These are the features that would potentially require more attention in the language classroom.

3.9 Summary and final considerations on our methodological approach

This chapter has presented our aims and research questions, as well as provided a description of our group of participants (20 Key Stage 4 English learners of Spanish) and evaluators (60 native Spanish speakers). Furthermore, some considerations on the ethical challenges of carrying out research in educational contexts have been put forward. An analysis of the tasks used in assessing participants' level of motivation, aptitude, and proficiency has also been examined. A motivation questionnaire will serve to measure participants' motivation towards pronunciation, while aptitude will be assessed through a group of tasks aimed at measuring subjects' working memory capacity and their ability for oral mimicry. Students' level of proficiency will be assessed using the DELE exam *Nivel Inicial*. In addition, we have provided a description of the design and selection of stimuli for the single word intelligibility test, as well as the elicitation tasks used in the assessment of intelligibility in the sentence (Harvard Psychoacoustic Sentences), passage (reading of a phonetically-balanced text) and semi-spontaneous production tasks. Table 8 below provides a summary of the intelligibility tests used in this dissertation. This table offers information on the type of elicitation task used at each level, as well as on the type of intelligibility assessment carried out by our 60 evaluators. Information on the scoring procedures for each task is also provided. Chapter 3 has also offered a description of the procedures involved in the recording and listening sessions, an analysis of the scoring procedures, as well as of the acoustic analysis carried out in the assessment of certain suprasegmental factors (speech rate, pause frequency and pause duration).

Table 8: Summary of intelligibility tests

Elicitation task	**Content**	**Intelligibility assessment**	**Scoring procedure**
Single word intelligibility test	199 minimal pairs	Choice between minimal pairs	Percentage of correctly identified target words
Harvard Psychoacoustic Sentences (Egan 1948; Valero 1991)	Series of 50 phonetically-balanced and semantically well-constructed sentences.	Orthographic transcription	Percentage of correctly transcribed key words
Reading of a phonetically-balanced text (Ortega *et al.* 2000)	Phonetically-balanced text: 180 words, 305 syllables and 712 phonemes	Orthographic transcription	Percentage of correctly transcribed key words.
Semi-spontaneous production task	Series of 50 sentences: grammatical task	Orthographic transcription	Percentage of correctly transcribed key words

From a methodological standpoint, the assessment of intelligibility at different levels (word, sentence, passage levels and in semi-spontaneous production task) was a crucial element in the design of the experimental part of this dissertation. It was also a priority to introduce a combined analysis of the possible impact of segmental and suprasegmental deviations, so that we could provide a thorough picture of the potential level of intelligibility loss in our group of English learners of Spanish. Chapter 3 has also examined the statistical procedures that will be used in our data analysis. These procedures will serve to establish the potential relationship between intelligibility loss in isolated words and connected speech. Our next chapter will present an analysis and discussion of the results, as well as some potential limitations to this study.

Chapter 4: Data analysis and results

This chapter will present our analysis of results in three parts. First, an analysis of students' scores on motivation, aptitude and level of proficiency will be put forward. Specifically, our analysis will centre on participants' performance on the motivation questionnaire, each of the sections of the DELE exam, as well as their results on the oral mimicry task and the four tasks measuring working memory capacity. Second, an analysis of learners' results on the intelligibility tests (word, sentence, passage and semi-spontaneous production tasks) will be presented. Third, we will put forward students' performance on the eight categories of phonemic errors of the single word intelligibility test, as well as a segmental error analysis with regard to the same eight categories of phonemic errors in the sentence, passage and semi-spontaneous production tasks. Acoustic analyses involving speech rate, pause frequency and pause duration in connected speech will also be presented. Finally, correlation and multiple-regression analyses will examine the possible relation between intelligibility scores at different levels, as well as between the segmental and suprasegmental variables under study and participants' scores on the intelligibility tasks. Furthermore, correlation analyses will also serve to explore the relation between intelligibility results and certain individual differences, namely, gender, L1, aptitude, motivation and level of proficiency. Based on the aforementioned analyses, an attempt will be made to answer the research questions originally formulated at the beginning of this study. Additionally, an account of some potential limitations will also be offered.

4.1 Assessment of individual differences

As pointed out in Chapter 3, 20 Key Stage 4 students completed all the recording sessions. Table 9 below presents participants' scores on each of the components of the DELE exam. An overall score, representing an average of the scores obtained on each section of the test, is also provided. Results are expressed in percentages to facilitate comparisons between the different sections. It should be noted that the speaking section of the exam yielded the lowest scores. Even though teachers claim to apply principles derived from the communicative approach, the reality in the classroom is

very different. Our own observation of several teaching sessions confirmed the fact that the instructor focused greatly on writing and reading exercises with only occasional work on more interactive and communicative activities. According to our analysis of the descriptive statistics, participants exhibited a better performance on the vocabulary and grammar sections of the exam. In addition, scores on the writing and reading sections were very similar (the average of the two tests was 60.05 and 59.40 respectively). Descriptive statistics for the overall scores were as follows: scores ranged from 45 to 74, with a mean of 59.71 and a standard deviation of 6.87. Additionally, we analysed the results to determine if skewness and kurtosis could prevent the data from following a normal distribution. Skewness serves to measure whether the data are distributed uniformly with respect to a central point. A positive value may indicate a positively skewed distribution (i.e. scores are piled at the bottom of the scale); while a negative one may point to the existence of a negatively skewed distribution, i.e. with scores piled up in the upper end of the scale. Kurtosis, on the other hand, measures whether the data are concentrated in the central part of the distribution. Both in the case of skewness and kurtosis, a normal distribution yields a statistical value of around zero. Skewness values for each section of the DELE exam ranged between -0.74 and 0.15 for each of the sections. The skewness for the total score was 0.02. As for the kurtosis, values ranged between -0.88 and 2.05. Both skewness and kurtosis values are therefore close to 0, and thus, data seem to be normally distributed from a statistical point of view. No statistically significant differences were detected, in terms of level of proficiency, between male and female students ($t = -0.83$, $p = 0.42$). However, the bilingual group of students performed significantly better than the monolingual English speakers in the DELE exam ($t = -6.06$, $p = 0\ .00$).

Table 9: DELE scores

Student	Reading	Writing	Listening	GramVocab	Speak	Total
1	58.00	50.00	55.00	60.00	37.00	52.00
2	62.00	56.00	52.00	58.00	47.00	55.00
3	54.00	52.00	50.00	56.00	53.00	53.00
4	60.00	64.00	68.00	62.00	61.00	63.00
5	72.00	78.00	76.00	74.00	70.00	74.00
6	58.00	56.00	60.00	62.00	64.00	60.00
7	48.00	46.00	60.00	56.00	50.00	52.00
8	62.00	60.00	56.00	64.00	48.00	58.00
9	40.00	42.00	48.00	50.00	45.00	45.00
10	64.00	66.00	68.00	62.00	65.00	65.00
11	56.00	60.00	58.00	60.00	51.00	57.00
12	64.00	60.00	62.00	66.00	63.00	63.00
13	66.00	62.00	60.00	64.00	68.00	64.00
14	74.00	72.00	70.00	70.00	64.00	70.00
15	60.00	64.00	60.00	66.00	60.00	62.00
16	55.00	52.00	60.00	62.00	48.00	62.00
17	62.00	56.00	64.00	60.00	40.00	55.40
18	60.00	58.00	58.00	62.00	54.00	56.40
19	62.00	66.00	68.00	70.00	55.00	58.40
20	64.00	68.00	72.00	68.00	62.00	64.20

Table 10 below presents participants' results on our motivation questionnaire. The questionnaire consisted of 18 statements that students rated on a scale of 1 to 5 ('strongly disagree to strongly agree'). Results were calculated by averaging the score yielded by the Likert scale for each statement. They were later converted into percentages to facilitate comparisons (a higher percentage indicates a higher degree of motivation). Descriptive statistics show that scores ranged from 48 to 90 with a mean of 68.60, a standard deviation of 12.88. Skewness and kurtosis values were 0.15 and –1.13 ensuring, therefore, the normality of the distribution. No statistically significant differences in terms of motivation were recorded between males and females ($t = -0.53$, $p = 0.60$). Again, the bilingual speakers proved to be more motivated than the monolingual ($t = -7.45$, $p = 0.00$).

Table 10: Scores on motivation

Student	Motivation
1	58.00
2	60.00
3	62.00
4	74.00
5	85.00
6	55.00
7	48.00
8	64.00
9	50.00
10	78.00
11	69.00
12	80.00
13	85.00
14	88.00
15	90.00
16	60.00
17	65.00
18	55.00
19	76.00
20	70.00

Aptitude was measured through a combination of an oral mimicry test and a battery of tasks to measure students' working memory capacity (Table 11). As pointed out in Chapter 3, there were ten tokens in the oral mimicry section. Scores ranged from 40 to 90, with a mean of 66.25 and a standard deviation of 17.23. Table 11 below presents participants' scores on the battery of tests that served to measure working memory capacity. Results were quite similar for the memory-updating, sentence-span and operation-span tasks. Results also showed that scores on the spatial short-term memory task were significantly higher than in the other three subtests. Overall working memory capacity was calculated as an average of participants' scores for all four subtests. Descriptive statistics for the overall score are as follows: scores ranged from 46.75 to 87.13 with a mean of 67.09 and a standard deviation of 13.66. The values for skewness and kurtosis are -0.02 and –1.28 respectively. Table 11 also presents an overall score for language aptitude, an average of the oral mimicry task and the battery of working memory tasks. This overall value, along with the overall DELE score and the results from the motivation questionnaire will be used later on in

the correlation analyses to explore the relation between intelligibility scores and participants' individual differences.

Table 11: Scores on oral mimicry and working memory tasks

Student	OralM	MU	OS	SS	SSTM	TotalWM	TotalAp
1	40.00	38.00	40.00	48.00	62.00	47.00	43.50
2	50.00	50.00	58.00	54.00	70.00	58.00	54.00
3	60.00	42.00	56.00	50.00	74.00	55.50	57.75
4	50.00	62.00	70.00	68.00	88.00	72.00	61.00
5	80.00	78.00	72.00	65.00	90.00	76.25	78.13
6	50.00	58.00	38.00	45.00	58.00	49.75	49.88
7	70.00	39.00	48.00	44.00	56.00	46.75	58.38
8	60.00	61.00	60.00	66.00	68.00	63.75	61.88
9	40.00	40.00	44.00	52.00	59.00	48.75	44.38
10	70.00	96.00	75.00	80.00	94.00	86.25	78.13
11	60.00	60.00	58.00	64.00	76.00	64.50	62.25
12	80.00	78.00	66.00	80.00	92.00	79.00	79.50
13	90.00	76.00	80.00	96.00	96.00	87.00	88.50
14	90.00	98.00	72.00	80.00	98.00	87.12	88.56
15	50.00	80.00	70.00	74.00	89.00	78.25	84.13
16	60.00	58.00	48.00	58.00	80.00	61.00	60.50
17	45.00	60.00	62.00	60.00	72.00	63.50	54.25
18	70.00	54.00	62.00	60.00	70.00	61.50	65.75
19	80.00	80.00	78.00	68.00	90.00	79.00	79.50
20	90.00	72.00	70.00	70.00	96.00	77.00	83.50

It is also noteworthy that Pearson's r was computed to examine the relationship between aptitude, motivation and overall DELE scores. There was a positive correlation between DELE scores and motivation, $r = 0.81$, $n = 15$, $p = 0.00$, as well as between DELE scores and aptitude, $r = 0.\ 81$, $n = 15$, $p = 0.00$. In other words, higher aptitude and motivation resulted in a higher level of proficiency.

4.2 Assessment of speech intelligibility

This section will present the intelligibility scores at word, sentence and passage levels, as well as in the semi-spontaneous production task. Intelligibility scores at word level were calculated by estimating the percentage of correctly identified words by our 60 listeners across participants. A single word intelligibility test probing eight categories of phonemic errors served to measure intelligibility at word level. As pointed out in

Chapter 3, the eight phonemic categories under study, based on the nature of the targeted phoneme probed by the different minimal pairs, were the following:

1. Affricates
2. Unstressed vowels
3. Stressed vowels
4. Fricatives
5. Liquids
6. Stops
7. Vowel combinations
8. Nasals

Table 12 below offers the intelligibility scores at sentence level (percentage of correctly identified content words from 50 sentences of the Spanish version of the Harvard Psychoacoustic Sentences), passage level (percentage of correctly identified content words from a phonetically-balanced text) and in the semi-spontaneous production task (percentage of correctly identified content words). Kendall's Coefficient of Concordance was used to examine inter-evaluators' reliability when assessing participants' degree of intelligibility. Kendall's coefficient was estimated at $W = 0.965$, which indicates a high degree of agreement.

Table 12: Intelligibility scores in the word, sentence, passage and semi-spontaneous tasks (average percentage per speaker)

Student	**Word**	**Sentence**	**Passage**	**Spont**	**Total**
1	57.35	64.70	61.85	82.00	64.47
2	75.49	76.47	76.28	74.00	75.56
3	78.23	82.50	79.38	88.00	82.03
4	78.15	88.23	82.46	86.00	83.71
5	91.31	95.29	92.78	90.00	92.35
6	89.26	88.45	90.72	88.00	89.11
7	61.06	85.88	72.16	84.00	75.77
8	68.02	85.29	83.50	84.00	80.20
9	61.01	71.76	74.22	68.00	68.75
10	82.04	95.88	97.93	92.00	91.96
11	78.92	87.64	89.69	88.00	86.06
12	86.24	96.29	98.96	94.00	93.87
13	83.74	82.35	82.47	80.00	82.14
14	81.01	84.70	87.62	90.00	85.83
15	82.54	87.05	92.35	82.00	85.98
16	84.59	85.29	85.53	92.00	87.60
17	88.08	82.35	85.56	86.00	85.50
18	84.85	90.00	91.75	94.00	90.15
19	90.25	88.23	91.78	90.00	90.06
20	86.73	94.11	97.93	88.00	91.69
NS1	98.33	99.41	100	100	99.43
NS2	100	98.82	100	100	99.70

Intelligibility scores in the single word test ranged from 57.35% to 91.31% (percentage of correctly identified words across listeners). The mean intelligibility score was 79.44% with a standard deviation of 10.09. Scores at the sentence level ranged from 64.70% to 96.29%, with a standard deviation of 7.87 and a mean of 85.62%, while scores at the passage level ranged from 61.85% to 98.96% with a standard deviation of 9.53 and a mean of 85.69%. Scores in the semi-spontaneous production task ranged from 68% to 94% with a mean of 86% and a standard deviation of 6.52. Correlation analyses between intelligibility scores and individual differences, i.e. motivation, aptitude and level of proficiency, will be presented in Section 4.3. It is important to note that an examination of possible statistically significant differences in intelligibility results according to gender and L1 was also undertaken (Appendix J: Figures J.1 and J.2). No statistically significant differences were found according to gender at word ($t(18) = -0.25$, $p = 0.80$), sentence ($t(18) = -0.66$, $p = 0.51$) and passage levels ($t(18) = 0.21$, $p = 0.83$), as well as in semi-spontaneous production ($t(18) = -0.68$, $p = 0.51$). However, in terms of L1, bilingual students (Polish and Punjabi) performed significantly better at word ($t(18) = -2.32$, $p = 0.02$), sentence ($t(18) = -2.75$, $p = 0.01$) and passage levels ($t(18) = -2.96$, $p = 0.008$). No significant differences were found between both groups in the semi-spontaneous production task ($t(18) = -1.26$, $p = 0.22$). Overall differences between L2 learners and Spanish native speakers are also presented in Appendix J (Figure J.3). Native Spanish speakers exhibited a mean of 99.16% at word level, 99.11% at sentence level, 100% at passage level and 100% in the semi-spontaneous production task. Differences between both groups were especially noticeable at word level (99.16% versus 79.57%) and in semi-spontaneous production (78.11% versus 100%). Differences were statistically significant at all levels of analysis ($p = 0.01$, $p = 0.02$, $p = 0.04$ and $p = 0.01$ respectively).

4.3 Error analysis: single word intelligibility test

Table 13 below presents participants' scores on each of the categories of phonemic contrasts. In terms of proportion of errors, the following hierarchy of difficulty can be established (from higher to lower level of difficulty): unstressed vowels (30% of

errors), liquids (27.66%), stressed vowels (19.88%), vowel combinations (18.82%), nasals (14.82%), stops (14.64%) and affricates (14.15%).

Table 13: Scores in each of the phonemic categories under study in the single word test (percentage of correctly identified contrasts per speaker)

Student	C1	C2	C3	C4	C5	C6	C7	C8
1	61.11	40.00	50.00	87.65	50.00	49.67	61.11	59.25
2	88.88	50.00	77.77	79.01	60.49	71.89	83.33	92.59
3	83.33	60.00	70.00	100.00	67.90	81.69	77.77	85.18
4	77.77	60.00	70.00	91.35	79.01	91.50	77.77	77.77
5	100.00	80.00	90.00	100.00	79.01	100.00	88.88	92.59
6	88.88	80.00	100.00	91.35	79.01	91.50	83.33	100.00
7	77.77	50.00	45.00	91.35	43.20	71.89	50.00	59.25
8	77.77	50.00	60.00	67.90	60.49	81.69	72.22	74.07
9	72.22	40.00	50.00	79.01	50.00	65.35	72.22	59.25
10	88.88	70.00	80.00	91.35	67.90	91.50	88.88	77.77
11	72.22	70.00	80.00	79.01	67.90	81.69	80.55	100.00
12	77.77	80.00	90.00	79.01	79.01	91.50	100.00	92.59
13	77.77	80.00	70.00	100.00	79.01	81.69	88.88	92.59
14	88.88	70.00	80.00	100.00	79.01	71.89	80.55	77.77
15	88.88	70.00	80.00	91.35	91.35	91.50	69.44	77.77
16	77.77	80.00	100.00	79.01	67.90	91.50	80.55	100.00
17	100.00	80.00	100.00	79.01	79.01	100.00	88.88	77.77
18	88.88	70.00	90.00	100.00	67.90	100.00	69.44	92.59
19	100.00	80.00	90.00	100.00	79.01	91.50	88.88	92.59
20	100.00	80.00	90.00	100.00	67.90	80.00	83.33	92.59
NS1	100.00	100.00	100.00	96.66	100.00	100.00	100.00	100.00
NS2	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Appendix K presents our participants' scores on each of the eight categories of phonemic contrasts according to gender and L1. A comparison of means for each contrast using a t-test was carried out to determine possible differences in performance. Results showed no statistically significant differences due to gender for any of the eight phonemic contrasts. However, when we examined scores with regard to L1, the Polish/Punjabi group performed significantly better in four categories of phonemic contrasts (contrasts 2, 4, 5 and 7, i.e. unstressed vowels, fricatives, liquids and vowel combinations). Significance values for these four categories of phonemic contrasts were as follows: $p = 0.02$, $p = 0.02$, $p = 0.009$ and 0.02 respectively.

The results that we have presented above do not provide any information on the importance of individual phonemic contrasts probed through the single word intelligibility test. In other words, the overall score in a specific category could be due

to deviances involving an individual phonemic contrast or arise as a result of errors across several contrasts. Appendix M (Figures M1-M8) offers the scores for each specific minimal pair. It is noteworthy that, as far as unstressed vowels are concerned, speakers' scores were especially low in one phonemic contrast: /i/–/e/ with a score of 44.51% (mean percentage across speakers). Less problematic were /i/–/o/, /i/–/u/, /o/–/u/ and /a/–/u/ with percentages that ranged from 77.77% to 88.88%. With regard to liquids, low scores were found in four of the phonemic contrasts: /r/–/ɾ/, /r/–/d/, /r/–Ø and /ɻ/–/l/. The rest of the contrasts in this category did not seem to cause major issues (percentages ranged from 88.88% to 100%). In the vowel combination category, /e/–/ei/, with a score across speakers of 55.55%, was the only contrast to exhibit a score lower than 80%. Only one minimal pair in the nasal category did not achieve a perfect score of 100%. Scores for /n/–/ɲ/ reached 54.52%. Similarly, the majority of minimal pairs in the stop category reached a score of over 80%: /p/–/f/, /p/–/m/, /p/–/k/, /b/–/m/, /b/–/f/, /d/–/θ/, /d/–/g/ and /t/–/k/, all reached a perfect score of 100%. On the other hand, there were two minimal pairs that exhibited scores lower than 70%: /g/–/x/ and /d/–/t/. This aforementioned trend of errors affecting only selected minimal pairs within a specific category can also be seen in the case of affricates, where only the pair /tʃ/–/k/ did not a perfect score of 100%. In the case of fricatives, errors seem to concentrate on three minimal pairs. The scores for /s/–/tʃ/, /ɻ/–/tʃ/ and /ɻ/–/ɲ/ were 77.77%, 77.81% and 66.66%. The rest of the contrasts reached a perfect score of 100%.

4.4 Segmental error analysis (sentence, passage, semi-spontaneous production task)

The single word intelligibility test enabled us to obtain information on specific categories of phonemic contrasts and their potential impact on intelligibility scores.

However, the tests used to measure intelligibility in the sentence, passage and semi-spontaneous production tasks offered an intelligibility score based on the number of content words correctly identified by our group of sixty evaluators. In other words, the intelligibility tests could not provide any information on the percentage of accurate phoneme production. In order to determine the potential impact of segmental deviations on intelligibility scores in connected speech, we undertook a segmental error analysis of our participants' productions in the sentence, passage and semi-spontaneous production tasks. Participants' utterances were phonemically and phonetically transcribed by the main researcher plus a second researcher with extensive training in phonetics. As far as transcription conventions are concerned and given that all participants shared the same native language, two sets of phones, corresponding to British English and Castilian Spanish, plus a set of diacritics were used in the transcription process. A point by point transcription agreement percentage was calculated using the following formula (Cucchiarini 1996):

Percentage agreement = number of agreements/number of disagreements + number of agreements * 100

The transcription agreement between both researchers was 95.50%. Those instances of disagreement were discussed after the initial transcription. In cases where the disagreement persisted, a third researcher was consulted.

4.4.1 Vowels

As far as stressed vowel production is concerned, scores ranged from 82% to 96% at sentence level (M = 85.25, SD = 3.75). Scores ranged from 84% to 93.25% (M = 88.62, SD = 2.75) in the passage task and from 80% to 89% (M = 84.50, SD = 3.25) in semi-spontaneous production. Accurate production of unstressed vowels at sentence level ranged from 62.40% to 87.80% (M =79.60, SD = 7.35), while, at passage level, accurate production ranged from 75.40% to 85%, (M = 79.44, SD = 2.75). In the semi-spontaneous production task, scores ranged from 60% to 81% (M = 70.50, SD = 2.89%). The following patterns of error were observed:

1. In terms of stressed vowels, scores at sentence level show that the percentage of accurate production for /a/ was 82.30%. Substitution of /a/ by /æ/ reached 14% of occasions, while substitutions by /ɑː/ occurred in 3.70% of instances. At passage level, /a/ was accurately produced on 80% of occasions. Substitution by /æ/ could be found in 20% of instances. Unstressed /a/ was replaced by /ə/ at all levels of analysis (20.40% at sentence level, 20.56% passage level and 29.50% in the semi-spontaneous production task).
2. Target-like production for stressed /e/ reached 85.40%. Substitution by /i/ occurred in 14.60% of cases. At passage level, accurate production was 82.10% with 7.40% of cases exhibiting a substitution by /i/. Replacement by a diphthong, /ei/, occurred in 10.50% of instances. In semi-spontaneous production, accurate production reached 80.50%, with substitutions by /i/ in 15% of instances and replacement by /ei/ in 4.50% of cases. In terms of unstressed /e/, we recorded substitutions by /ə/ in 25% (sentence), 29% (passage) and 31% (semi-spontaneous) of instances.
3. At sentence level, 83.10% was the accurate production of stressed /i/. Substitution by /ɪ / occurred in 10.50 % and by /ɑɪ/ in 6.40% of cases. At passage level, accurate production reached 79.20%, while substitutions by /ɪ/ were detected in 20.80% of instances. We recorded an accurate production of 78% in the semi-spontaneous production task. As far as unstressed /i/ is concerned, we recorded a substitution by /ə/ in 35% of cases (sentence), 32% (passage) and 38% of instances (semi-spontaneous production task).
4. At sentence level, results show that /o/ was accurately produced in 82.20% of cases. From a phonetic standpoint, it was replaced by [ɔː] in 17.80% of instances. At passage level, it was accurately pronounced in 78.40% of cases.

In 10.30% of cases, it was replaced by /ə/, in 11.30% of instances by the diphthong /ou/ and in 3.10% by [ɔː]. As far as unstressed /o/ is concerned, it was replaced by /ə/ in 25% (sentence), 29% (passage) and 32% (semi-spontaneous production) of instances.

5. At sentence level, /u/ was accurately produced in 66.20%. It was replaced by /ʌ/ in 16.20% of cases and by /ʊ/ in 17.60%. At passage level, accurate production was 76.30% with substitutions by /ʌ/ in 3.50%, by [uː] in 5% and /ʊ/ in 15.20% of the cases. In the semi-spontaneous production task, /u/ was replaced by /ʌ/ in 20% of instances. As far as unstressed /u/ is concerned, it was substituted by /ə/ in 30% of cases (sentence), 35% (passage) and 38% (semi-spontaneous task).

It is noteworthy that percentages of accurate production are very similar at sentence and passage levels. The following hierarchy of target-like production, from highest rate of accurate production to lowest, can be put forward for stressed vowels (sentence) : /a/> /i/> /e/> /o/>/u/. At passage level, the hierarchy is the following: /e/> /i/> /o/> /a/> /u/. In the semi-spontaneous production task, we recorded the following hierarchy: /e/ > /a/> /u/ > /i/> /o/. As far as unstressed vowels are concerned, we can put forward the following hierarchies of difficulty: /a/> /e/> /o/> /i/> /u/ (sentence level); /a/> /e/> /o/> /i/> /u/ (passage); /a/>/e/>/o/>/u/>/i/ (semi-spontaneous production).

4.4.2 Consonants

The following patterns of error were observed in the sentence, passage and semi-spontaneous production tasks:

1. /p/ was accurately produced in 100% of cases in the sentence, passage and

semi-spontaneous production tasks. From a phonetic/allophonic perspective, it is noteworthy that [p^h] occurred in 57% of instances (sentence), while in the passage and semi-spontaneous task, we found 56.30% and 59% of cases where [p] was replaced by [p^h].

2. /b/ was accurately uttered in 100% at all levels. Again, from a purely allophonic perspective, [β] was accurately targeted in 53.70% of instances. Substitutions by [b] occurred in 35% and by [v] in 11.30% of cases (sentence). At passage level, [β] was accurately produced in 49.60% of cases, with substitutions by [b] in 37% of instances and by [v] in 13.40%. In the semi-spontaneous production task, [β] was replaced by [b] in 40% of instances.
3. Both /m/ and /f/ were produced with perfect accuracy at all levels.
4. At sentence level, 84.40% was the accurate production for /θ/ with substitutions by /s/ in 15.60% of cases. At passage level, target-like production reached 72.30% of cases while substitutions by /s/ reached 27.70%. In the semi-spontaneous task, /θ/ was replaced by /s/ in 30% of cases.
5. /s/ reached 83.6% of target-like production with 10% of substitutions by /z/ and 7.40% by /θ/ (sentence). At passage level, accurate pronunciation was 78.80% with substitutions by /z/ occurring in 11% of cases and substitutions by /θ/ in 10.80%. In semi-spontaneous production, /s/ was replaced by /z/ in 15% of instances.
6. /t/ was correctly targeted in 100% of instances at all levels of analysis. From a phonetic standpoint, [t̪] reached an accurate production of 51.90% with substitution by [t^h] in 48.10% of cases (sentence level). At passage level, accurate production of [t̪] reached 50.40% with substitution by [t^h] in 40% and

by [ʔ] in 9.60% of instances. It is important to note that, in the semi-spontaneous production task, accurate production of [t̪] reached 48%.

7. /d/ was accurately produced in 100% of instances at all levels of analysis. From an allophonic perspective, target-like pronunciation of [d̪] was 48.10%, with substitutions by [d] in 52.90% of cases. At passage level, accurate production of [d̪] reached 46.60%, with substitutions by [d] in 40% of cases and by [ð] in 3.40% of instances. In the semi-spontaneous production task, target-like production of [d̪] was 42.50%. Furthermore, [ð] was accurately targeted in 66.60% of cases with substitutions by [d] in 33.40% of instances (sentence level). At passage level, target-like production was 80% with 20% of substitutions by [d]. Target-like production of [ð] in the semi-spontaneous production task reached 60%.
8. /n/ reached 100% of accuracy at all levels of analysis.
9. /l/ reached 100% of accuracy at all levels of analysis.
10. /r/ was accurately produced in 78.70% of cases, with substitutions by /ɹ/reaching 15% and a deletion of this sound in 6.3% of instances. At passage level, accurate production reached 71.80% with deletion of this sound in 15% of cases, and substitution by /ɹ/ in 13.20%. In the semi-spontaneous task, accurate production of /r/ reached 69.50% with replacement by /ɹ/ in 30.50% of instances.
11. /ɾ/ was accurately produced in 70.60% of cases with substitutions by /ɹ/ in 18% and by /r/ in 11.40% (sentence task). At passage level, correct production reached 68.40% with substitutions by /ɹ/ in 20% of cases and by /r/ in 11.60%. In semi-spontaneous conversation, /ɾ/ was accurately produced in 70% of instances.
12. At sentence level, /tʃ/ was accurately produced in 78.10% of cases with

substitutions by /ʃ/ in 21.90% of instances. At passage level, accurate production reached 72.06% with substitutions by /ʃ/ in 27.90% of occasions. Target-like production in semi-spontaneous conversation reached 75%, with 25% of instances with substitutions by /ʃ/.

13. /j/ was accurately produced in 77% of cases with substitutions by /l/ in 23% of cases. At passage level, accurate production reached 73.80% with substitutions by /l/ reaching 26.20%. In the semi-spontaneous production task, accurate production reached 75% of cases.
14. /ɲ/ was accurately produced in 86.3% of cases with substitutions by /n/ in 13.60% of instances (sentence task). At passage level, accurate production reached 84.10% of cases with substitutions by /n/ in 15.90% of instances. In semi-spontaneous production, target-like production reached 80% of cases, with substitutions by /n/ in 20% of instances.
15. /x/ reached a target-like production of 80.40% with substitutions by /j/ in 15% and by /dʒ/ in 4.60% of cases. At passage level, target-like production was 75.1% with substitutions by /j/ in 12% and by /dʒ/ in 12.90% of cases. In semi-spontaneous production, accurate production reached 72% of cases with substitutions by /j/ in 28% of instances.
16. /k/ reached 100% of accuracy at all levels of analysis. From a purely phonetic standpoint, there was a substitution by [kʰ] in 47.80% of instances. At passage level, accurate production reached 53.60% of cases with substitutions by [kʰ] occurring in 46.40% of cases.
17. /g/ was accurately produced at all levels of analysis.

We can put forward the following hierarchy of difficulty in connected speech from the highest percentage of accuracy to the lowest:

Sentence: /b/, /p/, /d/, /k/, /m/, /f/, /l/, /n/, /g/, all 100,> /ɲ/ (86.3)> /ɵ/ (84.40)> /s/ (83.60)>/x/ (80.40)> /r/ (78.70)> /tʃ/ (78.10)> /j/ (77)> /ɾ/ (70.60).

Passage: /b/, /n/, /m/, /f/, /l/, /g/, /d/, /k/, /p/ all 100, >/ɲ/ (84.1)> /s/ (78.80)> /j/ (73.80)> /tʃ/ (72.06)> /r/ 71.80)> /θ/ (72.30) > /ɾ/ (68.4)

Semi-spontaneous production task: /b/, /m/, /f/, /n/, /l/, /g/, /d/, /k/, /p/ all 100, >/ɲ/ (80)> /s/ (75)> /tʃ/ (75)> /x/ (72)> /r/> (70)> /ɾ/ (69.50).

4.4.3 Vowel combinations

1. At sentence level, the combination /ía/ (hiatus) was accurately pronounced on 76.6% of occasions, with reductions of /a/ to /ə/ in 17% of instances and substitutions of /i/ by /ɛ/ in 6.40% of cases. At passage level, it was accurately produced in 78.80% of instances. Substitutions of /i/ by /ɛ/ occurred in 4% and reductions of /a/ to /ə/ in 17.2%. In the semi-spontaneous production task, accurate production reached 75%, with reductions of /a/ to /ə/ in 25% of cases.
2. At sentence level, [je] was correctly uttered in 74.40% of instances with a creation of a hiatus instead of a diphthong in 25.60%. At passage level, accurate production reached 78.50% and the production of a hiatus combination in 22.50% of cases. In the semi-spontaneous task, accurate production reached 70% of instances.
3. /aj/ only appeared at sentence level and recorded an accurate production of

90%. It was replaced by /ɛ/ in 10% of cases.

4. Target-like pronunciation of [jó] occurred in 81.60% of cases with a creation of a hiatus instead of a diphthong in 19.40% of cases (sentence level). At passage level, accurate production reached 85% of cases, while in the semi-spontaneous task accurate production reached 78% of instances.
5. /ei/ appeared only at sentence level and was accurately produced in 88% of cases.
6. [we] was correctly targeted in 74.20% of cases. The pronunciation of a hiatus instead of a diphthong appeared on 25.80% of occasions. Accurate production reached 77% at passage level and 72% of instances in the semi-spontaneous task.
7. /oi/ appeared only at sentence level and was accurately produced in 100% of instances.
8. At sentence level /ea/ was correctly uttered in 71.06% of instances while substitutions by /ɪə/ occurred in 29.04% of instances. In the semi-spontaneous task, accurate production reached 75%. This combination did not appear at passage level.
9. At sentence level, [ja] reached an accurate production of 80% with a production of a hiatus instead of a diphthong in 20% of cases. At passage level, accurate production reached 92% of instances, while production in the semi-spontaneous task reached 82%.

4.4.4 Overall segmental deviations by error category

In order to facilitate the correlation and multiple-regression analyses that will be carried out in Section 4.5, as well as to enable us to directly compare the results with those obtained at word level, segmental deviations were grouped in eight categories of phonemic errors based on the type of phoneme targeted in each instance. The eight chosen categories of errors are similar to the ones used in the single word intelligibility test:

1. Errors affecting affricates
2. Errors affecting unstressed vowels
3. Errors affecting stressed vowels
4. Errors affecting fricatives
5. Errors affecting liquids
6. Errors affecting stops
7. Errors affecting vowel combinations
8. Errors affecting nasals

Appendix N presents the scores per speaker for each category at sentence and passage levels, as well as in the semi-spontaneous production task. It is important to note that there are close similarities between the results obtained at the three levels of analysis. The highest accurate production was found for fricatives, affricates and nasals. On the other hand, the lowest target-like production involved unstressed vowels, liquids and vowel combinations. As pointed out above, the hierarchy of difficulty was very similar for phonemic errors in the sentence, passage and semi-spontaneous production tasks. Only minor percentage differences were found in the ranking of unstressed vowels, liquids and vowel combinations, as well as in the ranking of fricatives, affricates and nasals.

4.5 Assessment of speech rate, pause frequency and pause duration

Appendix O (Tables O.1-3) shows the mean results in connected speech for speakers with regard to speech rate, pause frequency and pause duration. At sentence level, descriptive statistics were as follows: speech rate (M = 3.21, SD = 0.44), pause frequency (M = 1.79, SD = 0.95) and pause duration (M = 0.33, SD = 0.14). At passage level, we recorded M = 3.23, SD = 0.53 for speech rate, while values for pause frequency and pause duration were as follows: M = 1.76, SD = 1.01; M = 0.37, SD = 0.16. In the semi-spontaneous production task, we recorded M = 3.22 and SD = 0.54 for speech rate, M = 1.76 and SD = 1.02 for pause frequency, and M = 0.42, SD = 0.19. Mean comparisons of results for different subgroups according to gender, L1 and native versus non-native speakers are presented in Appendix P. It was expected

that L2 learners would exhibit a slower speech rate when compared to native Spanish speakers. At sentence level, the average speech rate for native speakers was higher than for L2 learners (4.20 versus 3.21). A t-test confirmed than those differences were statistically significant ($t = -3.044$, $p = 0.006$). There were no significant differences due to gender ($t = 0.60$, $p = 0.55$). Bilingual students again performed significantly better ($t = -2.44$, $p = 0.02$). As far as pause frequency and pause duration are concerned, no significant differences were found due to gender for both variables ($t = 0.09$, $p = 0.92$; $t = -1.02$, $p = 0.31$) or between native and non-native speakers for pause frequency ($t = 1.8$, $p = 0.07$), while there were significant differences between native and non-native speakers for pause duration ($t = 2.38$, $p = 0.02$) and between English and Polish/Punjabi speakers for both pause frequency and pause duration ($t = 2.33$, $p = 0.03$; $t = 2.19$, $p = 0.04$).

At passage level, there were statistically significant differences between native Spanish speakers and non-native speakers with regard to speech rate ($t = 3.34$, $p = 0.03$), pause frequency ($t = -6.49$, $p = 0.00$) and pause duration ($t = -3.49$, $p = 0.02$). Again, there were no significant differences between male and female speakers for any of three suprasegmentals under study ($t = 0.73$, $p = 0.47$; $t = 0.04$, $p = 0.96$; $t = -1.57$, $p = 0.13$). There were, however, statistically significant differences between bilingual speakers (English/Polish or English/Punjabi speakers) and monolingual English speakers for speech rate ($t = -2.30$, $p = 0.03$) and pause frequency ($t = 2.26$, $p = 0.03$).

In the semi-spontaneous production task, there were statistically significant differences in performance between the native and non-native Spanish speakers for speech rate and pause frequency ($t = 3.54$, $p = 0.02$; $t = -6.77$, $p = 0.00$) but not for pause duration ($t = 0.47$, $p = 0.89$). As far as gender differences are concerned, no differences were found between males and females for any of the variables under study ($t = 0.64$, $p = 0.62$; $t = -0.10$, $p = 0.91$; $t = -1.01$, $p = 0.32$). In a similar way as in the sentence and passage tasks, differences in performance were found between bilingual and monolingual speakers for speech rate and pause frequency ($t = -2.28$, $p = 0.03$; $t = 2.26$, $p = 0.03$) but not for pause duration ($t = 1.88$, $p =076$).

To summarise, we did not observe any significant differences due to gender in terms of performance in any of the three variables under consideration. However, native Spanish speakers, as well as bilingual English/Polish and English/Punjabi speakers, performed in a significantly different manner with regard to speech rate and

speech frequency. Both groups exhibited a faster speech rate and less frequent pauses. Furthermore, even though both groups exhibited shorter pauses, this difference cannot be considered statistically significant.

4.6 Correlation analyses

One of the goals of this study is to examine the relationship between speech intelligibility at word level and in connected speech. Moreover, this dissertation aims at exploring the impact of segmental and suprasegmental deviations on speech intelligibility. In addition, the possible impact of a series of individual differences on intelligibility scores will also be assessed. Correlation and multiple-regression analyses will serve to explore the potential relationship between the aforementioned variables. They will also be used to identify the set of segmental and suprasegmental factors that best predict intelligibility scores at word level and in connected speech.

Before undertaking the appropriate statistical analyses, a preliminary examination of the variables under study was carried out in order to ensure that data were normally distributed from a statistical point of view. The following variables were therefore examined for skewness, kurtosis and normality of the distribution: aptitude, motivation, DELE scores, eight categories of phonemic contrasts, three suprasegmental variables (speech rate, pause frequency and pause duration), and intelligibility scores at word, sentence, passage and in the semi-spontaneous production task.

4.6.1 Correlation coefficients

Correlation analyses will be presented in four parts. First, we will explore the relation between individual differences and intelligibility scores at word level and in connected speech. Second, we will analyse the correlation between intelligibility scores at word level and in connected speech. Third, following Rogers (1997), we will attempt to validate the usefulness of the single word intelligibility test by exploring the relation between intelligibility scores in connected speech (sentence, passage and semi-spontaneous production) and the scores obtained in the eight categories of

phonemic errors under study in the single word intelligibility test. Finally, the relation between intelligibility scores in connected speech and the segmental and suprasegmental error analysis carried out at each level of analysis will be examined.

Table 14 below presents the correlation between aptitude, motivation, level of proficiency and intelligibility scores at word, sentence and passage levels, as well as in the semi-spontaneous production task. Pearson's r was used to perform the different correlation analyses. The Pearson correlation coefficient, designed for quantitative variables, is an index that measures the degree of co-variation between two variables. Pearson's r is an index of straightforward calculation and interpretation. Its absolute values range between –1 and +1. The magnitude of the relationship is specified by the numeric value of the coefficient, the sign reflecting the direction of such value. There are, however, no precise guidelines as to how to interpret this correlation coefficient. Much seems to depend on the nature of the investigation and the size of the sample. In general terms, Bisquerra (2004) indicates that a correlation between 0.40 and 0.60 can be considered as moderate, while between 0.60 and 0.80 as high. In this study, level of proficiency (DELE scores) proved to be highly correlated with intelligibility at word level (r = 0.68) and in connected speech (r = 0.69, 0.69, 0.55). Moderate correlations were found between motivation and intelligibility scores at the word, sentence and passage levels (r = 0.51, 0.45, 0.55). Correlation between aptitude and intelligibility at the same three levels of analysis was slightly higher (r = 0.55, 0.63 and 0.67). It is noteworthy that no significant correlation was found between aptitude or motivation and intelligibility scores in the semi-spontaneous production task. In summary, while correlations were significant between intelligibility scores at word, sentence and passage levels and the individual differences under study (level of proficiency, motivation and aptitude), no statistically significant correlations were recorded when assessing intelligibility in the semi-spontaneous production task.

Table 14: Correlations between level of proficiency, motivation, aptitude and intelligibility at word level and in connected speech

		Word	**Sentence**	**Passage**	**Semi-sp**
DELE	Pearson Correlation	.68	.69	.69	.55
	Sig. (2-tailed)	.00	.00	.00	.01
	N	20	20	20	20
Motivation	Pearson Correlation	.51	.45	.55	.29

	Sig. (2-tailed)	.02	.04	.01	.21
	N	20	20	20	20
Aptitude	Pearson Correlation	.55	.63	.67	.42
	Sig. (2-tailed)	.01	.003	.00	.07
	N	20	20	20	20

Table 15 shows the correlation coefficients between intelligibility scores at word level and in connected speech. The correlation between sentence and passage scores is highly significant (r = 0.89), while correlation between intelligibility scores at word level and in connected speech (sentence, passage and semi-spontaneous production) is high (r = 0.70, 0.82, 0.63 respectively). In other words, a high score in the single word intelligibility tests translates into high scores at sentence, passage and in semi-spontaneous production. It is noteworthy that the correlation between the three scores measuring intelligibility in connected speech is significantly high. As pointed out in Chapter 3, previous intelligibility studies within the areas of speech disorders and foreign-accented speech have yielded contradictory results when it comes to establishing a relation between intelligibility at word level and in connected speech.

Table 15: Correlation between intelligibility scores in the word, sentence, passage and semi-spontaneous production tasks

		Word	**Sentence**	**Passage**	**Semi-spont**
Word	Pearson Correlation	1.00	.70	.82	.63
	Sig. (2-tailed)		.00	.00	.00
	N	20	20	20	20
Sentence	Pearson Correlation	.70	1.00	.89	.72
	Sig. (2-tailed)	.00		.00	.00
	N	20	20	20	20
Passage	Pearson Correlation	.82	.89	1.00	.64
	Sig. (2-tailed)	.00	.00		.00
	N	20	20	20	20
Semi-Spont	Pearson Correlation	.63	.72	.64	1.00
	Sig. (2-tailed)	.00	.00	.00	
	N	20	20	20	22

4.6.1.2 Validating the single word intelligibility test

Following Rogers (1997), we decided to validate the usefulness of the single word intelligibility test by exploring the possible correlation between the intelligibility scores in connected speech and the scores on the different categories of phonemic errors yielded by the single word test. Table 16 below shows the correlation between the eight categories of phonemic errors and intelligibility scores in the sentence, passage and the semi-spontaneous production tasks. Significant differences can be seen when it comes to correlations between the eight categories of phonemic errors and intelligibility in connected speech. The correlation coefficient is generally higher in the passage task, when compared to sentence scores, for all phonemic contrasts except for categories 4 and 6. Several categories only exhibited a moderate degree of correlation (in the 0.5 or 0.6 range). This was the case for categories number 1, 3, 5, 7 and 8, at the sentence level, and 7 and 8 at the passage level.

A high degree of correlation was shown between intelligibility at the sentence level and one specific category of phonemic errors: contrasts based on the reduction of unstressed vowels (r = .70). As far as passage scores are concerned, a high level of correlation can be found with regard to contrasts based on the reduction of unstressed vowels, stressed vowels and liquids. Moderate correlations were found for the category related to vowel combinations.

Only five categories of phonemic contrasts showed significant correlations with intelligibility scores in the semi-spontaneous production task (contrasts 2, 3, 5, 6, and 8). Furthermore, those statistically significant correlations were only moderate (between r = 0.46 and 0.66), especially when compared to the correlations recorded at the word, sentence and passage levels. If we examine the correlation coefficients across all three levels of analysis (sentence, passage and semi-spontaneous production), it is noteworthy that category 4 (fricatives) did not show any significant correlations at any of the levels under study, while contrast 2, affecting the reduction of unstressed vowels, was highly correlated with intelligibility scores at all levels. Moreover, there seems to be a clear difference between the correlations recorded at sentence and passage levels and those recorded in semi-spontaneous production.

Table 16: Correlation between scores on eight categories of phonemic contrasts from the single word test and intelligibility scores in connected speech

		Sentence	Passage	Semi-spont
C1	Pearson Correlation	.55	.60	.37
	Sig. (2-tailed)	.01	.00	.11
	N	20	20	20
C2	Pearson Correlation	.70	.80	.66
	Sig. (2-tailed)	.00	.00	.00
	N	20	20	20
C3	Pearson Correlation	.55	.71	.55
	Sig. (2-tailed)	.01	.00	.01
	N	20	20	20
C4	Pearson Correlation	.30	.26	.32
	Sig. (2-tailed)	.19	.26	.16
	N	20	20	20
C5	Pearson Correlation	.53	.71	.46
	Sig. (2-tailed)	.01	.00	.04
	N	20	20	20
C6	Pearson Correlation	.73	.71	.60
	Sig. (2-tailed)	.00	.00	.00
	N	20	20	20
C7	Pearson Correlation	.46	.62	.33
	Sig. (2-tailed)	.03	.00	.15
	N	20	20	20
C8	Pearson Correlation	.53	.63	.51
	Sig. (2-tailed)	.01	.00	.02
	N	20	20	20

4.6.1.3 Correlation between segmental error analysis and intelligibility scores in connected speech

Table 17 below shows the correlation coefficients between the intelligibility scores in the sentence task and the scores in each category of phonemic errors. Two categories did not show any significant correlation with intelligibility scores: affricates and nasals. A high significant correlation was found for category 2 (unstressed vowels), 3 (stressed vowels) and 7 (vowel combinations). A moderate degree of correlation was observed for fricatives, liquids and stops.

Table 17: Correlation between error categories at sentence level and intelligibility scores in the sentence task

		Sentence
C1	Pearson Correlation	.43
	Sig. (2-tailed)	.05
	N	20
C2	Pearson Correlation	.71
	Sig. (2-tailed)	.00
	N	20
C3	Pearson Correlation	.65
	Sig. (2-tailed)	.00
	N	20
C4	Pearson Correlation	.48
	Sig. (2-tailed)	.02
	N	20
C5	Pearson Correlation	.52
	Sig. (2-tailed)	.01
	N	20
C6	Pearson Correlation	.46
	Sig. (2-tailed)	.00
	N	20
C7	Pearson Correlation	.66
	Sig. (2-tailed)	.00
	N	20
C8	Pearson Correlation	.26
	Sig. (2-tailed)	.25
	N	20

As we can see in Table 18 below, results for the passage task were very similar to those obtained at sentence level. Three categories did not show significant correlations with the set of intelligibility scores: categories 1, 4 and 8. Again, the highest degree of correlation was found for unstressed vowels (category 2), stressed vowels (category 3) and vowel combinations (category 7). Moderate coefficients were recorded for categories 5 and 6, i.e. liquids and stops.

Table 18: Correlation between error categories at passage level and intelligibility scores in the passage task

		Passage
C1	Pearson Correlation	.06
	Sig. (2-tailed)	.78
	N	20
C2	Pearson Correlation	.66
	Sig. (2-tailed)	.00
	N	20
C3	Pearson Correlation	.63
	Sig. (2-tailed)	.00
	N	20
C4	Pearson Correlation	.18
	Sig. (2-tailed)	.44
	N	20
C5	Pearson Correlation	.52
	Sig. (2-tailed)	.01
	N	20
C6	Pearson Correlation	.42
	Sig. (2-tailed)	.00
	N	20
C7	Pearson Correlation	.62
	Sig. (2-tailed)	.00
	N	20
C8	Pearson Correlation	.43
	Sig. (2-tailed)	.05
	N	20

Table 19 below shows that the pattern of results obtained in the sentence and passage tasks repeats itself in the semi-spontaneous production task. Again categories 1, 4 and 8 (affricates, fricatives and nasals) did not show any significant correlation with intelligibility scores. The highest correlation coefficients were yielded by unstressed vowels (category 2), and vowel combinations (category 7), while moderate correlations were recorded for stressed vowels (category 3) and stops (category 6).

Table 19: Correlation between error categories and intelligibility scores in the semi-spontaneous production task

		Semi-spontaneous
C1	Pearson Correlation	.26
	Sig. (2-tailed)	.25
	N	20
C2	Pearson Correlation	.63
	Sig. (2-tailed)	.00
	N	20
C3	Pearson Correlation	.49
	Sig. (2-tailed)	.02
	N	20
C4	Pearson Correlation	.43
	Sig. (2-tailed)	.05
	N	20
C5	Pearson Correlation	.32
	Sig. (2-tailed)	.16
	N	20
C6	Pearson Correlation	.54
	Sig. (2-tailed)	.01
	N	20
C7	Pearson Correlation	.62
	Sig. (2-tailed)	.00
	N	20
C8	Pearson Correlation	.17
	Sig. (2-tailed)	.47
	N	20

It has become evident that certain patterns of results seem to repeat themselves across the different levels of analysis. The highest degree of correlation at all levels was obtained for errors affecting unstressed vowels, stressed vowels and vowel combinations. Affricates, fricatives and nasals did not show any significant correlation with intelligibility scores. Moderate correlations were recorded for liquids and stops.

4.6.1.4 Correlation between intelligibility in connected speech and suprasegmentals

Table 20 below presents the correlation coefficients for the three suprasegmental variables under study (speech rate, pause frequency and pause duration) and intelligibility scores in connected speech (sentence, passage and semi-spontaneous production task). As shown in Table 20, correlation coefficients are extremely high between the three suprasegmental variables and the intelligibility scores at the sentence and passage levels. Furthermore, a high level of correlation was found among the three suprasegmental variables. As explained below in Section 4.7, this high level of intercorrelation led to a multicollinearity problem when running the multiple-regression analyses. The three suprasegmental variables exhibited correlations above 0.8 with both intelligibility at the sentence and passage level, and above 0.57 with intelligibility scores in the semi-spontaneous production task. Correlation coefficients were generally higher than those found between the eight phonemic contrasts and intelligibility scores in connected speech. The highest correlation coefficient was yielded by pause frequency (r = -0.73) followed by speech rate (r = 0.70) and pause duration (r = -0.57).

Table 20: Correlation between speech rate, pause frequency, pause duration, and intelligibility scores in connected speech

		Sentence	Passage	Semi-spont
Speech rate	Pearson Correlation	.88	.82	.70
	Sig. (2-tailed)	.00	.00	.00
	N	20	20	20
Pause frequency	Pearson Correlation	-.93	-.82	-.73
	Sig. (2-tailed)	.80	.69	.69
	N	20	20	20
Pause duration	Pearson Correlation	-.80	-.73	-.57
	Sig. (2-tailed)	.00	.00	.01
	N	20	20	20

4.7 Multiple-regression analysis

This section will put forward a series of multiple-regression analyses between the different categories of phonemic errors, the suprasegmental variables under study and the intelligibility scores in connected speech. Our goal is to analyse the interaction between the different factors, while attempting to select the best possible model (best possible subset of variables) that can predict intelligibility scores at different levels of analysis.

A multivariate regression analysis is a statistical method that aims at establishing a mathematical relationship between a set of predictor factors and a dependent variable. Therefore, a regression model attempts to obtain an equation in order to 'predict' the value of the dependent variable from those of the predictor variables. It also attempts to quantify the potential relationship between both sets of variables.

The inclusion of variables in the regression equation is a fundamental problem that arises when constructing a multivariate model. This selection will have a major impact on the estimation of the best possible model of predictive factors. It becomes necessary, initially, to define what we mean by 'best model'. If our goal is to find a predictive model, we will then strive to search for a model that can provide us with reliable and accurate predictions. If our goal is to build an explanatory model, we will then seek to obtain accurate estimates of the equation coefficients on which to base our own inferences.

An important step in building a regression model revolves, therefore, around the selection of variables. The mechanisms for selecting variables are not easy to specify because they largely depend on the type of model (predictive or explanatory), the context of use and the characteristics of the analysed process. One rule that can be used in the variable selection process is the principle of parsimony, which posits that if we have to choose between two possible models, the simplest model, i.e. the one requiring less number of assumptions, should be adopted.

There are various systematic procedures for choosing a multiple-regression model. We can start by including all possible independent variables and then eliminate those whose presence does not improve the quality of the model according to a specified criterion (backward model selection criteria). On the other hand, we

could start with a single independent variable and add those variables that significantly improve the regression model (forward model selection procedure). Another alternative, not always feasible if the number of variables is large, is to evaluate all possible regression models with all possible combinations of variables (best subsets procedure).

A stepwise regression, frequently used in the literature of speech disorders, is a modified version of the forward regression process in which the equation variables are introduced one by one. First, the predictor variable exhibiting the highest degree of correlation with the dependent variable is selected. Second, partial correlations are calculated between the other independent variables and the criterion variable. The variable that is neutralised by the partial correlation is chosen first. Then, the second selected variable is the one with the highest partial correlation. Upon introducing each new variable, the statistical significance of the variables already in the model is reassessed and this process may lead to the removal of those variables that are no longer statistically significant. The use of stepwise regression is widespread in the field of speech disorders (see Weismer 2008, for important caveats and methodological flaws stemming from this statistical procedure).

As Rumsey (2007: 128) points out: 'Because of its versatility and the comprehensive way it looks at all possible models, the best subsets model is generally the model of choice by statisticians'. Given some of the criticism levelled against the stepwise procedure and considering that it is feasible for most well-known statistical packages to carry out a large number of statistical calculations, the best subsets procedure was adopted here in the assessment of the relation between the independent variables under study and the intelligibility scores.

Before presenting the data related to the three multiple-regression analyses undertaken in this study, two important issues need to be addressed: first, the question of the number of observations required to perform a multiple-regression analysis and, second, the issue of multicollinearity.

The number of participants in this study is slightly smaller than the number usually recommended by statisticians for multiple-regression analysis (Tabachnick and Fidell 1996). However, after examining some of the literature available in the field, examples of studies with a similar or even smaller number of cases are rather frequent (Whitehill 1997). This fact, along with the strong R square coefficients

obtained during the analysis (see Tables 21, 22 and 23 below), certainly justifies the choice of this statistical procedure in this study.

Multiple-regression assumes that there is not an exact linear relationship among the predictor variables. The problem of multicollinearity refers, in particular, to the existence of a quasi-linear relationship among the independent variables. Perfect multicollinearity is not commonly seen in practice, unless the model is poorly designed. Instead, it is more common to see the existence of a quasi-linear relationship among the predictor variables. It is precisely this relationship that makes it difficult to accurately quantify the effects of each factor.

A high correlation coefficient among all or some of the predictor variables points to the possible existence of multicollinearity in a multiple-regression model. Furthermore, a variance inflation factor (VIF) of 5 or above usually points in the same direction. In our case, the correlation coefficients among the three suprasegmental variables were very high (Pearson's r was above 0.8) and the VIF values were definitely over 5.

Several procedures are usually suggested in order to solve the multicollinearity problem. Multicollinearity can be mitigated if the predictor variables that are more affected by the high degree of intercorrelation are eliminated from the model. The problem with this solution is that, if the original model was correct, estimates for the new model could result in biased results. Increasing the sample size or replicating the study with a different set of participants from the same population may be an alternative to simply dropping some of the variables. However, this is not always possible given the constraints generated in certain research contexts. Combining some of the variables using ratios or calculating an average appears to be rather simple in its implementation. However, additional problems may ensue as a result of such transformation since the properties of the original model are implicitly modified. This may in turn lead to certain statistical distortions.

As pointed out above, we had a case of multicollinearity affecting the three suprasegmental variables under study. Increasing the sample size or validating the study with a different sample was ruled out due to some of the many constraints that we encountered during this project. Combining some of the variables did not seem initially to be feasible, given the nature and the different scales and measurement procedures used when collecting the data. From a conceptual standpoint, all three variables are related to temporal aspects of speech production and, therefore, it was

not surprising to find high levels of correlation among them. As pointed out in Section 3.8.4.1, speech rate was initially defined in this study as the number of syllables per unit of time excluding pauses. In an attempt to combine the three suprasegmentals under study and avoid the multicollinearity issue, we decided to drop pause frequency and pause duration from the statistical analysis and use a conceptualisation of speech rate that included pauses in the different multiple-regressions. The goal was to encompass in a single measurement aspects from three different temporal variables of speech production.

Three multiple-regression analyses were conducted in this study. Tables 21, 22 and 23 present the results of the statistical analyses between the dependent and independent variables. Each table presents the subsets of independent variables (categories of phonemic errors and one suprasegmental variable) that best predict the intelligibility scores in connected speech. For each subset, the number of selected variables as well as the R square, R square adjusted and Mallow's Cp coefficients are offered. These coefficients served to evaluate the quality of each predictive model. Specifically, R square refers to the proportion of variability in the intelligibility scores that the model is capable of explaining (the closer the value to 100%, the better). R square adjusted refers to the same notion and its value results from the adjustment of R square according to the number of variables present in the model. R square adjusted is considered, therefore, more useful than the simple R square when examining a specific regression model. Mallow's Cp refers to the amount of error that each subset is unable to explain when compared to the error left unexplained in the full model-the lower its value, the better in terms of predictability of the model.

Before presenting the results of each model, it is noteworthy that we examined the relationship among all variables involved through scatter plots and correlation analyses. Furthermore, the conditions for multiple-regression, i.e. independent residuals with the same variance and a normal distribution, were checked through standardised residual plots.

Table 21 presents the multiple-regression analysis between seven predictor variables (six categories of phonemic errors and speech rate) and intelligibility scores at sentence level. Error categories 1 and 8 (affricates and nasals) were not included in the analysis due to their lack of significant correlation with intelligibility scores. R square adjusted for the full set of variables was 85.5%. However, it is subset 4 that seems to predict best the variance in intelligibility scores at this level. Subset 4

includes: phonemic errors affecting unstressed vowels, stops, vowel combinations and speech rate. It must be pointed out that the R square for speech rate was, by itself, 77.9%. Furthermore, speech rate in combination with errors affecting unstressed vowels accounts for 86.8% of the variance. This seems to reinforce the importance of vowel categories in intelligibility loss as noted in the literature on speech disorders. It is also interesting to note that error categories affecting liquids, fricatives and vowel combination seem to explain a very small percentage of variance when added to the model made up of speech rate and unstressed vowels.

Table 21: Multiple-regression (best subsets): six categories of phonemic errors, speech rate and intelligibility scores at the sentence level

Subset	**R-Sq**	**R-sq (adj)**	**Cp**
Full model (C2, C3, C4, C5, C6, C7, speech rate)	90.9	85.5	8.0
Subset 6 (C1, C2, C5, C6, C7, C8)	90.9	86.2	6.0
Subset 5 (C2, C3, C4, C6, C7, speech rate)	90.5	87.1	4.5
Subset 4 (C2, C6, C7, speech rate)	89.9	87.6	3.2
Subset 3 (C2, C6, speech rate)	89.3	87.3	2.0
Subset 2 (C2, speech rate)	88.2	86.8	1.5
Subset 1 (speech rate)	79.1	77.9	11.5

Table 22 below presents a multiple-regression analysis between five categories of phonemic errors, one suprasegmetal variable (speech rate) and the intelligibility scores at passage level. Categories 1, 4 and 8 were excluded from this model since they did not show any significant correlation with speakers' intelligibility scores. The R square adjusted for the full model is 92.2%. However, the best predictor model, as indicated by the R square adjusted coefficient, is subset number 5, which includes speech rate and categories 2, 3, 5 and 6, i.e. unstressed vowels, stressed vowels, liquids and stops. In fact, the suprasegmental factor (speech rate) explains by itself

66.2% of variance in intelligibility scores at this level of analysis. The addition to the model of one category related to vowels increases the predictability by 13.5%. Again, the suprasegmental variable under study seems to play a more important role than the segmental categories when it comes to predicting the intelligibility scores in connected speech. Furthermore, it is important to point out that the different patterns of results resemble those obtained when analysing intelligibility at sentence level.

Table 22: Multiple-regression between five categories of phonemic errors, speech rate and intelligibility scores at passage level

Subset	**R-Sq**	**R-sq (adj)**	**Cp**
Full model (C2, C3, C5, C6, C7, speech rate)	94.7	92.2	7.0
Subset 5 (C2, C3, C5, C6, speech rate)	94.7	92.7	5.0
Subset 4 (C2, C5, C6, speech rate)	92.9	91.0	7.3
Subset 3 (C3, C5, speech rate)	87.3	84.9	19
Subset 2 (C3, speech rate)	81.9	79.7	61.8
Subset 1 (speech rate)	68.0	66.2	61.8

Table 23 presents a multiple-regression analysis between the intelligibility scores in the semi-spontaneous production task (dependent variable) and six predictor factors (five categories of phonemic errors and one suprasegmental variable). Categories 1, 4 and 8 were excluded from this model, given that they did not seem to be significantly correlated with intelligibility scores. R square adjusted for the full model was 39.6%. The suprasegmental variable in this model was able to explain 45.7% of the variance in intelligibility scores. R square for the best subset of variables was 53.3%, which in fact is a higher percentage than for the full set of variables. The best subset of variables includes speech rate and error categories affecting vowels in unstressed syllables.

Table 23: Multiple-regression between five categories of phonemic errors, speech rate and intelligibility scores in the semi-spontaneous production task

Subset	**R-Sq**	**R-sq (adj)**	**Cp**
Full model (C2, C3, C5, C6, C7, speech rate)	59.7	39.6	7.0
Subset 5 (C2, C5, C6, C7, speech rate)	59.7	44.2	5.0
Subset 4 (C2, C3, C7, speech rate)	59.6	48.1	3.0
Subset 3 (C2, C3, speech rate	58.9	50.7	0.6
Subset 2 (C2, speech rate)	58.5	53.3	0.6
Subset 1 (speech rate)	48.7	45.7	0.3

Results in the multiple-regression analyses for intelligibility scores at sentence and passage levels are very similar. The suprasegmental variable in the model seems to account for the highest percentage of intelligibility scores at both levels. The best subset of variables at both levels includes speech rate plus those phonemic categories concerning the reduction of vowels in unstressed syllables. However, results in the multiple-regression for intelligibility in the semi-spontaneous production task significantly differ from the other two. In fact, the predictability of the best subset of variables is significantly lower for intelligibility at this level of analysis suggesting that other factors not included in the model should be able to predict the rest of the variability in intelligibility scores (more on this issue will be discussed in Section 4.8).

4.8 Discussion and limitations to this study

In light of the data and statistical analyses presented in our previous section, we will now address the research questions formulated in Chapter 1.

The first obvious question is to determine if there is indeed an intelligibility loss for the L2 learners involved in this study. If this was not the case, and the

intelligibility loss was not significant enough, then a special treatment of this issue, particularly from a pedagogical perspective, would not be justified. The results presented in the previous sections confirmed the existence of statistically significant differences between the native and the non-native speakers with regard to their performance in the intelligibility tests at all levels. The intelligibility loss at all four levels of analysis can be quantified at 19.59%, 13.49%, 14.31% and 21.89% respectively. These percentages seem to be slightly higher than those found in previous intelligibility studies involving L2 learners (e.g. Rogers 1997). It is, however, very difficult to establish any valid comparisons due to the different language combinations, participants' level of proficiency and the different range of variables under consideration in each study. Nevertheless, a percentage that approaches or even exceeds 20% of intelligibility loss seems to warrant, at least in our opinion, a certain degree of attention and a specific pedagogical treatment. It is noteworthy that the highest percentage of intelligibility loss occurred in the semi-spontaneous production task. It is precisely communicative effectiveness in spontaneous conversation that seems to be one of the main goals of communicative language teaching. This underlines even further the potential importance of intelligibility loss for L2 learners even within the framework provided by the communicative approach.

One of the goals of this study is the exploration of a possible correlation between certain individual differences (level of proficiency, motivation, aptitude, gender and L1) and the intelligibility scores at word level and in connected speech. Gender, operationalised from a biological point of view, did not prove to be correlated with intelligibility scores. In other words, no statistically significant differences between males and females were detected in intelligibility scores at the four levels of analysis. Results seem to confirm previous studies exploring the relationship between pronunciation accuracy and gender (see, for example, Elliott 1995 or Flege and Fletcher 1992). It is also true that, in the last decade, gender has been operationalised not from a biological but from a social constructivist approach. As Hansen Edwards (2008: 255) points out: 'when gender is framed and investigated as a social construct, it does appear to impact the level of access learners have to L2 use opportunities and therefore the ability to get L2 input and negotiate meaning, which appear to affect L2 development'. It did not seem feasible, however, to include in this study a conceptualisation of the notion of gender from a social perspective, given the nature

of our research and the instruments used for data collection. In any case, results support the idea that gender, at least when considered from a biological perspective, does not correlate with intelligibility scores at word level or in connected speech.

As far as possible advantages for the bilingual speakers (English/Polish and English/Punjabi) over the monolingual ones in terms of degree of intelligibility, it must be pointed out that previous empirical research has yielded contradictory results on this issue. On the one hand, bilingual students seem to be at an advantage when it comes to learning an additional language (see, for example, Cenoz 1991; Sanz 2000). On the other hand, when researchers have examined the acquisition of specific areas such as L3 phonology, some results point to advantages in terms of speech production or perception (e.g. Cohen *et al*. 1967; Enomoto 1994), while others could not find a statistically significant difference in performance for bilingual over monolingual learners (e.g. Werker 1986). As Cenoz (2003) notes, a wide array of factors seems to exert an influence on the possible advantages of bilingual over monolingual learners, e.g. the particular area under investigation, the level of instruction received by bilingual speakers in both languages, their level of proficiency or the typological distance between the L1, L2 and L3. It must be noted that in this study the actual level of proficiency of bilingual students in their L1 was not examined. Nevertheless, our preliminary questionnaire determined that they were all born in England and they spoke either Polish or Punjabi at home with their parents. As reported in Section 4.2, statistically significant differences were found between bilingual and monolingual learners for intelligibility scores at word, sentence and passage levels. It is also worth noting that intelligibility scores in the semi-spontaneous production task were higher for bilingual than for monolingual learners (mean percentage). However, the difference was not large enough to become statistically significant. Our results seem to contradict Werker (1986) and clearly point to an advantage of bilingual speakers even in specialised areas such as phonological acquisition. It could be hypothesised that early exposure to an additional phonological system allows for an increase in one's sensitivity in the perception of new phonetic patterns. In any case, we must be cautious when interpreting these results because the interaction of a wide range of variables makes it very difficult to reach any firm conclusions.

Another goal of this study centred on exploring the possible relationship between intelligibility scores at different levels of analysis and certain individual differences such as aptitude, motivation and level of proficiency. In the general field

of SLA, aptitude and motivation account for a significant degree of variance in learners' achievement. As noted by Dörnyei and Skehan (2003), correlations usually range between 0.20 and 0.60. Specifically, as far as the acquisition of L2 phonology is concerned, certain studies have linked degree of motivation to pronunciation accuracy or perceived degree of foreign accent (e.g. Moyer 2007), while others have only been able to show a very low correlation between both variables (Flege *et al*. 1995) or even no correlation whatsoever (Thompson 1991). In the case of aptitude, both aptitude for oral mimicry and working memory capacity or phonological working memory have been linked to pronunciation accuracy (e.g. Flege *et al*. 1999; Ioup *et al.* 1994; Purcell and Suter 1980). In this study, motivation was operationalised as students' scores on a motivation questionnaire, while aptitude was quantified as a combination of scores from a battery of tests that served to measure our participants' working memory capacity, as well as from a test that was designed to measure their ability for oral mimicry. Level of proficiency was operationalised as participants' scores on the *Instituto Cervantes*' DELE exam. Level of proficiency was highly correlated with intelligibility at all levels (Pearson's r was above 0.6). In the case of aptitude and motivation, a significant correlation with intelligibility scores was found at the word, sentence and passage levels. The correlation was not statistically significant in the semi-spontaneous task even though Pearson's r was above 0.2. In general, the correlation coefficient of the three variables (aptitude, motivation and level of proficiency) and intelligibility scores was higher than in previous studies, where the variables involved were pronunciation accuracy or degree of perceived foreign accent. It must also be noted that there was a significant level of intercorrelation among the three variables. This fact prevented us from running a multiple-regression analysis that would have quantified the effect of each individual variable on degree of intelligibility. In any case, we did expect a high correlation between level of proficiency and intelligibility in semi-spontaneous production, given that it is precisely at this level where a higher degree of proficiency could ensure more accurate speech production. This, in turn, could compensate for a possible decrease in attention to form (Tarone 1978, 1979). It is indeed possible that higher levels of proficiency may increase the likelihood of allocating a bigger number of cognitive resources away from the syntactic or semantic level and onto the phonological sphere. As far as aptitude is concerned, it is not surprising to find a significant correlation between degree of intelligibility and ability for oral mimicry, as well as working

memory capacity. It is, in fact, quite consistent with those studies that have explored the impact of aptitude and L2 phonological acquisition (e.g. Flege *et al*. 1999; Purcell and Suter 1980). With regard to motivation, the level of correlation yielded by this study seems to be consistent with the coefficients found in other areas of SLA (see Ellis 2005). To summarise, it is very difficult to deny the relationship between speech intelligibility and the individual differences that have been treated in this study. However, the existence of intercorrelation among the different variables makes it difficult to quantify the exact influence of each factor on intelligibility scores.

One of our initial research questions centred on exploring the possible relation between intelligibility scores at word level and in connected speech. A strong correlation between both levels would validate the use of a single word intelligibility test as a potential diagnostic tool in assessing L2 learners' degree of intelligibility. In the literature on speech disorders, it is not unusual to find that one level of analysis is able to predict intelligibility scores at another (Yorkston and Beukelman 1978). Results in this study seem to support this idea, as a high level of correlation was found between intelligibility at the word, sentence, passage and spontaneous conversation levels. The highest correlation (Spearman's r above 0.8) was recorded between intelligibility at sentence and passage levels. Correlation with the semi-spontaneous production task was above 0.6. If we look at the overall scores for each level of analysis, the highest degree of intelligibility was reached in connected speech. In fact, very similar mean percentage scores were recorded at sentence, passage and in the semi-spontaneous task. It is very likely that the role of contextual information is responsible for the score increase in connected speech. As pointed out in our review of the literature, intelligibility scores are affected by the type of stimuli, their linguistic complexity and their presentation in an isolated manner or within higher units of discourse. In general terms, research indicates that intelligibility scores are higher in connected speech than at word level (Kempler and Van Lancker 2002; Santos Barreto and Zazo Ortiz 2010).

This leads us to the examination of the possible influence of stylistic variation on intelligibility scores, i.e. does the amount of attention L2 learners pay to speech production have any effects on their degree of intelligibility? As noted in Chapter 2, empirical research in L2 phonology seems to indicate that attention to form conditions the production of target-like phonological features: a more formal style, which translates into more attention to form, shows a higher production of target-like

features, while a more spontaneous style, with less attention to form, results in the production of a smaller proportion of target-like features. In our case, intelligibility scores in the semi-spontaneous production task did not result in lower scores when compared to word, sentence and passage scores. It is very likely that two competing forces were in action here: (i) the contextual information provided to listeners in connected speech and (ii) a possible decrease in target-like productions by L2 learners as a result of stylistic variation. In this study, the former seems to have prevailed over the latter. In fact, certain effects derived from stylistic variation may account for the differences recorded during the multiple-regression analyses. We will expand on this idea below.

One of the main goals of this study was to analyse the potential impact of segmental and suprasegmental deviations on speech intelligibility at word level and in connected speech. Specifically, we aimed at assessing the impact of eight categories of phonemic errors on intelligibility scores at word level. Moreover, we also explored the possibility for those same categories of phonemic errors of predicting variability in intelligibility scores in connected speech. In addition, the contribution of three suprasegmental variables (speech rate, pause frequency and pause duration) was also examined. Correlation and multiple-regression analyses served to identify those subsets of segmental and suprasegmental variables that best predicted variances in intelligibility scores. As noted in Chapter 2, empirical research has yielded inconclusive results on the possible impact of segmental and suprasegmental deviations on speech intelligibility. While some studies point to the impact of segmental errors (mainly vowels) on degree of intelligibility (e.g. Bent *et al.* 2007; Rogers 1997; Whitehill 1997), others have revealed weak correlations between intelligibility or comprehensibility scores and segmental deviations (e.g. Ertmer 2010). In terms of the importance of suprasegmental variables, duration (e.g. Tajima, Port and Dalby 1997) and intonation (e.g. Crocker 2010) have proven to be correlated with degree of speech intelligibility. However, the importance of these variables has not been fully validated by further empirical studies (e.g. Holm 2008). In one of the few studies involving English learners of Spanish, Schairer (1992) found high correlations between comprehensibility and overall vowel production (r = .92), consonant production (r = .62) and consonant linkage (r .84). Speech rate did not show any significant correlation. It is also important to emphasise once again that correlation does not imply causation. In other words, we cannot establish a causal link

between the segmental or suprasegmental variables under study and the reduction in intelligibility scores. However, it is also true that the patterns of correlation are strong enough to suggest, at least, a connection between both dimensions.

In this study, the single word intelligibility test revealed that unstressed vowels (30%) and liquids (27.66%) presented the highest percentage of errors[51], while nasals (14.82%), stops (14.64%) and affricates (14.15%) offered the least amount of difficulty for our group of speakers. Moreover, following Rogers (1997), we attempted to validate the single word intelligibility test by correlating the scores on the eight categories of phonemic errors from the single word test with speakers' results in the intelligibility tests in connected speech. Correlation analyses suggest that certain categories of phonemic errors only exhibited a moderate degree of correlation with intelligibility in connected speech. This was, in fact, the case for categories number 1, 3, 5, 7 and 8 at the sentence level, and 7 and 8 at the passage level. A high degree of correlation was shown between intelligibility at the sentence level and one specific category of phonemic errors: contrasts based on the reduction of unstressed vowels ($r = .70$). As far as passage scores are concerned, a high level of correlation was again found with regard to contrasts based on the reduction of unstressed vowels, stressed vowels and liquids. Only five categories of phonemic contrasts showed significant correlations with intelligibility scores in the semi-spontaneous production task (contrasts 2, 3, 5, 6, and 8). Furthermore, those statistically significant correlations were only moderate (between $r = 0.46$ and 0.66), especially when compared to the correlations recorded at sentence and passage levels. If we examine the correlation coefficients across all three levels of analysis (sentence, passage and semi-spontaneous production), it is noteworthy that category 4 (fricatives) did not show any significant correlations at any of the levels under study, while category 2, affecting the reduction of unstressed vowels, was highly correlated with intelligibility scores at all levels.

If we turn now our attention to the segmental and suprasegmental error analyses in the sentence, passage and semi-spontaneous tasks, one can certainly point out that there is a great similarity in the results obtained in the sentence and passage tasks. This seems to confirm previous studies both in the field of speech disorders and

[51] In this case, an 'error' must be understood as those instances incorrectly identified by our group of sixty evaluators in the single word intelligibility test.

foreign-accented speech (e.g. Rogers 1997). Nevertheless, comparisons between different studies, as pointed out on several occasions, are very difficult to make due to the different methodological choices, language combinations, populations of participants, etc. In the sentence and passage tasks, the suprasegmental variable under study (speech rate) accounted for most of the variance (77.9% at the sentence level and 66.2% at passage). It is also important to point out that the three suprasegmental variables measured in this study were highly correlated with intelligibility scores in connected speech. Furthermore, a high degree of intercorrelation was also found among the same variables. As pointed out in Section 4.4.2, this resulted in multicollinearity and led us to reformulate our definition of speech rate and only include this variable in the subsequent multiple-regression analysis. The predictive ability of the eight phonemic categories and speech rate is very similar for intelligibility scores at sentence and passage levels (over 80%). The best possible subset of variables for intelligibility scores at the sentence level included speech rate, errors affecting the reduction of unstressed vowels and vowel combinations. At passage level, the best possible subset of variables included speech rate, errors affecting stressed vowels, unstressed vowels, liquids and stops. At both levels of analysis, speech rate and phonemic errors affecting stressed and unstressed vowels seem to be responsible for most of the variance in intelligibility scores.

In the semi-spontaneous production task, only five categories of phonemic errors were significantly correlated with intelligibility scores: errors affecting unstressed vowels, stressed vowels, liquids, stops and vowel combinations. Even though intelligibility scores were very similar to those obtained at the sentence and passage levels, the predictability of the model was considerably lower. At his level of analysis, the suprasegmental variable still accounts for most of the variance (45.7%). However, the best possible subset of variables (speech rate and contrasts affecting the reduction of unstressed vowels) explains 53.5% of scores. Given that intelligibility scores have differed in the semi-spontaneous production task when compared to scores at sentence and passage level, one must conclude that the source of intelligibility loss at this level of analysis arises from other variables not included in our study. A possible explanation could stem from variation due to style shifting, i.e., participants paid less attention to form in spontaneous conversation and focused their efforts on conveying a communicative message. As a result of this, there could be an increase in lexical or grammatical inaccuracies. Nevertheless, the importance of

speech rate in combination with the category of phonemic errors involving the reduction of vowels in unstressed syllables, suggests that intelligibility, at this level of analysis, may be more related to issues of fluency than to mere segmental deviations. The exact quantification of lexical and grammatical errors and their impact on speech intelligibility could certainly represent an avenue for further research, as we will argue in Chapter 5. Fluency is certainly an important notion in determining a speaker's communicative ability. Very often, this concept refers to elements that are difficult to measure such as pronunciation, intonation, speed, naturalness, etc. The majority of studies have focused on the analysis of certain linguistic and paralinguistic components, namely, quantifiable aspects such as pauses, repetitions and self-corrections (Ejzenberg 2000) or the importance of non-linguistic elements such as gestures and facial expressions (Bavelas 2000). Some studies have also explored the criteria used by judges in the assessment of fluency (e.g. Freed 2000). Others have attempted to analyse the social and cultural aspects associated with the notion of fluency. An additional research trend has focused on the cognitive implications that the notion of fluency has in psycholinguistic studies, specifically in terms of the possible identification of constraints in the human capacity for speech planning (Segalowitz 2000). Certain pedagogical implications derived from the impact of fluency on intelligibility will be discussed in Chapter 5.

In assessing the potential limitations of this study, we will refer to four notions that are commonly used to evaluate research quality: construct validity, internal validity, external validity and reliability (Yin 2003). Brown (1988: 29) defines validity as 'the degree to which a study and its results correctly lead to, or support, exactly what is claimed'. Specifically, construct validity refers to whether the instruments of data collection provide an accurate measure of the theoretical constructs which they purport to identify. This can be especially problematic in the case of speech intelligibility, given the difficulty involved in defining this type of construct. In our study, as a consequence of utilising quantitative instruments, the measurement of speech intelligibility has been undertaken at word level and in connected speech (sentence, passage and semi-spontaneous production). As pointed out in Chapter 3, the stimuli used at sentence and passage level were phonetically balanced. This was not the case for the speech samples obtained in the semi-spontaneous production task and could partially account for the results yielded by the multiple-regression analysis involving intelligibility scores at this level. In spite of the

difficulties that arise when attempting to use spontaneous or semi-spontaneous speech samples, we decided to include such samples in this study, as this was precisely one of the research gaps encountered in our review of the literature. While we aimed at introducing as much control as possible over the selection of lexical items for the semi-spontaneous task, it must be acknowledged that a certain degree of variability is unavoidable. On the other hand, each particular instrument used in this study presents limitations in itself. This certainly applies to the instruments used to collect information on students' individual differences. Aptitude was measured through a working memory capacity task and an oral mimicry test. Even though both components have proven to be related to aptitude for pronunciation in an L2, additional instruments of data collection could have provided a more thorough picture of this construct. Motivation, on the other hand, was measured through a purely quantitative questionnaire. The use of a quantitative instrument of data collection presents certain intrinsic limitations, which may result in disregarding more dynamic conceptualisations of the notion of motivation in language learning (Piller 2002). In addition, there are indeed other individual and cognitive differences that may influence the acquisition of Second Language phonology. As in any other research project, we strived to find a balance between the use of accurate and thorough instruments of data collection and the physical and temporal constraints associated with this type of research.

'Internal validity refers to whether the results of a study are due solely to those variables being identified and compared within the study' (Brown 1988: 40). It is noteworthy that only eight categories of phonemic errors and three suprasegmental variables were taken into account in the analysis of intelligibility at word level and in connected speech. Furthermore, the influence of grammatical or lexical accuracy was not taken into consideration. Moreover, potential listener-related variables that are known to influence intelligibility scores were not examined. Nevertheless, an attempt was made to exercise a certain level of control in the selection of our group of evaluators. First, our selection consisted exclusively of a group of native Spanish speakers with a very limited knowledge of English and, second, their contact with native English speakers had been very limited. On a purely statistical level, the presence of multicollinearity among the three suprasegmental variables forced us to redefine our notion of speech rate and prevented us from including all these factors in the multiple-regression models.

It is also important to point out the existence of potential limitations arising as a consequence of using written material in the elicitation tasks. The influence of orthographic input on L2 pronunciation has been widely acknowledged in the literature (see, for example, Bassetti 2009). We certainly recognise that certain instances of non-targetlike pronunciations may be due to orthographic influence. Nevertheless, this dissertation does not focus on the reasons behind the segmental and suprasegmental deviations exhibited by our group of participants; it merely centres on examining their potential impact on speech intelligibility. It is true that we initially considered the use of alternative tasks such a delayed repetition test. However, given that this study centres on students' L2 speech production and not on their level of perception in the L2, it was subsequently decided to use the reading of written material in the assessment of intelligibility at all levels of analysis. Furthermore, our participants' ability for oral mimicry had already been assessed through the use of an oral mimicry task (see Section 3.7.2.3).

Campbell and Stanley (1966) use the term 'external validity' to refer to the extrapolation of the results of an investigation to other individuals, groups, contexts and situations. External validity is usually associated to the idea of statistical and representative sampling. First of all, it must be pointed out that our study focuses on a mixed-ability class of Key Stage 4 learners of Spanish. As far as statistical generalisation is concerned, it must be acknowledged that a sample size of 20 participants presents a limited potential for generalisation. In fact, some consider it to be less than the desirable amount of cases to carry out a multiple-regression analysis. Regression models can be validated by using another set of data with similar characteristics and drawn from the same population. Alternatively, when working with large samples, the sample is divided randomly into two groups that can be used to devise two models for comparison. In this study, we did not have the opportunity of validating the results with a second sample. Furthermore, the battery of tests that participants and evaluators had to go through is of such an intensive nature that it would have been very difficult to carry out the same type of analysis with a larger sample of the population in question. In any case, results seem to be robust enough to, at least, warrant the generalisation of certain patterns to the entire population under study.

Reliability demonstrates that the operations of a study can be repeated with the same results. It is linked to the notion of quality and consistency in data measurement.

Therefore, the main goal should always be to minimise errors and biases in the study. It is important to determine a protocol in order to specify the steps to follow in the research process and to keep a database with all the information collected. In our case, an initial protocol was established and a database with the results of each intelligibility test was kept throughout this study. It must also be borne in mind that additional limitations may result from the fact that students were recorded in a quiet room and not in a soundproof laboratory. The same applies to the listening sessions where evaluators had to perform the different transcription tasks. Moreover, the recording sessions with each student did not occur at the same time of day and, therefore, an array of additional factors could have affected each student's performance in the intelligibility tasks.

It is also important to emphasise the idea that correlation does not mean causation. This study has correlated certain categories of phonemic errors, as well as certain temporal variables of speech, with an intelligibility loss in our group of English learners of Spanish. Even though some of the correlations' coefficients obtained in the statistical analyses are significantly high, we cannot in any way infer that the intelligibility loss is caused by the aforementioned segmental and suprasegmental factors. One can certainly appreciate a relationship between intelligibility scores and some of the categories of phonemic errors and, especially, between intelligibility values in connected speech and the three temporal aspects of speech analysed here. It is likely that the variables under study are only partially responsible for our participants' degree of speech intelligibility. This seems to be the case for intelligibility results in the semi-spontaneous production task, where our multiple-regression analysis model could not account for the majority of variance in intelligibility scores.

Chapter 5: Summary, Implications and Conclusions

5.1 Summary of this study

This study has offered an assessment of the non-native speech intelligibility of a group of English learners of Spanish at different levels of analysis (word, sentence, passage and in a semi-spontaneous production task). Specifically, we aimed at analysing the impact of certain categories of phonemic errors and some temporal aspects of speech (pause frequency, pause duration and speech rate) on intelligibility scores. In addition, the possible correlation between degree of intelligibility and certain individual factors (gender, level of proficiency, motivation, aptitude and L1) was also examined.

Our review of the literature centred on the notion of speech intelligibility (see 2.1), as well as on the methodological challenges involved in offering a precise measurement of intelligibility loss (see 2.4). We explored the wide array of variables that seem to exert an influence on intelligibility measurements (see 2.3). In addition, we reviewed the scarce number of studies that have focused on speech intelligibility within the area of Second Language Acquisition (see 2.4.2). This paucity of research becomes even more evident when we look outside the field of English as a second/foreign language. Consequently, the possible applications of this type of research onto the field of pronunciation instruction have been rather limited thus far.

Sixty evaluators, native speakers of Peninsular Spanish, transcribed different speech samples belonging to a group of 20 Key Stage 4 English learners of Spanish. The elicitation of the speech samples served to assess intelligibility at word level (through the use of a single word intelligibility test), at the sentence and passage levels (using the Spanish version of the Harvard Psychoacoustic Sentences and the reading of a phonetically-balanced text) and in a semi-spontaneous production task (see Chapter 3 for a description of participants, evaluators and general methodological considerations). Results revealed an intelligibility loss for the subjects under study both at word level and in connected speech (see 4.2). Moreover, the intelligibility scores in the single word test were highly correlated with results in the sentence, passage and semi-spontaneous production tasks (see 4.6). In addition, participants' results on eight categories of phonemic errors plus one suprasegmental factor (speech

rate) were able to predict over 90% of variance in intelligibility scores at sentence and passage level. In both cases, speech rate accounted for most of the variance (77.9% and 66.2% respectively). It is also noteworthy that the other two temporal variables of speech studied here, i.e. pause frequency and pause duration, were not included in the different multiple-regression analyses due to the appearance of multicollinearity. When we turn to intelligibility scores in the semi-spontaneous production task, the predictability of the model is significantly lower (53.3%). Thus, it can be suggested that additional variables (syntactic, lexical, grammatical accuracy etc.) play a significant role in predicting degree of speech intelligibility at this level of analysis. In any case, the high correlation of the temporal variables of speech, as well as the high predictability of 'speech rate', when entered into the different multiple-regression analyses, seem to underscore the importance of suprasegmentals over segmentals as far as intelligibility loss is concerned. Moreover, it is argued here that the notion of fluency, as a concept intrinsically linked to the suprasegmental variables examined in this study, is central in explaining instances of intelligibility loss.

With regard to the role played by individual differences (see 4.1), it seems clear that gender did not exert any influence on intelligibility scores. Aptitude and motivation, on the other hand, were significantly correlated with intelligibility at all levels of analysis. It is also noteworthy that the bilingual students (Polish/Punjabi) performed significantly better than monolingual native English speakers both at word level and in connected speech.

5.2 Pedagogical implications

Our introductory chapter highlighted the importance of the *intelligibility principle* for pronunciation instruction. Section 1.2 also served to examine the major methodological trends with regard to the teaching of L2 pronunciation. Chapter 1 provided therefore a general background in which to frame any pedagogical intervention targeting segmental or suprasegmental features. We will now attempt to derive some specific pedagogical insights from the empirical results analysed in Chapter 4.

5.2.1 Individual differences and intelligibility: pedagogical implications

As pointed out in Chapter 2, research on individual variables reveals that L2 phonological acquisition is not solely a linguistic matter, but rather, there are a number of extralinguistic aspects that seem to play a major role in the learning process. Our empirical results confirm the importance of motivation, aptitude and L1 background (bilingual versus monolingual speakers) for speech intelligibility. A high degree of correlation was found between the aforementioned variables and intelligibility scores both at word level and in connected speech. Results seem to confirm that degree of speech intelligibility benefits from a high level of motivation and a high degree of aptitude. In addition, amount of L2 use, amount of L2 input and age at the onset of the acquisition process are also variables that, according to the literature available (see Chapter 2), have proven to be important in the process of L2 phonological acquisition. In this study, aptitude was operationalised as a combination of ability for oral mimicry and working memory capacity. It is noteworthy that there are very few proposals for the treatment of working memory capacity in the foreign language classroom. It seems logical to assume that working memory capacity may play an important role in communicative teaching contexts where students must process a large amount of oral information. Working memory may also play an important role in processing written and audiovisual material. In addition, it has also been posited that the relationship between L2 competence and working memory capacity is based on the idea that an increased capacity of the latter will allow L2 learners to release a series of cognitive resources that would otherwise be engaged in input processing. On this basis, one can assume that exercises that increase overall L2 fluency could compensate for a low working memory capacity. Priority should be given to classroom activities that will help improve the automatic recognition of words. Moreover, we believe it extremely useful to offer learners enough opportunities to process, reproduce and automate the production of words and structures. Additionally, we could also make use of activities to improve strategies aimed at linking pre-existing L1 and L2 knowledge to new L2 input. If we increase the use of what we might call ‘long-term knowledge’, the burden on working memory will be drastically reduced. Other techniques that may serve to compensate for a low working memory capacity include the use of visual aids. Visual input can be used to

replace working memory's temporary storage function allowing for the processing of a greater amount of information. It would be advisable for teachers to increase their learners' use of written and visual aids during their communicative interactions, especially in the early stages of the learning process. Our knowledge of the role of working memory in foreign language learning is still relatively small. Further research is certainly needed to determine whether teaching techniques that focus on increasing working memory capacity can have a positive impact on L2 learning. We also need more research to determine whether educational intervention techniques aimed at compensating for a reduced working memory capacity can actually improve the learning process. The use of exercises designed to improve students' degree of fluency or the use of visual and written aids are just some of the pedagogical tools that can lead to an improvement in L2 phonological acquisition. All these suggestions will need empirical validation and should offer possible avenues for further research in the field.

With regard to motivation, L2 learners should be encouraged to develop positive attitudes towards pronunciation. It is important that they become aware of the positive effects that an improved pronunciation may have on communication. The more instructors are aware of these variables, the better they will be able to anticipate the level of proficiency their students may reach. In other words, they will be able to predict, with a reasonable degree of reliability, their students' progress in L2 phonological acquisition.

5.2.2 Implications in terms of curriculum sequencing

It is widely accepted that those elements that play a major role in reaching the 'intelligibility threshold' should be considered a priority in any pedagogical treatment. As Kenworthy (1996) notes, difficulties affecting intelligibility should become a priority for pronunciation instruction, while the treatment of those features that merely characterise foreign-accented speech but do not result in intelligibility loss are merely optional. One of the problems is that those difficulties or deviations that result in intelligibility loss have not been properly identified by the empirical research available thus far (see e.g. Derwing and Munro 2009 and our Chapter 2). Our own correlational studies between speech intelligibility and accurate production of English

learners of Spanish seem to reveal two general conclusions regarding the nature of the interaction between speech intelligibility and segmental and temporal aspects of L2 speech: First, temporal/prosodic aspects of speech, at least the ones examined in this study, seem to be more strongly correlated with intelligibility scores than mere segmental deviances. These temporal aspects should therefore play a central role in any instructional programme. Secondly, in spite of the apparent simplicity of the Spanish vowel system, deviations affecting vowel combinations and vowels are more critical for intelligibility scores than errors affecting consonant production. It is important to note that these results apply to the patterns of error exhibited by our class of Key Stage 4 learners of Spanish. In other words, patterns of errors and intelligibility scores may differ for learners with, for example, a different level of proficiency in Spanish. The priority of suprasegmental elements has been emphasised by numerous authors (e.g. Cantero 2003; Gil 2007; Kenworthy 1996; Llisterri 2003a). In fact, Dalton and Seidlhofer (1994) note that, from a pedagogical perspective, the easiest area to teach, i.e. segmental elements, is the least relevant for communication. The truth is that both segmental and suprasegmental elements matter. Suprasegmentals are certainly important in order to integrate individual sounds in the speech chain, interpret the message appropriately and reach a satisfactory level of fluency. Stress and rhythm are a high priority when the L1 and L2 exhibit different patterns. In these cases, it is enormously helpful for students to be aware of the existence of such rhythmic changes.

The empirical results of this dissertation have shown a high level of correlation between the temporal aspects under study and intelligibility scores in connected speech. Thus, it might be warranted to advocate a sequence of contents that prioritises the suprasegmental level as far as pronunciation instruction is concerned. This methodological decision supports the view of authors such as Cantero (2003), Gil (2007)[52] and Listerri (2003a) when positing that suprasegmental elements, i.e. stress, rhythm and intonation, are the most important in ensuring the effectiveness of

[52] Gil (2007: 161) proposes the following hierarchy of contents with regard to pronunciation instruction:

1. Base de articulación general
2. Aspectos suprasegmentales: acento, ritmo y entonación
3. Realización fonética de los fonemas vocálicos y consonánticos: alófonos y variantes
4. Cuestiones de detalle fonético: coarticulación
5. Grado de fluidez

the communicative act. Based on our empirical results, it is argued here that suprasegmental factors and the notion of fluency are at the cornerstone of intelligibility loss for this specific population of L2 learners under study. The pedagogical treatment of the notion of fluency will be addressed in Section 5.3. It is therefore important to emphasise that any treatment of specific segmental elements will have to be placed within a more general pedagogical framework that aims at improving L2 learners' level of fluency. Bearing this in mind, we can point out that our empirical results, for both isolated word level and in connected speech, stress the importance of vowels over consonants and, specifically, the relevance of the phenomenon of reduction of vowels in unstressed syllables. In this respect, one can say that the main difficulties encountered by English speakers derive more from the negative influence of their L1 system than from the intrinsic complexity of the Spanish vowel system. With regard to consonants, their importance for the level of speech intelligibility of the participants under study has proven to be rather limited. In fact, the importance of consonant elements seems to decline when the task at hand requires a higher degree of spontaneous production. In any case, insightful techniques of phonetic correction, as well as pedagogical suggestions aimed at specific segmental elements can be found in Llisterri (2003a) or Gil (2007) and will not be repeated here. Suffice it to say that any pedagogical intervention targeted at specific segmental elements will not be successful unless it is properly framed within a much broader treatment of fluency and prosodic elements. An emphasis on prosodic elements and fluency is precisely what, according to the empirical part of this dissertation, will have the highest impact on the degree of speech intelligibility for our population of L2 learners.

5.2.3 Fluency and its pedagogical treatment

Based on our empirical results, it is argued here that intelligibility, at least for this language combination, level of proficiency and population of L2 learners, is more highly correlated with level of fluency than with accurate production of segmental elements in connected speech. The high correlation of speech rate, pause frequency and pause duration with intelligibility scores seems to point in this direction.

We believe that any suggestions with regard to the treatment of suprasegmentals need to be framed within a more general pedagogical approach for the teaching of fluency. As we have seen in Chapter 2, researchers have found it difficult to come to terms with a widely accepted definition of spoken fluency. We are going to adopt here some of the ideas put forward by Guillot (1999) and McCarthy (2009), as a possible working framework whereby specific pedagogical interventions related to suprasegmentals and segmentals could be understood. In this sense, the idea of 'awareness-raising' will play an essential part in the development of this framework. In other words, L2 learners need to become aware, through the systematic study of spoken data, of their own individual patterns of fluency. A systematic study of spoken data through the use of certain techniques derived from discourse analysis not only promotes the student's autonomy in their own learning process, but also as Guillot (1999: 61) puts it:

> It can facilitate the emergence of individual paradigms of fluency, enable students to identify the features and strategies of greatest relevance to them as learners and communicators, and concurrently, help them to exploit both their strengths and weaknesses more efficiently.

By drawing on discourse and conversation analysis, L2 learners are able to examine a wide variety of spoken data and embark on a journey of discovery of their own verbal interactions. They build on what they already know and, thanks to an inductive approach in the analysis of spoken data, are able to discover the features of verbal communication (see Guillot 1999 for further pedagogical suggestions and specific activities on the treatment of fluency in the foreign language classroom). Obviously, the different strategies used by instructors to implement the ideas posited in our general pedagogical framework are varied and depend largely on the imagination, experience and creativity of the teacher in question. In general terms, specific actions aimed at the development of learners' oral fluency need to take into consideration the use of appropriate linguistic models, the development of receptive skills and the role of the teacher in the L2 classroom. One of the most important goals centres, as we have already pointed out, on raising the learner's awareness with regard to the structure and specific "sound" of the target language. Thanks to the extrapolation of real communicative situations into the foreign language classroom, students are able to develop the strategies needed to communicate in an effective manner. In addition,

teachers should encourage group work, given that it could serve to reproduce many of the facets that characterise genuine situations of oral interaction. It has also been shown that group activities can provide a clear opportunity for negotiation of meaning. Another advantage of this type of activities centres on the promotion of the learner's autonomy. In this sense, students should take ownership of their own learning, while teachers need to act as facilitators and sometimes as partners in communicative activities.

Communication strategies are also essential to the notion of fluency and could represent a potential goal for classroom instruction. Communication strategies are restricted almost exclusively to oral production and are used by L2 learners to solve specific communication problems, e.g. paraphrasing, avoidance, restructuring, code-switching, literal translation, etc. Some authors (e.g. Manchón Ruiz 1985) distinguish between communication strategies based on the L1, such as literal translations, and those based on the target language, such as the formation of new words, or the use of paraphrasing. It is assumed here that strategies are mechanisms that serve to resolve a communicative problem. The presentation of a pedagogical proposal for the treatment of communication strategies in the L2 classroom goes beyond the scope of this dissertation. However, it is important to point out that an appropriate integration of this issue in the foreign language classroom could result in an improvement in one's level of fluency and, therefore, an increase in one's level of speech intelligibility.

As far as the specific treatment of temporal aspects of speech is concerned, it has already been noted that L2 learners pause more often than L1 speakers in connected speech. From a pedagogical point of view, it would seem advisable for instructors to introduce activities and tasks that focus on the perception and production of pauses both under control and spontaneous conditions. This should enable L2 learners to become more accustomed to identify and acquire the skill of using pauses in places that are acceptable and admissible in the target language. We would, therefore, argue for the introduction of activities that focus learners' attention on pause placement and underline the distinctive value that some of those pauses may possess during the communicative exchange. With regard to the phenomenon of speech rate, we recommend the introduction of tasks aimed at increasing the sensitivity of L2 learners in relation to the temporal structure of the target language. They should be made aware of the close relation between speech rate and the transmission of specific emotions and attitudes, e.g. the expression of anger may

result in a higher speech rate. From a pedagogical point of view, it is also important for instructors to set the goal of achieving an appropriate acquisition of the L2 rhythmic patterns. According to Renard (1979), the assimilation of these patterns is initially difficult because L1 habits in the suprasegmental area are firmly established. As far as English learners of Spanish are concerned, they need to be made aware that Spanish vowels do not modify their quality and duration in unstressed syllables, at least not to the same extent as their English counterparts. Learners should be encouraged to approximate the rhythmic, intonation and temporal features of the L2[53].

To summarise, this section has examined the potential implications of our research in terms of individual differences and curriculum sequencing. More importantly, we have stressed here the importance played by temporal prosodic elements in the degree of speech intelligibility of the population of participants under study. It has also been argued that any pedagogical treatment of speech intelligibility in the L2 classroom needs to be framed within a more general pedagogical approach towards spoken fluency. Some considerations for the treatment of fluency based on Guillot's (1999) ideas of using a systematic approach to the assessment of spoken data by drawing on discourse and conversation analysis have also been presented.

5.3 Contributions of this dissertation

In spite of the fact that intelligibility has been widely posited as the goal for pronunciation instruction, the number of studies analysing specific factors affecting L2 speech intelligibility is rather scarce, especially when compared to other areas of SLA (see Chapter 2). This dissertation offers new empirical data in an attempt to fill a research gap for a specific population of L2 learners receiving formal instruction within a certain educational setting.

One major contribution of this study focuses on offering an objective measurement of speech intelligibility at word level and in connected speech. The high correlation between intelligibility scores at both levels of analysis seems to validate

[53] See McNerney and Mendelson (1992) for specific examples of activities aimed at treating issues relating to pause frequency and pause duration. Activities dealing with issues related to speech rate can be found in Gil (2007). See Celce-Murcia *et al*. (1996), Gil (2007) and Terrell (1989) for examples of activities related to vowel sounds. See Odisho (2003) and Pica (1984) for activities dealing with consonant sounds. Numerous examples of activities addressing suprasegmentals, segmentals, as well as issues related to phonetic setting can be found in Gil (2007).

the usefulness of utilising a single word intelligibility test in the measurement of speech intelligibility.

Another major goal has centred on showing the strong relationship between individual differences and speech intelligibility. All individual differences under study with the exception of gender have proven to be highly correlated with speech intelligibility. This fact underlines the possible relevance of individual-related factors in any potential pedagogical treatment of this issue. The superior performance at all levels of analysis of bilingual (Polish/Punjabi speakers) over monolingual English speakers is especially significant. In addition, we have stressed here the importance of suprasegmentals over segmentals, as well as the relevance of the notion of fluency for speech intelligibility. At a segmental level, deviations affecting vowels seem to play a more important role than inaccuracies in consonants both at word level and in connected speech. The predictability of our multiple-regression models has been high for speech samples obtained at sentence and passage levels. However, multiple-regression models for intelligibility in semi-spontaneous production exhibited a more limited capability in predicting variation in students' scores. Results suggest the existence of additional variables affecting intelligibility at this level of analysis.

This dissertation has also attempted to offer some limited, but we believe very necessary, pedagogical insights for the treatment of speech intelligibility in the foreign language classroom. It has been argued that any successful instructional treatment of speech intelligibility will depend on an appropriate integration of prosodic elements and spoken fluency within the time devoted to this issue in the L2 classroom.

5.4 Suggestions for further research

In the general area of L2 phonological acquisition, studies have, for the most part, focused on the acquisition of specific phonological elements. It is for this reason that one needs to be cautious when evaluating empirical results. Additional research is certainly needed in order to assess the importance of suprasegmentals, given that they seem to play a major role in speech production and perception. Moreover, researchers seem to have focused their efforts on highly controlled contexts and a small number of participants. This makes very difficult any attempt to generalise their findings to

other populations of L2 speakers. Furthermore, the majority of researchers is based in English-speaking countries and carries on their work within the area of English as a Second Language.

If we turn to the area of L2 speech intelligibility, the amount of research devoted to this type of studies within the area of SLA is rather scarce. Further research should be able to provide a more thorough picture of the variables involved in intelligibility loss. We also need further intelligibility studies that focus on different language combinations. Moreover, researchers should increase the number of studies with participants receiving formal instruction at different levels of proficiency. Special attention should also be devoted to the relation between intelligibility at word level and in connected speech. In particular, it seems essential to gain a better understanding of the multiple factors that have an impact on intelligibility in spontaneous conversation. It is noteworthy that while our selection of segmental and suprasegmental variables was highly correlated with intelligibility scores at sentence and passage level, the predictability of our multiple-regression model was significantly lower for intelligibility scores obtained in the semi-spontaneous production task. Therefore, further exploration at this level of analysis seems to be warranted. More attempts should be made to bridge the gap between purely empirical research and possible pedagogical applications in the foreign language classroom. It is now essential to determine if instructional programmes focusing on those elements identified by researchers as responsible for higher levels of intelligibility loss result in an improvement of speech intelligibility for the specific population of L2 learners under study. This is especially relevant, given that the 'intelligibility principle' has been widely adopted as the main teaching goal by current approaches to pronunciation instruction. From a methodological point of view, we need more studies that make use of acoustic modifications of participants' speech samples at a suprasegmental level. This type of technique could offer new insights for speech intelligibility. Furthermore, the inclusion of additional individual differences, as well as the impact of task and environment-related variables, will need to be addressed under different contextual circumstances. Another interesting avenue for further research concerns the role of the listener for speech intelligibility. In the last few years, some studies have started to explore this area of research (see Chapter 2). It is expected that this strand of research will serve to complete our picture of non-native speech intelligibility.

Bibliography

Abercrombie, D. (1967) *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Abercrombie, D. (1968) Paralanguage. *British Journal of Disorders of Communication* 3: 55-59.

Abrahamsson, N. and Hyltenstam, K. (2009) Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning* 59: 249-306.

Alarcos, E. (1965) *Fonología española*. Madrid: Gredos.

Allen, C., Nikolopoulos, T. P., Dyar, D. and O' Donoghue G. M. (2001) Reliability of a rating scale for measuring speech intelligibility after pediatric cochlear implantation. *Journal of Otology and Neurotology* 22, 5: 631-633.

Andersen, R. (1983) Transfer to somewhere. In S. Gass and L. Selinker (eds.) *Language Transfer in Language Learning.* Rowley, MA: Newbury House. pp. 177-201.

Anderson-Hsieh, J. and Koehler, K. (1988) The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning* 38, 4: 561-613.

Anderson-Hsieh, J., Johnson, R. and Koehler, K. (1992) The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning* 42: 529-555.

Anthony, E. (1963) Approach, method and technique. *English Language Teaching* 17: 63-57.

Aoyama, K., Guion, S., Flege, J., Yamada, T. and Akahane-Yamada, R. (2008) The first years in an L2 speaking environment: A comparison of Japanese children and adults learning American English. *International Journal of Applied Linguistics* 46: 61-90.

Archibald, J. (1992) Transfer of L1 parameter settings: Some empirical evidence from Polish metrics. *Canadian Journal of Linguistics / Revue canadienne de Linguistique* 37: 301-339.

Archibald, J. (1994) A formal model of learning L2 prosodic phonology. *Second Language Research* 10, 3: 215-240.

Asher, J. J. (1977) *Learning another Language through Actions. The Complete Teachers' Guide Book.* Los Gatos: Sky Oaks Productions, Inc.

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

Bachman, L. (1990) *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Baddeley, A. D. (2000) The episodic buffer: A new component of working memory? *Trends in Cognitive Science* 4, 11: 417-423.

Baddeley, A. D. (2007) *Working memory, thought and action.* Oxford: Oxford University Press.

Baddeley, A. D. and Hitch, G. (1974) *Working memory.* In G.H. Bower (ed.) *The psychology of learning and motivation: Advances in research and theory* (Vol. 8). New York: Academic Press. pp 47-89.

Bagarić, V. and Djigunović, J. M. (2007) Defining communicative competence. *Metodika* 8: 94-103.

Bamkin, S. (2010) How Not to Fix a Problem: Misapplications of *Pronunciation* Theory. *Gateway Papers*, Vol 1.

Barrera Pardo, D. (2004) Can pronunciation be taught? A review of research and implications for teaching. *Revista Alicantina de Estudios Ingleses* 17: 31-44.

Bassetti, B. (2009) Orthographic input and second language phonology. In P. Thorsten and M. Young-Scholten (eds.) *Input Matters in SLA*. Clevedon, UK: Multilingual Matters. pp. 191-206.

Battistella, E. L. (1990) *Markedness: The evaluative superstructure of language*. Albany: State University of New York Press.

Bavelas, J. B. (2000) Nonverbal aspects of fluency. In H. Riggenbach (ed.) *Perspectives on fluency*. Ann Arbor: University of Michigan Press. pp. 91-101.

Beebe, L. (1987) Myths about interlanguage phonology. In G. Ioup and S. Weinberger (eds.) *Interlanguage. Phonology*. Cambridge: Newbury House. pp.165-175.

Benoît, C. (1990) An intelligibility test using semantically unpredictable sentences: towards the quantification of linguistic complexity. *Speech Communication* 9, 4: 293-304.

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

Bent, T., Bradlow, A. R. and Smith, B. L. (2007) Segmental errors in different word positions and their effects on intelligibility of non-native speech: All's well that begins well. In O-S. Bohn and M. J. Munro (eds.) *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam: John Benjamins. pp. 331-347.

Best, C. (1995) A direct realist view of cross-language speech perception. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*. Timonium, MD: York Press Inc. pp. 171-206.

Best, C. and Tyler M. (2007) Nonnative and second-language speech perception: Commonalities and complementarities. In O-S. Bohn and M. J. Munro (eds.) *Language Experience in Second language Speech Learning. In honor of James Emil Flege.* Amsterdam: John Benjamins. pp. 13-34.

Beukelman, D. R. and Yorkston, K. M. (1979) The relationship between information transfer and speech intelligibility of dysarthric speakers. *Journal of Communication Disorders* 12: 189-196.

Birdsong, D. (1999) Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D. Birdsong (ed.) *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates. pp 1-22.

Birdsong, D. (2006) Age and second language acquisition and processing: A selective overview. *Language Learning* 56: 9-49.

Bisquerra, R. (ed.) (2004) *Metodología de la* investigación *educativa*. Madrid: La Muralla.

Blanken, G., Dittman, J., Haas, J-C. and Wallesch, C-W. (1987) Spontaneous speech in senile dementia and aphasia. Implications for a neurolinguistic model of language production. *Cognition* 27: 247-274.

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

Bohn, O-S. and Flege, J. E. (1996) Perception and production of a new vowel category by second-language learners. In A. James and J. Leather (eds.) *Second-Language Speech: Structure and Process*. Berlin: Walter de Gruyter. pp. 53-74.

Bongaerts, T., Planken, B. and Schils, E. (1995) Can late learners attain a native accent in a foreign language? A test of the Critical Period Hypothesis. In D. Singleton and Z. Lengyel (eds.) *The Age Factor in Second Language Acquisition*. Clevedon: Multilingual Matters. pp. 30-50.

Bongaerts, T., Summeren, C., Planken, B. and Schils, E. (1997) Age and ultimate attainment in the production of foreign language. *Studies in Second Language Acquisition* 19: 447-465.

Boomer, D. S. (1965) Hesitation and grammatical encoding. *Language and Speech* 8**:** 148-158.

Boula de Mareüil, P. and Vieru-Dimulescu, B. (2006) The contribution of prosody to the perception of foreign accent. *Phonetica* 63: 247-267.

Bowen, D. J. (1956) A comparison of the intonation patterns of English and Spanish. *Hispania* 39, 1: 30-35.

Bradlow, A. R. and Bent, T. (2002) The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America* 112, 1: 272-284.

Breitkreutz, J., Derwing, T. M. and Rossiter, M. J. (2002) Pronunciation teaching practices in Canada. *TESL Canada Journal* 19: 51-61.

Brennan S. E. and Williams W. (1995) The feeling of another's knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language* 34: 383-398.

Broselow, E. (1983) Non-obvious transfer: On predicting epenthesis errors. In S. Gass and L. Selinker (eds.) *Language transfer in language learning*. Rowley, MA: Newbury House. pp. 269-280.

Broselow, E. and Finer, D. (1991) Parameter setting in second language phonology and syntax. *Second Language Research* 7: 35-59.

Broselow, E., Chen, S. and Wang, C. (1998) The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20: 261-280.

Brown, A. (ed.) (1991) *Teaching English Pronunciation: A Book of Readings*. London: Routledge.

Brown, C. (1998) The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14: 136-193.

Brown, J. D. (1988) *Understanding research in second language learning*. Cambridge: Cambridge University Press.

Brown, J. D. and Rodgers, T. S. (2002) *Doing second language research*. Oxford: Oxford University Press.

Bruck, M. and Genesee, F. (1995) Phonological awareness in young second language learners. *Journal of Child Language* 22: 307-324.

Brumfit, C. J. (1984) *Communicative Methodology in Language Teaching: The Roles of Fluency and Accuracy.* Cambridge*:* Cambridge University Press.

Bruyninckx*,* M.*,* Harmegnies*,* B.*,* Llisterri*,* J. and Poch*,* D. (1994) Language-induced voice quality variability in bilinguals. *Journal of Phonetics* 22: 19-31.

Bucks, R. S., Singh, S. Cuerden, J. M. and Wilcock, G. K. (2000) Analysis of spontaneous conversational speech in DAT: Evaluation of an objective technique for analysing lexical performance. *Aphasiology* 14: 71-91.

Burleson, D. F. (2007) Intonation patterns in English tag questions of Japanese speakers of English as a second language. *IULC Working Papers,* 7.

Butterworth, B. (1980) *Language Production Vol. 1: Speech and Talk*. London: Academic Press.

Campbell, D. T. and Stanley, J. C. (1966) *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Campione, E. and Veronis, J. (2002) A Large-Scale Multilingual Study of Silent Pause Duration. In *Speech Prosody 2002*, Aix-en-Provence, France.
<http:/www.iscaspeech.org./archive>.
(accessed 03/04/2011)

Canale, M (l983) From communicative competence to communicative language pedagogy. In J. Richards and R. Schmidt (eds.) *Language and Communication*. London: Longman. pp 2-27.

Canale, M. and Swain, M. (l980) Theoretical bases of communicative approaches to Second language teaching and testing. *Applied Linguistics* 1, 1: 1-48.

Cantero, F. J. (1998) Conceptos clave en lengua oral. In A. Mendoza (ed.) *Conceptos clave en didáctica de la lengua y la literatura*. Barcelona: Horsori. pp. 141-154.

Cantero, F. J. (2003) Conceptos básicos en Didáctica de la Lengua y la Literatura. In A. Mendoza (ed.) *Didáctica de la lengua y la literatura*. Barcelona: Prentice Hall. pp. 33-79.

Carbó, C., Llisterri, J., Machuca, M-J., de la Mota, C., Riera, M. and Ríos, A. (2003) Estándar oral y enseñanza de la pronunciación del español como primera lengua y como lengua extranjera. *ELUA, Estudios de Lingüística de la Universidad de Alicante* 17: 161-180.

Cárdenas, D. N. (1960) *Introducción a una comparación fonológica del español y del inglés.* Washington: Centre for Applied Linguistics.

Carlisle, R. S. (1991) The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics* 12: 76-95.

Castino, J. (1996) El impacto de un curso de fonética en la adquisición de la fonología española. *Cuadernos Cervantes de la Lengua Española* 6: 44-47.

Catford, J. C. (1950) The background and origins of Basic English. *English Language Teaching* 5: 36-47.

Catford, J. C. (1987) Phonetics and the teaching of pronunciation. In J. Morley (ed.) *Current Perspectives on Pronunciation: Practices Anchored in Theory*. Washington DC: TESOL. pp. 83-100.

Cauldwell, R. (2002) The functional irrhythmicality of spontaneous speech: A discourse view of speech rhythms. *Applied Language Studies 2*: 1-24.
<http://www.solki.jyu.fi/apples/020201/main.htm> (accessed 02/04/2010).

Celce-Murcia, M. (2007) Rethinking the Role of Communicative Competence in Language Teaching. In E. A. Soler and M. P. S. Jordà (eds.) *Intercultural Language Use and Language Learning*. Springer Netherlands. pp. 41-57.

Celce-Murcia, M., Brinton, D. and Goodwin, J. (1996) *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.

Cenoz, J. (1991) *Enseñanza-aprendizaje del inglés como L2 o L3*. Donostia: Universidad del País Vasco.

Cenoz, J. (2003*) The additive effect* of bilingualism on third language acquisition: A review. *International Journal of Bilingualism* 7, 1: 71-87.

Cenoz, J. and Perales, J. (2000) Las variables contextuales y el efecto de la instrucción en la adquisición de segundas lenguas. In C. Muñoz (ed.) *Segundas lenguas. Adquisición en el aula*. Barcelona: Ariel. pp. 109-125.

Clark, J. and Yallop, C. (1990) *An introduction to phonetics and phonology*. Oxford: Basil Blackwell.

Cohen, L., Manion, I. and Morrison, K. (2005) *Research Methods in Education.* 5th edition. London: Routledge.

Cohen, S. P., Tucker, R., and Lambert, W. E. (1967) The comparative skills of monolinguals and bilinguals in perceiving phoneme sequences. *Language and Speech* 10: 159-168.

Cook, V. (2001) Using the First Language in the Classroom. *Canadian Modern Language Review* 57, 3: 402-423.

Cook, V. (2008) *Second Language Learning and Language Teaching*. London: Arnold.

Cook, M. and Lalljee, M. (1970) The interpretation of pauses by the listener. *British Journal of Social and Clinical Psychology* 9: 375-376.

Cortés Pomacóndor, S. (1999) Production and perception of English sounds by Catalan/Spanish speakers. *Proceedings of the First Congress of Experimental Phonetics*. Tarragona: Universitat de Barcelona. pp.165-170.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Couper, G. (2006) The short and long-term effects of pronunciation instruction. *Prospect* 21, 1: 6-66.

Crocker, L. (2010) Intelligibility of Hindi-accented English: The role of Duration and Intonation. Unpublished master's thesis. University of Minnesota.

Crystal, D. and Davy, D. (1975) *Advanced conversational English*. London: Longman.

Crystal, D. and Quirk, R. (1964) *Systems of prosodic and paralinguistic features in English.* The Hague: Mouton.

Cucchiarini, C. (1996) Assessing transcription agreement: methodological aspects. *Clinical Linguistics and Phonetics* 10, 2: 131-155.

Cucchiarini, C., Strik, H. and Boves, L. (2002) Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America* 111: 2862-2873.

Dalbor, J. B. (1980) *Spanish Pronunciation: Theory and Practice. An Introductory Manual of Spanish Phonology and Remedial Drills.* New York: Holt, Rinehart and Winston.

Dalton, P. and Hardcastle, W. (1977) *Disorders of fluency*. London: Edward Arnold.

Dalton, C. and Seidlhofer, B. (1994) *Pronunciation*. Oxford: OUP.

Daneman, M. and Carpenter, P. A. (1980) Individual differences in working memory and reading. *Journal of Learning and Verbal Behavior* 19: 450-466.

Daneman, M. and Green, I. (1986) Individual differences in comprehending and producing words in context. *Journal of Memory and Language* 25: 1-18.

Daneman, M., and Merikle, P. M. (1996) Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin and Review* 3: 422-433.

Dechert, R. and Lennon, S. (1989) Collocational blends of advanced second language learners. In S. Weigle (ed.) *Learning Vocabulary for Different Purposes*. Oxford: Pergamon Press. pp. 98-116.

Deese, J. (1984) *Thought into Speech: The Psychology of a Language*. Englewood Cliffs, NJ: Prentice-Hall.

De Graaff, R. and Housen, A. (2009) Investigating the effects and effectiveness of L2 instruction. In C. J. Doughty and M. H. Long and (eds.) *The Handbook of Language Teaching.* Oxford: Blackwell. pp. 726-755.

De Jong, N. H. and Wempe, T. (2009) Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods* 41, 2: 385-390.

Deng, J., Holtby, A., Howden-Weaver, L., Nessim, L., Nicholas, B., Nickle, K., Pannekoek, C., Stephan, S., and Sun, M. (2009) English pronunciation research: The neglected orphan of second language acquisition studies. *Prairie Metropolis Centre. Working Paper Series,* WP05-09, Edmonton, AB.

DeKeyser, R. (2000) The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499-533.

Delattre, P. (1965) *Comparing the phonetic features of English, French, German and Spanish: An interim report*. Heidelberg: Julius Groos Verlag.

Derwing, T. M. (2008) Curriculum issues in teaching pronunciation to second language learners. In J. G. Hansen Edwards and M. L. Zampini (eds.) *Phonology and second language acquisition.* Amsterdam: John Benjamins. pp. 347-369.

Derwing, T. M. and Munro, M. J. (1997) Accent, intelligibility and comprehensibility. *Studies in Second Language Acquisition* 19, 1: 1-16.

Derwing, T. M. and Munro, M. J. (2005) Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly* 39: 379-397.

Derwing, T. M. and Munro, M. J. (2009) Putting accent in its place: Rethinking obstacles to communication. *Language Teaching* 42: 276-490.

Derwing, T. M. and Rossiter, M. J. (2003) The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2-accented speech. *Applied Language Learning* 13: 1-18.

Derwing, T. M., Munro, M. J. and Wiebe, G. E. (1997) Pronunciation instruction for "fossilized" learners: Can it help? *Applied Language Learning* 8: 217-235.

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

Derwing, T. M., Rossiter, M. J., Munro, M. J. and Thomson, R. I. (2004) L2 fluency: Judgments on different tasks. *Language Learning* 54: 655-679.

Dlaska, A and Krekeler, C. (in press). The short-term effects of individual corrective feedback on L2 pronunciation. *System. An International Journal of Educational Technology and Applied Linguistics*.

Dörnyei, Z. and Skehan, P. (2003) Individual differences in second language learning. In C. J. Doughty and M. H. Long (eds.) *The handbook of second language acquisition*. Oxford: Blackwell. pp. 589-630.

Dörnyei, Z. and Thurrell, S. (1991) Strategic competence and how to teach it. *ELT Journal* 45: 16-23.

Doughty, C. J. (2003) Instructed SLA: Constraints, Compensation, and Enhancement. In C. J. Doughty and M. H. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford: Blackwell. pp. 256-310.

Doughty, C. and Williams, J. (1998) Issues and terminology. In C. Doughty and J. Williams (eds.) *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University Press. pp 1-13.

Dowd, A., Smith, J. and Wolfe, J. (1997) Learning to produce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time. *Language and Speech* 41: 1-20.

Dubno, J., Dirks, D. and Morgan, D. (1984) Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of Am*erica 76: 87-96.

Dubno, J. R., Horwitz, A. R. and Ahlstrom, J. B. (2003) Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. *Journal of the Acoustical Society of America* 113, 4: 2084-2094.

Eckman, F. (1977) Markedness and the Contrastive Analysis Hypothesis. *Language Learning* 27: 315-330.

Eckman, F. (1991) The structural conformity hypothesis and the acquisition of consonants clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition* 13: 23-41.

Eckman, F. (2008) Typological markedness and second language phonology. In J. G. Hansen Edwards and M. L. Zampini (eds.) *Phonology and second language acquisition.* Amsterdam: John Benjamins. pp. 95-116.

Egan, J. P. (1948) Articulation testing methods. *Laryngoscope* 58, 9: 955-991.

Ehrlich, S. (1997) Gender as social practice: Implications for second language acquisition. *Studies in Second Language Acquisition* 19: 421-446.

Ejzenberg, R. (2000) The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In H. Riggenbach (ed.) *Perspectives on fluency*. Michigan: The University of Michigan Press. pp. 287-314.

Elliott, A. (1995) Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *Modern Language Journal* 79: 356-371.

Elliott, A. (1997) On the teaching and acquisition of pronunciation within a communicative approach. *Hispania* 80: 96-108.

Elliott, A. (2003) Staking out the territory at the turn of the century: Integrating phonological theory, research, and the effect of formal instruction on pronunciation in the acquisition of Spanish as a Second Language. In B. Lafford and R. Salaberry (eds.) *Spanish Second Language Acquisition: State of the Science.* Washington, D.C.: Georgetown University Press. pp. 19-46.

Ellis, R. (1985) Sources of variability in interlanguage. *Applied Linguistics 6*: 118-131.

Ellis, R. (1994) *The study of second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (2005) *Instructed Second Language Acquisition: A literature review.* Wellington: Ministry of Education New Zealand.

Ellis, N. and Schmidt, R. (1997) Morphology and longer distance dependencies: laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition* 19: 145-171.

Ellis, N. and Sinclair, S.G. (1996) Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology* 49-A: 234-250.

Enomoto, K. (1994) L2 perceptual acquisition: The effect of multilingual linguistic experience on the perception of a "less novel" contrast. *Edinburgh Working Papers in Applied Linguistics* 5: 15-29.

Ertmer, D. (2010) Relationships between speech intelligibility and word articulation scores in children with hearing loss. *Journal of Speech, Language, and Hearing Research* 53*:* 1075-1086.

Escudero, P. (2007) Second Language Phonology: The role of perception. In M. Pennington (ed.) *Phonology in Context* (Palgrave Advances Series). London: Palgrave MacMillan. pp. 109-134.

Escudero, P. and Boersma, P. (2004) Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551-585.

Fayer, J. and Krasinski, E. (1987) Native and nonnative judgments of intelligibility and irritation. *Language Learning* 37: 313-326.

Fillmore, C. J. (1979) On fluency. In C. J. Fillmore, D. Kempler and W. S. Y. Wang (eds.) *Individual differences in language ability and language behavior*. New York: Academic Press. pp. 85-102.

Firth, S. (1992) Pronunciation syllabus design: A question of focus. In P. Avery and S. Ehrlich (eds.) *Teaching American English Pronunciation*. Oxford: Oxford University Press. pp. 173-183.

Fischman, M. (2001) Informed Consent. In B. Sales and S. Folkman (eds.) *Ethics in research with human participants*. Washington: APA. pp. 35-48.

Flege, J. E. (1984) The detection of French accent by American listeners. *Journal of the Acoustical Society of America* 76: 692-707.

Flege, J. E. (1987) The development of skill in producing word-final English stops. *Journal of the Acoustical Society of America* 82, 2: 433-447.

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

Flege, J. E. (1991) Age of learning affects the authenticity of voice-onset-time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89, 1: 395-411.

Flege, J. E. (1995) Second-language speech learning: Theory, findings and problems. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*. Timonium, MD: York Press Inc. pp. 233-272.

Flege, J. E. (1999) Age of learning and second-language speech. In D. P. Birdsong (ed.) *Second Language Acquisition and the Critical Period Hypothesis*. Hillsdale, N.Y.: Lawrence Erlbaum. pp. 101-132.

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

Flege, J. E. and Davidian, R. (1984) Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics* 5: 323-347.

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

Flege, J. E. and Fletcher, K. (1992) Talker and listener effects on the perception of degree of foreign accent. *Journal of the Acoustical Society of America* 91: 370-389.

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

Flege, J. E., Bohn, O-S. and Jang, S. (1997) Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25: 437-470.

Flege, J. E., Munro, M. J. and MacKay, I. R. A. (1995) Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97, 5: 3125-3134.

Flege, J. E., Yeni-Komshian, G. H. and Liu, Serena (1999) Age constraints on second-language acquisition. *Journal of Memory and Language* 4*:* 78-104.

Flipsen, P., Jr. (2006) Measuring the intelligibility of conversational speech in children. *Clinical Linguistics and Phonetics* 20, 4*:* 202-312.

Folkman, S. (2001) Privacy and Confidentiality. In B. Sales, B. and S. Folkman (eds.) *Ethics in research with human participants*. Washington: APA, pp. 49-58.

Fox Tree J. E. (1995) The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34: 709-738.

Fox Tree, J. E. (2002) Disfluencies in spoken language. In *Encyclopedia of Cognitive Science*. London: Macmillan.

Freed, B. F. (2000) Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach (ed.) *Perspectives on fluency*. Michigan: The University of Michigan Press. pp. 243-265.

Freed, B. F., Segalowitz, N. and Dewey D. (2004) Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition* 26: 275-301.

French, L. M. (2009) Phonological memory, intensive language instruction and L2 oral fluency development. Paper presented at the meeting of the American Association of Applied Linguistics, Denver, CO.

Fries, C. C. (1945) *Teaching and Learning English as a Foreign Language*. Ann Arbor: University of Michgan Press.

Fullana, N. (2005) Age-related effects on the acquisition of a foreign language phonology in a formal setting. Unpublished doctoral dissertation. Universitat de Barcelona.

García, J. and Cannito, M. (1996) Influence of verbal and nonverbal contexts on the sentence intelligibility of a speaker with dysarthria. *Journal of Speech and Hearing Research* 39: 750-760.

García Lecumberri, M. L. and Gallardo Del Puerto, F. (2003) English FL sounds in school learners of different ages. In M. P. García Mayo and M. L. García Lecumberri (eds.) *Age and the Acquisition of English as a Foreign Language*. pp. 115-135.

Gardner, R. C. and Lambert, W. E. (1959) Motivational Variables in Second-Language Acquisition. *Canadian Journal of Psychology* 13: 266-272.

Garman, M. (1990) *Psicolinguistics*. Cambridge: Cambridge University Press.

Gass, S., Roots, R. and Lee, J. (2006) Inhibition and working memory capacity in a second language. Paper presented at the 16th European Second Language Association Conference, Antalya, Turkey.

Gass, S. and Selinker, L. (2008) *Second Language Acquisition: An Introductory Course*. New York, NY: Routledge.

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

Gil, J. (2007) *Fonética para profesores de español: De la teoría a la práctica*. Madrid: Arco/Libros.

Gili Gaya, S. (1950) *Elementos de fonética general*. Madrid: Gredos.

Gimson, A. C. (1989) *An Introduction to the Pronunciation of English*. London: Hodder Arnold.

Gleitman, L. R., and Newport, E. L. (1995) The invention of language by children: Environmental and biological influences on the acquisition of language. In L. R. Gleitman and M. Liberman (eds.) *An Invitation to Cognitive Science*. Vol 1 *Language.* Cambridge: MIT Press. pp 1-24.

Goldman-Eisler, F. (1961) The significance of changes in the rate of articulation. *Language and Speech* 4, 4: 171-174.

Goldman-Eisler F. (1968) *Psycholinguistics: Experiments in Spontaneous Speech.* New York: Academic Press.

González-Bueno, M. (1997) The effects of formal instruction on the acquisition of Spanish stop consonants. In W. Glass and A.T. Pérez-Leroux (eds.) *Contemporary Perspectives on the Acquisition of Spanish*. *Volume 2: Production, Processing, and Comprehension*. Somerville, MA: Cascadilla. pp 57-75.

Graham, R. (1978) Intonation and emphasis in Spanish and English. *Hispania* 61: 95-101.

Grosjean, F. (1980) Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics* 28: 267-283.

Grotjahn, R. (1987) On the methodological basis of introspective methods. In C Faerch and G. Kasper (eds.) *Introspection in Second Language Research*. Clevedon, Avon: Multilingual Matters.

Guillot, M.-N. (1999) *Fluency and its teaching*. Clevedon, Avon: Multilingual Matters.

Guion, S. G., Flege, J. E., Liu, S. H., and Yeni-Komshian, G. H. (2000) Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics* 21: 205-228.

Gut, U. (2007) Foreign Accent. In C. Muller (ed.) *Speaker Classification*. Berlin: Springer.

Gutierrez Díez, F. (2001) The acquisition of English timing by native Spanish learners of English. An empirical study. *International Journal of English Studies* 1, 1: 93-115.

Hahn, L. D. (2004) Primary stress and intelligibility: research to motivate the teaching of suprasegmentals. *TESOL Quarterly* 38, 2: 201-223.

Hansen, Edwards J. G. (2006) *Acquiring a non-native phonology: Linguistic constraints and social barriers.* London: Continuum Publishers.

Hansen Edwards, J. G. (2008) Social factors and variation in production in L2 phonology. In J. G. Hansen Edwards and M. L. Zampini (eds.) *Phonology and Second Language Acquisition.* Amsterdam: John Benjamins Publishing Company. pp 251-279.

Hardison, D. (2010) Visual and auditory input in second-language speech processing [Research Timeline]. *Language Teaching* 43*:* 84-95.

Hardman, J. B. (2010) The *intelligibility* of Chinese-accented English to international and American students at a U.S. university. Unpublished doctoral dissertation. Ohio State University.

Hardy, J. E. (1993) Phonological learning and retention in second language acquisition. In F. Eckman (ed.) *Confluence: Linguistics, L2 acquisition and speech pathology*. Amsterdam: John Benjamins. pp. 235-247.

Harley, T. A. (2001) *The psychology of language* (2nd. Ed.). Hove: Psychology Press.

Harrington, M. W. and Sawyer, M. (1992) L2 working memory capacity and L2 reading skills. *Studies in Second Language Acquisition* 14, 1: 25-38.

Harris, K. C., Eckert, M. A., Ahlstrom, J. B. and Dubno, J.R. (2010) Age-related differences in gap detection: Effects of task difficulty and cognitive ability. *Hearing Research* 264: 21-29.

Harris, K. C., Mills, J. H., He, N.-J. and Dubno, J. R. (2008) Age-related differences in sensitivity to small changes in frequency assessed with cortical evoked potentials. *Hearing Research* 243: 47-56.

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

Hernández Sampieri, R., Fernández Collado, C. and Baptista, L. (1997) *Metodología de la Investigación*. Mexico: McGraw Hill.

Hodge, M. N. and Whitehill, T. L. (2010) Intelligibility impairments. In J. S. Damico, M. J. Ball, and N. Muller (eds.) *The handbook of language and speech disorders*. Oxford: Blackwell Publishers. pp. 99-114.

Holm, S. (2008) Intonational and durational contributions to the perception of foreign-accented Norwegian: An experimental phonetic investigation. Unpublished doctoral dissertation. Norwegian University of Science and Technology.

Hualde, J. I. (2005) *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Hughes, A. (2003) *Testing* for *language* teachers (2nd ed.). Cambridge, New York: Cambridge University Press.

Hustad, K. C. (2008) The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech Language and Hearing Research* 51, 3: 562-573.

Hustad, K. and Beukelman, D. (2001) Effects of linguistic cues and stimulus cohesion on intelligibility of severely dysarthric speech. *Journal of Speech Language and Hearing Research* 44: 497-510.

Hustad K and Beukelman D. (2002) Listener comprehension of severely dysarthric speech: effects of linguistic cues and stimulus cohesion. *Journal of Speech, Language, and Hearing Research* 45: 545-558.

Hustad, K., Dardis, C. and Kramper, A. (2011) Use of listening strategies for the speech of individuals with dysarthria and cerebral palsy. *Augmentative and Alternative Communication* 27, 1: 5-15.

Hyltenstam, K. and Abrahamsson, N. (2003) Maturational constraints in SLA. In C. J. Doughty and M. H. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford: Blackwell. pp. 539-588.

Illes, J. (1989) Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's and Parkinson's. *Brain and language* 37: 628-642.

Ioup, G. (1984) Is there a structural foreign accent? A comparison of syntactic and phonological errors in second language acquisition. *Language Learning* 34: 1-17.

Ioup, G., Boustagi, E., El Tigi, M. and Moselle, M. (1994) Re-examining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16: 73-98.

Iruela, A. (2004) Adquisición y enseñanza de la pronunciación en lenguas extranjeras. Unpublished doctoral dissertation. Universitat Autònoma de Barcelona.

Iruela, A. (2007) Principios didácticos para la enseñanza de la pronunciación en lenguas extranjeras. *MarcoELE. Revista de Didáctica del Español como Lengua Extranjera 4*.
http://www.marcoele.com/num/4/02e3c098b90d8d60d/pronunciacion.pdf
(accessed on 10-2-2010)

Isaacs, T. (2008) Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *Canadian Modern Language Review* 64, 4: 555-580.

James, A. R. (1983) Transferability and dialect phonology: Swabian: Swabian English. In A. James and B. Kettemann (eds) *Dialektphonologie und Fremdsprachenerwerb*

[*Dialect phonology and foreign language acquisition*]. Tübingen, Germany: Narr. pp. 162-188.

Jeng, J., Weismer, G. and Kent, R. D. (2006) Production and perception of Mandarin tone in adults with cerebral palsy. *Clinical Linguistics and Phonetics* 20, 1: 67-87.

Jenkins, J. (2000) *The phonology of English as an international language: new models, new norms, new goals*. Oxford: Oxford University Press.

Jeng, J. Y., Weismer, G. and Kent, R. D. (2006) Production and perception of mandarin tone in adults with cerebral palsy. *Clinical linguistics and phonetics* 20, 1: 67-87.

Johnson, J. S. and Newport, E. L. (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60-99.

Keintz, C. K., Bunton, K., and Hoit, J. D. (2007) Influence of visual information on the intelligibility of dysarthric speech. *American Journal of Speech Language Pathology* 16: 222-234.

Kellerman, E. (1995) Crosslinguistic influence: Transfer to nowhere? *Annual review of applied linguistics* 15: 125-150.

Kempler, D. and Van Lancker, D. (2002) The effect of speech task on intelligibility in dysarthria: A case study of Parkinson's Disease. *Brain and Language* 80: 449-464.

Kendall, T. (2009) Speech Rate, Pause, and Linguistic Variation: An Examination Through the Sociolinguistic Archive and Analysis Project. Unpublished doctoral dissertation. Duke University.

Kennedy, S. and 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, 3: 459-489.

Kent, R. D., Weismer, G., Kent, J. F. and Rosenbek, J. C. (1989) Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders* 54: 482-499.

Kenworthy, J. (1996) *Teaching English Pronunciation*. Harlow: Longman.

Klasner, E. R., and Yorkston, K. M. (2005) Speech intelligibility in ALS and HD dysarthria: The everyday listener's perspective. *Journal of Medical Speech-Language Pathology* 13: 127-139.

Kluge, D. C., Rauber, A. S., Reis, M. S. and Bion, R. A. H (2007) The relationship between the perception and production of English nasal codas by Brazilian learners of English. *Proceedings of Interspeech 2007*. Antwerp, Belgium. pp. 2297-2300.

Kormos, J. (2006) *Speech production and second language acquisition. Cognitive sciences and second language acquisition*. Mahwah, N.J: Lawrence Erlbaum Associates.

Kowal, S. and O'Connell, D. (1980) Pausological research at Saint Louis University. In H. Dechert and M. Raupach (eds.) *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. The Hague: Mouton. pp 61-66.

Krashen, S. (1981) *Second language acquisition and second language learning*. Oxford: Pergamon Press.

Krashen, S. and Terrell**,** T. D**. (**1983) *The Natural Approach: Language Acquisition in the Classroom.* Oxford: Pergamon Press.

Kuhl, P. K. (1993) Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. deBoysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage and J. Morton (eds.) *Developmental neurocognition: Speech and face processing in the first year of life.* Dordrecht, Netherlands: Kluwer Academic Publishers. pp. 259-274.

Kuhl, P. K. (2008) Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo and L. Freund (eds.) *Infant pathways to language: Methods, models, and research directions*. New York: Erlbaum. pp. 213-243.

Labov, W. (1990) The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 205-254.

Labov, W (2001) *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.

Ladefoged, P. (1975) *A Course in Phonetics*. Orlando: Harcourt Brace.

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

Lambacher, S. (2001) A brief guide to resources for developing expertise in the teaching of pronunciation. *Prospect: Australian Journal of TESOL* 16, 1: 63-74.

Laroy, C. (1995) *Pronunciation*. Oxford: Oxford University Press (Resource Books for Teachers).

Larsen-Freeman, D. and Long, M. H. (1991) *An Introduction to Second Language Acquisition Research*. London: Longman.

Laver, J. (1980) *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Leather, J. (1999) Second Language Speech Research: An Introduction. *Language Learning* 49, S1: 1-56.

Leather, J. and James, A. R. (1991) The acquisition of second language speech. *Studies in Second Language Acquisition* 13: 305-341.

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

Lennon, P. (1990) Investigating fluency in EFL: A quantitative approach. *Language Learning* 40: 387-417.

Levelt W. J. M (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levis, J. (2005) Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly* 39, 3: 369-377.

Levis, J. (2007) Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics* 27: 1-19.

Lewandowsky, S., Oberauer, K., Yang, L.-X. and Ecker, U. K. H. (2010) A working memory test battery for Matlab. *Behavioral Research Methods* 42: 571-585.

Liss, J. M., Spitzer, S. M., Caviness, J. N., and Adler, C. (2002) The effects of familiarization on intelligibility and lexical segmentation of hypokinetic and ataxic dysarthria. *Journal of the Acoustical Society of America* 112: 3022-3031.

Llisterri, J. (1995) Relationships between Speech Production and Speech Perception in a Second Language. In K. Elenius and P. Branderud (eds.) *Proceedings of the XIIIth International Congress of Phonetic Sciences.* Stockholm, Sweden, Vol. 4. pp. 92-99.
<http://liceu.uab.cat/~joaquim/publicacions/LListerri_95_L2_Production_Perception.pdf>
(accessed on 12-08-2010).

Llisterri, J. (2003a) La enseñanza de la pronunciación. *Revista del Instituto Cervantes en Italia* 4, 1: 91-114.

Llisterri, J. (2003b) La evaluación de la pronunciación en la enseñanza del español como segunda lengua. In M. V. Reyzabal (ed.) *Perspectivas teóricas y metodológicas: Lengua de acogida, educación intercultural y contextos inclusivos*. Madrid: Dirección General de Promoción Educativa, Consejería de Educación, Comunidad de Madrid. pp. 547-562.

Lombardi, L. (2003) Second Language Data and Constraints on Manner: Explaining Substitutions for the English Interdentals. *Second Language Research* 19, 3: 225-250.

Long, M. H. (1990) Maturational constraints on language development. *Studies in Second Language Acquisition* 12: 251-285.

Lord, G. (2005) (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania* 88: 557-567.

Lord, G. (2006) Defining the indefinable: Study abroad and phonological memory abilities. In C. A. Klee and T. L. Face (eds.) *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages*. Somerville, MA: Cascadilla Proceedings Project. pp. 40- 46.

Lybeck, K. E. (2002) The role of acculturation and social networks in the acquisition of second language pronunciation. Unpublished doctoral dissertation. University of Minnesota.

Maasen, B. and Povel, D.-J. (1985) The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech. *Journal of the Acoustical Society of America* 78: 877-886.

MacCarthy, P. (1978) *The Teaching of Pronunciation*. Cambridge: Cambridge University Press.

MacDonald, S. (2002) Pronunciation views and practices of reluctant teachers. *Prospect* 17, 3: 3-18.

McGarr, N. (1981) The Effect of Context on the Intelligibility of Hearing and Deaf Children's Speech. *Language and Speech* 24: 255-263.

Major, R. C. (1986) The ontogeny model: evidence from L2 acquisition of Spanish *r*. *Language Learning* 36: 453-504.

Major, R. C. (1987) A model for interlanguage phonology. In G. Ioup and S. Weinberger (eds.) *Interlanguage phonology: the acquisition of a second language sound system*. New York: Newbury House/Harper and Row. pp. 101-25.

Major, R. C. (2001) *Foreign accent: The Ontogeny and Phylogeny of Second Language Phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Major, R. C. (2007) Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition* 29: 539-556.

Major, R. C. (2008) Transfer in second language phonology: A review. In J. G. Hansen Edwards and M. L. Zampini (eds.) *Phonology and Second Language Acquisition.* Amsterdam: John Benjamins Publishing Company. pp. 63-95.

Major, R., Fitzmaurice, S., Bunta, F. and Balasubramania, C. (2002) The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly* 36: 173-190.

Manchón Ruiz, R. (1985) Estudios de interlengua: análisis de errores, estrategias de aprendizaje y estrategias de comunicación. *Revista Española de Lingüística Aplicada* 1: 55-75.

Marghany, M. M. (2002) Some phonetic features of Egyptian English: A descriptive and pedagogical perspective. In A. James and J. Leather (eds.), *New Sounds 2000:*

Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech, Klagenfurt, Austria: University of Klagenfurt. pp 231-235.

Marinova-Todd, S. H., Marshall, D. B. and Snow, C. E. (2000) Three misconceptions about age and L2 learning. *TESOL Quarterly* 34: 9-34.

McAllister, R., Flege, J. E. and Piske, T. (2002) The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30: 229-258.

McCarthy, M. (2009) Rethinking spoken fluency. *ELIA* 9: 11-29. <http://institucional.us.es/revistas/revistas/elia/pdf/9/3.%20McCarthy.pdf> (accessed on 24/02/2010).

McGarr, N. (1983) The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research* 26: 451-458.

McGroarty, M. (2001) Situating Second Language Motivation. In Z. Dornyei, and R. W. Schmidt (eds.) *Motivation and Second Language Acquisition*. Honolulu, University of Hawai'i. pp. 69-91.

McNab, F., Varrone, A., Farde, L., Jucaite, A., Bystritsky, P., Forssberg, H. and Klingberg, T. (2009) Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* 323: 800-802.

McNerney, M. and Mendelsohn, D. (1992) Suprasegmentals in the Pronunciation Class: Setting Priorities. In P. Avery and S. Ehrlich (eds.) *Teaching American English Pronunciation*. Oxford: Oxford University Press. pp. 185-196.

Meyerhoff, M. (2011) *Introducing Sociolinguistics* (2nd Edition). Oxford: Routledge.

Miyake, A. and Friedman, N. (1998) Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy and L. Bourne (eds.) *Foreign language learning: Psycholinguistic studies on training and retention*. London: Lawrence Erlbaum Associates. pp. 339-365.

Monsen, R. B. (1978) Toward measuring how well hearing impaired children speak. *Journal of Speech and Hearing Research* 21: 197-219.

Monsen, R. B. (1983) The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders* 48: 286-296.

Morley, J. (1991) The Pronunciation Component in Teaching English to Speakers of Other Languages. *TESOL Quarterly* 25, 3: 481-520.

Moulton, W. G. (1962) Toward a Classification of Pronunciation Errors. *Modern Language Journal* 46: 101-109.

Moyer, A. (1999) Ultimate attainment in L2 phonology: The critical factors of age, motivation and instruction. *Studies in Second Language Acquisition* 21: 81-108.

Moyer, A. (2004) *Age, Accent and Experience in Second Language Acquisition. An Integrated Approach to Critical Period Inquiry*. Clevedon: Multilingual Matters.

Moyer, A. (2007) Empirical Considerations on the Age Factor in L2 Phonology. *Issues in Applied Linguistics* 15, 2: 109-127.

Muñoz, C. (2008) Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics* 29, 4: 578-596.

Muñoz, C. and Singleton, D. (2011) A critical review of age-related research on L2 attainment. *Language Teaching* 44, 1: 1-35.

Munro, M. J. (1995) Non-segmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition* 17: 17-34.

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. J. (2003) A primer on accent discrimination in the Canadian context. *TESL Canada Journal* 20, 2: 38-51.

Munro, M. J. (2008) Foreign accent and speech intelligibility. In J. G. Hansen Edwards and M. L. Zampini (eds.) *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins. pp. 193-218.

Munro, M. J. (2011) Intelligibility: Buzzword or buzzworthy?. In. J. Levis and K. LeVelle (eds.) *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. Ames, IA: Iowa State University. pp.7-16.

Munro, M. J. and Derwing, T. M. (1994) Evaluations of foreign accent in extemporaneous and read material. *Language Testing* 11: 253-266.

Munro, M. J. and Derwing, T. M. (1995) Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning* 45, 1: 73-97.

Munro, M. J. and Derwing, T. M. (2001) Modelling perceptions of the comprehensibility and accentedness of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition* 23: 451-468.

Munro, M. J. and Derwing, T. M. (2006) The functional load principle in ESL pronunciation instruction: An exploratory study. *System* 34: 520-531.

Munro, M. J. and Derwing, T. M. (2008) Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning* 58: 479-502.

Munro, M. J. and Derwing, T. M. (2011) The foundations of accent and intelligibility in pronunciation research. *Language Teaching* 44: 316-327.

Munro, M. J., Derwing, T. M. and Burgess, C. S. (2010) Detection of nonnative speaker status from content-masked speech. *Speech Communication* 52: 626-637.

Munro, M. J., Derwing, T. M. and Morton, S. L. (2006) The mutual intelligibility of foreign accents. *Studies in Second Language Acquisition* 28: 111-131.

Munro, M. J., Derwing, T. M. and Sato, K. (2006) Salient accents, covert attitudes: Consciousness-raising for pre-service second language teachers. *Prospect: An Australian Journal of TESOL* 21: 67-79.

Nakata, H. (2002) Correlations between musical and Japanese phonetic aptitudes by native speakers of English. *Reading Working Papers in Linguistics* 6: 1-23.

Navarro Tomás, T. (1932) *Manual de pronunciación española*. Madrid: C.S.I.C.

Navarro Tomás, T. (1946) Escala de frecuencia de los fonemas españoles. *Estudios de fonología española. Syracuse*, 15-30.

Nelson, C. L. (2008) *Intelligibility* since 1969. *World Englishes* 27: 297-308.

Neufeld, G. (1979) Toward a theory of language learning ability. *Language Learning* 29: 227-241.

Neufeld, G. (1988) Phonological asymmetry in second-language learning and performance. *Language Learning* 38, 4: 531-559.

Nunan, D. (1992) *Research Methods in Language Learning.* Cambridge: Cambridge University Press.

Obler, L. K. and Hannigan, S. (1996) Neurolinguistics of second language acquisition and use. In W. C. Ritchie and T. K. Bhatia (eds.) *Handbook of Second Language Acquisition.* San Diego: Academic Press.

O'Connor, J. D. (1973) *Phonetics*. Harmondsworth: Penguin Books.

OCR (2009) *GSCE French/German/Spanish*. Nottingham: OCR Publications.

Odisho, E. (1992) A comparative study of English and Spanish vowel systems: Theoretical and practical implications for teaching pronunciation (Report No. ED352836). http://www.eric.ed.gov/ERICDocs/data/ (accessed 12/03/2009).

Odisho, E. (2003) *Techniques of teaching pronunciation in ESL, bilingual and foreign language classes*. München: Lincom-Europa.

Odlin, T. (2003) Cross-linguistic Influence. In C. Doughty and M. H. Long (eds.) *Handbook on Second Language Acquisition*. Oxford: Blackwell. pp. 436-486.

Ohara, Y. (2001) Finding one's voice in Japanese: A study of the pitch levels of L2 users. In A. Pavlenko, A. Blackledge, I. Piller and M. Teutsch-Dwyer (eds.)

Multilingualism, second language learning, and gender. Berlin: Mouton de Gruyter. pp. 231-256.

Oller, J. and Ziahosseini, S. (1970) The contrastive analysis hypothesis and spelling errors. *Language Learning* 20, 2: 183-189.

Ortega, L. (2009) *Understanding second language* acquisition. London: Hodder Arnold.

Ortega, J., González, J. and Marrero, V. (2000) AHUMADA: A large corpus in Spanish for speaker characterization and identification. *Speech Communication* 31, 2-3: 255-264.

Osaka, M. and Osaka, N. (1992) Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society* 30, 4: 287-289.

Osburne, A. G. (1996) Final consonant cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics* 178: 164-181.

Oyama, S. (1976) A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research* 5: 261-283.

Pennington, M. (1994) Recent research in L2 phonology: Implications for practice. In J. Morley (ed.) *Pronunciation pedagogy and theory. New views, new directions.* Alexandria, VA: Teachers of English to Speakers of Other Languages. pp. 92-108.

Pennington, M. and Richards, J. (1986) Pronunciation revisited. *TESOL Quarterly* 20, 2: 207-225.

Perlmutter, M. (1989) Intelligibility rating of L2 speech pre- and postintervention. *Perceptual and Motor Skills* 68: 515-521.

Peterson, L. R. and Peterson, M. J. (1959) Short-term retention of individual verbal items. *Journal of Experimental Psychology* 58: 193-198.

Phillips, S. L., Gordon-Salant, S., Fitzgibbons, P. J. and Yeni-Komshian, G. H. (1994) Auditory duration discrimination in young and elderly listeners with normal hearing. *Journal of the American Academy of Audiology* 5: 210-215.

Pica, T. (1984) Pronunciation activities with an accent on communication. *English Teaching Forum* 22, 3: 2-6.

Pike, K. L. (1945) *The Intonation of American English*. Ann Arbor: The University of Michigan Press.

Piller, I. (2002) Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics* 6, 2: 179-206.

Piske, T., MacKay, I. R. A. and Flege, J. E. (2001) Factors affecting the degree of foreign accent in an L2: A review. *Journal of Phonetics* 29, 2: 191-215.

Poch, D. (1992) The rain in Spain. *Cable: Revista de didáctica del español como lengua extranjera* 10: 5-9.

Pointon, G. (1980) Is Spanish really syllable-timed? *Journal of Phonetics* 8: 293-305.

Polivanov, E. (1931) La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4: 79-96.

Porter, D. and Garvin, S. (1989) Attitudes to Pronunciation in ESL. *Speak Out* 5: 8-15.

Price P. J, Ostendorf M., Shattuck-Hufragel S. and Fong C. (1991) The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90, 6: 2956-2970.

Puigvert Ocal, A. (2001) Fonética contrastiva español-alemán, español-inglés, español-francés y su aplicación a la enseñanza de la pronunciación española. *Carabela, 49*: 17-37.

Puigví, D. and Fernández, J. M. (1992) *Estudi de la durada i la situació de les pauses en castellà*. Unpublished manuscript. Departament de Filologia Espanyola, Universitat Autònoma de Barcelona.

Puigví, D. and Fernández, J. M. (1993) *Las pausas no marcadas: situación y duración (Las pausas no marcadas ortográficamente en castellano)*. Unpublished manuscript. Departament de Filologia Espanyola, Universitat Autònoma de Barcelona.

Pujol, J., Soriano-Mas C., Ortiz H., Sebastián-Gallés N., Losilla J. M. and Deus J. (2006) Myelination of language-related areas in the developing brain. *Neurology* 66: 339-343.

Purcell, E. and Suter, R. (1980) Predictors of pronunciation accuracy: A reexamination. *Language Learning* 30, 2: 271-87.

Qualification and Curriculum Authority (2004) *Modern foreign languages in the Key Stage 4 curriculum.*
<www.archive.teachfind.com/qcda/orderline.qcda.gov.../1847212859.pdf>
(accessed on 2/04/2011).

Quené, H. and van Delft, L. E. (2010) Non-native durational patterns decrease speech intelligibility. *Speech Communication* 52, 11-12: 911-918.

Quilis, A. and Fernández, J. (1996) *Curso de Fonética y Fonología Españolas,* Madrid: CS1C.

Rajagopalan, K. (2010) The Soft Ideological Underbelly of the Notion of Intelligibility in Discussions about 'World Englishes.' *Applied Linguistics* 31: 465-470.

Rasmus, F., Nespor, M. and Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-292.

Renard, R. (1979) *Introduction à la méthode verbo-tonale de correction phonétique.* Troisième édition entièrement refondue. Bruxelles/Centre International de Phonétique Appliquée: Didier/Mons.

Reves, T. (1983) What makes a good language learner? Unpublished doctoral dissertation. Hebrew University of Jerusalem.

Riazantseva, A. (2009) Second language proficiency and pausing: A Study of Russian speakers of English. *Studies in Second Language Acquisition* 23, 4: 497-526.

Richards, J. C. and Rodgers, T. S. (1998) *Approaches and Methods in Language Teaching.* Cambridge: Cambridge University Press.

Riney, T. J. and Flege, J. E. (1998) Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition* 20: 213-243.

Rivers, W. M. (1981*) Teaching-Foreign Language Skills*. Chicago: The University of Chicago Press.

Roach, P. (1998) Some languages are spoken more quickly than others. In L. Bauer and P. Trudgill (eds.) *Language Myths*. London: Penguin. pp. 150-158.

Robb, M, Maclagan, M. and Chen, Y. (2004) Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics and Phonetics* 18: 1-15.

Robinson, P. (1995) Attention, memory, and the "noticing" hypothesis. *Language Learning* 45: 283-331.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics* 25: 45-73.

Roca, I. and Johnson, W. (1999) *A Course in phonology*. Oxford: Blackwell.

Rogers*,* C. L. (1997) Intelligibility of Chinese-accented English. Unpublished doctoral dissertation. Indiana University.

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

Rumsey, D. (2007) *Intermediate Statistics for Dummies*. Hoboken, NJ: Wiley Publishing.

Saiegh-Haddad, E., and Geva, E. (2008) Morphological awareness, phonological awareness, and reading in English-Arabic bilingual children. *Reading and Writing: An Interdisciplinary Journal* 21: 481-504.

Saito, K. and Lyster, R. (2012) Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning* 62*:* 595-633.

Salmons, J., Jacewicz, E. and Allen Fox, R. (2008) Fast talkers versus slow talkers: Speech rate across dialect, generation, and gender. Paper presented at the American Dialect Society Annual Meeting 2008: Chicago, IL.

Samar, V. J. and Metz, D. E. (1988) Criterion validity of speech intelligibility rating-scale procedures for the hearing-impaired population. *Journal of Speech, Language and Hearing Research* 31, 3: 307-316.

Santos Barreto, S. and Zazo Ortiz, K. (2010) Intelligibility measurements in speech disorders: a critical review of the literature. *Pró-fono: revista de atualização científica* 20, 3: 201-206.

Sanz, C. (2000) Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics* 21: 23-44.

Sato, C. (1984) Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language and Learning* 34: 43-57.

Schairer, K. E. (1992) Native speaker reaction to non-native speech. *Modern Language Journal* 76, 3: 309-319.

Scherer, K. R. (1984) Emotion as a multi-component process: A model and some cross-cultural data. In P. Schaver (ed.) *Review of personality and social psychology: Vol. 5, Emotions, relationships and health.* Beverly Hills, CA: Sage. pp. 37-63.

Schiavetti, N. (1992) Scaling procedures for the measurement of speech intelligibility. In R. Kent (ed.) *Intelligibility in speech disorders: Theory, measurement and management*. Philadelphia: John Benjamins. pp. 11-34.

Schiavetti, N., Metz, D. E. and Sitler, R. W. (1981) Construct validity of direct magnitude estimation and interval scaling of speech intelligibility: Evidence from a study of the hearing impaired. *Journal of Speech, Language and Hearing Research*, 24, 3: 441.

Schmidt, R. (1990) The role of consciousness in second language learning. *Applied Linguistics* 11: 129-58.

Schmidt, R. (1992) Psychological Mechanisms Underlying Second Language Fluency. *Studies in Second Language Acquisition* 14: 357-385.

Schmidt, A. M. and Flege, J. E. (1996) Speaking rate effects in stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica* 53, 3: 162-179.

Schneider, B. and Hamstra, S. (1999) Gap detection thresholds as a function of tonal duration for younger and older listeners. *Journal of the Acoustica*l *Society of America* 106: 371-380.

Schoetz, S (2002) Linguistic and paralinguistic phonetic variation in speaker recognition and text-to-speech Synthesis. GSLT, Speech Technology I.
http://www.speech.kth.se/~rolf/gslt_papers/SusanneSchotz.pdf
(accessed on 6-1-2009).

Schumann, J. (1978) The acculturation model for second language acquisition. In R. Gingras (ed.) *Second Language Acquisition and Foreign Language Teaching*. Arlington, VA: Center for Applied Linguistics. pp. 27-50.

Schumann, J. (1999) A perspective on affect. In J. Arnold (ed.) *Affect in language learning*. Cambridge: Cambridge University Press. pp. 28-41.

Scovel, T. (1988) *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* New York, NY: Newbury House/ Harper and Row.

Scovel, T. (2000) A critical review of the critical period research. *Annual Review of Applied Linguistics* 20: 213-223.

Segalowitz, N. (2000) Automaticity and attentional skill in fluent performance. In H. Riggenbach (ed.) *Perspectives on fluency.* Ann Arbor: University of Michigan Press. pp. 200-219.

Selinker, L. (1972) Interlanguage. *IRAL* 10: 210-231.

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

Sell, D., Grunwell, P., Mildinhall, S., Murphy, T., Cornish, T. C., Williams, A., Bearn, D., Shaw, W. C., Murray, J. and Sandy, J. (2001) The Clinical Standards Advisory Group (CSAG) Study. *The Cleft Palate–Craniofacial Journal* 38: 30-37.

Seubsunk, S. (2000) The pronunciation of English stop-liquid and stop-glide clusters in word-initial position by native speakers of two different dialects of Thai. Unpublished doctoral dissertation. University of Delaware.

Sewell, A. (2010) Research methods and *intelligibility* studies. *World Englishes* 29: 257-269.

Seyfeddinipur, M., Kita, S. and Indefrey, P. (2008) How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. *Cognition* 108, 3: 837-842.

Silva Corvalán, C. (2001) *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.

Singleton, D. (1989) *Language acquisition. The age factor*. Clevendon: Multilingual Matters.

Skehan, P. (1989) *Individual Differences in Second-Language Learning*. London: Arnold Hodder.

Skehan, P. (2002) Theorising and updating aptitude. In P. Robinson (ed.) *Cognition and second language instruction*. Amsterdam: John Benjamins. pp. 69-93.

Slevc, L. R. and Miyake, A. (2006) Individual differences in second language proficiency: Does music ability matter? *Psychological Science* 17: 675-681.

Smith C. R. (1975) Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research* 18: 795-811.

Smith, L. E. and Bisazza, J. A. (1982) The comprehensibility of three varieties of English for college students in seven countries. *Language Learning* 32: 259-269.

Smith, L. E. and Nelson, C. (1985) International intelligibility of English: Directions and resources. *World Englishes* 4: 333-342.

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

Stevens, S. S. (1975) *Psychophysics*. New York: Wiley.

Stevick, E. W. (1978) Towards a practical philosophy of pronunciation: another view. *TESOL Quarterly* 12, 2: 145-150.

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

Stockwell, R. and Bowen, J. (1965) *The sounds of English* and *Spanish*. Chicago: University of Chicago Press.

Strange, W. and Shafer, V. L. (2008) Speech perception in second language learners: the re-education of selective perception. In J. G. Hansen Edwards and M. L. Zampini (eds.) *Phonology and Second Language Acquisition*. Amsterdam: John Bejamins. pp 153-193.

Subtelny, J. D., Vanhattum, R. J. and Myers, B. B. (1972) Ratings and measures of cleft palate speech. *Cleft Palate Journal* 9: 18-27.

Suter, R. (1976) Predicators of Pronunciation accuracy in second language learning. *Language Learning* 26: 233-53.

Tabachnick, B. G. and Fidell, L. S. (1996) *Using multivariate statistics* (3rd ed.). New York: HarperCollins.

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

Tajima, K., Port, R. and Dalby, J. (1997) Effects of speech timing on intelligibility of foreign accented English. *Journal of Phonetics* 25: 1-24.

Tanaka, A. and Nakamura, K. (2004) Auditory memory and proficiency of second language speaking: A latent variable analysis approach. *Psychological Reports* 95: 723-734.

Tarone, E. (1978) The phonology of interlanguage. In J. Richards (ed.) *Understanding second and foreign language learning*. Rowley, MA: Newbury House.

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

Tench, P. (1997) Towards a design of a pronunciation test. *Speak Out!* 20: 29-43.

Thompson, I. (1991) Foreign Accents Revisited: The English *Pronunciation* of Russian Immigrants. *Language Learning* 41: 177-204.

Tikofsky, R. S. and Tikofsky, R. P. (1964) *Intelligibility* measures of dysarthric speech. *Journal of Speech and Hearing Research* 7: 325-333.

Timmis, I. (2002) Native-speaker norms and international English: A classroom view. *ELT Journal* 56: 240-249.

Torres, J. R. (2006) Adecuación de los manuales de pronunciación de inglés para hispanohablantes. *Phonica* 2: 1-27.

Towell, R., (2002) Relative degrees of fluency: A comparative case study of advanced learners of French. *International Review of Applied Linguistics in Language Teaching* 40, 2: 117-150.

Trofimovich, P. and 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: 1-30.

Trofimovich, P., Lightbown, P. M., Halter, R., and Song, H. (2009) Comprehension-based practice: The development of L2 pronunciation in a listening and reading program. *Studies in Second Language Acquisition* 31: 609-639.

Trubetzkoy, N. S. (1939) *Grundzüge der Phonologie*. Travaux du Cercle Linguistique de Prague 7.

Tulving, E. (2002) Episodic memory: From mind to brain. *Annual Review of Psychology* 53: 1-25.

Ushioda, E. (2001) Language learning at university: Exploring the role of motivational thinking. In Z. Dörnyei and R. Schmidt (eds.) *Motivation and second language acquisition*. Honolulu, HI: University of Hawaii Press. pp 93-125.

Valero, A. (1991) *El corpus de las frases psicoacústicas de Harvard: una adaptación al castellano*. Unpublished manuscript. Universitat Autònoma de Barcelona, Departamento de Filología Española.

van den Noort, M., Bosch, P. and Hugdahl, K. (2006) Foreign language proficiency and working memory capacity. *European Psychologist* 11: 289-296.

van Wijngaarden, S. J., Steeneken, H. J. M. and Houtgast, T. (2002) Quantifying the intelligibility of speech in noise for non-native talkers. *Journal of the Acoustical Society of America* 112, 6: 3004-3013.

Varonis, E. and Gass, S. (1982) The comprehensibility of non-native speech. *Studies in Second Language Acquisition* 4: 114-136.

Vázquez, G. (2000) *La destreza oral*. Madrid: Edelsa.

Venkatagiri, H. S. and Levis, J. (2007) Metaphonological knowledge and comprehensibility: An exploratory study. *Language Awareness* 16: 263-277.

Visser, T. (2004) Comprehensibility: A potential measure for improvement in communication. Unpublished master's thesis. University of British Columbia.

Wang, M., and Geva, E. (2003) Spelling acquisition of novel English phonemes in Chinese children. *Reading and Writing* 4: 325-348.

Wardhaugh, R. (1970) The Contrastive Analysis Hypothesis. *TESOL Quarterly* 4, 2: 123-130.

Walsh, M. and Diller, A. (1981) Neuro-linguistic Considerations of the Optimum Age for Second Language Learning. In A. Diller (ed.) *Second Language Learning and Neurolinguistics*. N.Y: Copp Clark. pp 3-21.

Weinreich, U. (1953) *Languages in Contact*. The Hague: Mouton.

Weismer, G. (2008) Speech intelligibility. In M. J. Ball, M. Perkins, N. Müller and S. Howard (eds.) *The Handbook of Clinical Linguistics*. Oxford: Blackwell Publishers. pp. 568-582.

Weismer, G. and Martin, R. E. (1992) Acoustic and perceptual approaches to the study of intelligibility. In R. D. Kent (ed.) *Intelligibility in speech disorders: Theory, measurement, and management*. Philadelphia: John Benjamins. pp. 67-118.

Wen, T., (2005) *The Role of Motivation in Second Language Pronunciation*. Unpublished master's thesis. University of North Texas.

Werker, J. (1986) The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics* 7: 141-156.

Weston, A. D., and Shriberg, L. D. (1992) Contextual and linguistic correlates of intelligibility in children with developmental phonological disorders. *Journal of Speech and Hearing Research* 35: 1316-1332.

Whitehill, T. L. (1997) Speech intelligibility in Cantonese speakers with congenital dysarthria. Unpublished doctoral dissertation. The University of Hong Kong.

Whitehill, T. L. and Chau C. (2004) Single-word intelligibility in speakers with repaired cleft palate. *Clinical Linguistics and Phonetics* 18: 341-355.

Whitehill, T. L. and Ciocca V. (2000) Speech errors in Cantonese speaking adults with cerebral palsy. *Journal of Clinical Linguistics and Phonetics* 14, 2: 111-130.

Whitehill, T. L. and Wong, C., C-Y. (2006) Contributing factors to listener effort for dysarthric speech. *Journal of Medical Speech-Language Pathology* 14: 335-341.

Whitley, M. (2002) *A course in Spanish linguistics: Spanish/English contrasts* (2nd ed.). Washington, D.C.: Georgetown University Press.

Williams, J. N. and Lovatt, P. J. (2003) Phonological Memory and Rule Learning. *Language Learning* 53*,* 1: 67-121.

Wilson, E. O. and Spaulding, T. J. (2010) Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. *Journal of Speech, Language and Hearing Research* 53, 6: 1543-1554.

Wode, H. (1976) Developmental principles in naturalistic L1 acquisition. *Arbeitspapiere zum Spracherwerb*, 16. Kiel: Department of English, Kiel University.

Yin, R. K. (2003) *Case study research, design and methods*. Newbury Park: Sage Publications.

Yorkston, K. M., and Beukelman, D. (1978) A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders* 11: 499-512.

Yorkston, K. M. and Beukelman, D. (1981) Ataxic dysarthria: Treatment sequences based on intelligibility and prosodic considerations. *Journal of Speech and Hearing Disorders* 46: 398-404.

Yorkston, K. M., Strand, E. A., and Kennedy, M. R. (1996) Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology* 5: 55-66.

Yule, G. and MacDonald, D. (1995) The different effects of pronunciation teaching. *IRAL* 33: 345-350.

Young-Scholten, M. (1985) Interference reconsidered: The role of similarity in second language acquisition. *Selecta* 6: 6-12.

Zampini, M. L. (1996) Voiced stop spirantization in the ESL speech of native speakers. *Applied Psycholinguistics* 17: 335-354.

Zielinski, B. (2008) The listener: No longer the silent partner in reduced intelligibility. S*ystem* 36*,* 1: 69-84.

Zsiga, E. C. (2003) Articulatory timing in a second language. *Studies in Second Language Acquisition* 25: 399-432.

APPENDIX A: CONSENT FORM

Dear Parents,

At Queen Mary, University of London, we are studying the pronunciation problems of Year 11 English learners of Spanish. The study consists of having your son or daughter perform certain pronunciation tasks in Spanish. Your child will be recorded while reading in Spanish a list of words, sentences, a passage, as well as during a semi-spontaneous production task. Testing will last approximately 1 hour. Results of the tests will be made available to you.

If you give permission for your child to participate in this study, please sign the consent form.

Sincerely,

Ángel Osle
PhD candidate
Queen Mary, University of London
a.osleezquerra@qmul.ac.uk
Tel 07895328561

PURPOSE: Your son/daughter is invited to participate in a study of pronunciation problems of English learners of Spanish. The aim of the study is to collect information on Key Stage 4 students of Spanish. This information will be used to better understand intelligibility problems of English learners of Spanish and to improve materials for pronunciation instruction.

PARTICIPANT SELECTION: Your child was selected to be a participant in this study because he/she is studying Spanish at KS4 and has reported no hearing or speech problems.

PROCEDURES: Your child will be recorded while reading in Spanish a list of words, sentences, passage and during a semi-spontaneous production task. Testing will last approximately 1 hour. Breaks will be provided during the recording of the testing material if at any time your child indicates that he/she is tired.

CONFIDENTIALITY: Any information obtained in this study will remain strictly confidential and will only be disclosed with your permission. Your child will remain anonymous and will be assigned and represented by a number in order to protect his or her privacy.

WITHDRAWAL: Participation in this study is voluntary and your child will be free to withdraw from at any time without penalty.

CONTACT: Do not hesitate to contact us if you have any further questions:

Ángel Osle
a.osleezquerra@qmul.ac.uk
Tel. 07895328561

You and your child are under no obligation to participate in this study. Your signature indicates that you have read the information provided and have voluntarily decided to participate.

_____ _____
Signature of Parent or Legal Guardian Date

_____ _____
Student's signature Date

APPENDIX B: STUDENTS' PRELIMINARY QUESTIONNAIRE

1. Name _____

2. Age _____

3. Sex_____

4. Place of birth_____

5. Native language_____

6. Father's native language _____

7. Mother's native language _____

8. Languages spoken at home_____

9. Number of years studying Spanish_____

10. Have you studied any other foreign languages? _____ How many years?

11 Have you spent any time in a Spanish speaking country?

4. Never
5. Less than a month (e.g. family holidays, school trips etc.)
6. More than a month
7. Between 1 and 3 months
8. More than 6 months

12. How much time do you spend practising Spanish outside of the classroom?

Less than 1 hour a week
Between 1 and 2 hours a week
More than 2 hours a week
More than 3 hours a week
Over 4 hours a week

APPENDIX C: EVALUATORS' QUESTIONNAIRE

1. Nombre_____

2. Edad_____

3. Sexo_____

4. Lugar de nacimiento_____

5. Nivel de estudios_____

6. Profesión_____

7. ¿Es usted hablante nativo de español?_____

8. ¿Son sus padres hablantes nativos de español?_____

9. ¿Habla usted alguna lengua extranjera? _____Si es así, ¿a qué nivel _____

10. ¿Tiene algún contacto con hablantes de lengua inglesa? _____ ¿Con qué frecuencia?_____

11. ¿Ha pasado algún periodo de tiempo en algún país de habla inglesa? Seleccione la opción que corrresponda

a. Nunca

b. Periodo muy breve días o semanas

c. Menos de 3 meses

d. Entre 3 y 6 meses

e. Más de 6 meses

APPENDIX D: MOTIVATION QUESTIONNAIRE (adapted from Wen 2005)

Read each statement and circle the number that best reflects your opinion

1= Strongly disagree
2= Disagree
3= Uncertain
4= Agree
5= Strongly agree

1. Spanish is not important for me to learn about because it is not necessary in the world.

1……2……3…….4……5

2. I feel happy if people tell me that I have great pronunciation.

1……2……3…….4……5

3. The majority of my experience of learning Spanish has been pleasant.

1……2……3…….4……5

4. I like to watch TV, listen to the radio, listen to the songs or watch movies in Spanish.

1……2……3…….4……5

5. I want to pronounce as best as I can.

1……2……3…….4……5

6. Pronouncing like a NS of Spanish is important for me.

1……2……3…….4……5

7. I seldom watch or listen to any kind of Spanish programmes.

1……2……3…….4……5

8. Better pronunciation helps me in my career plan or study.

1……2……3…….4……5

9. I seek chances to speak Spanish.

1……2……3…….4……5

10. I pay careful attention to how Spanish speakers pronounce words.

1……2……3…….4……5

11. When somebody teases me for my pronunciation, I feel a little embarrassed but it's ok.

1……2……3…….4……5

12. I avoid speaking Spanish if somebody teases me for my pronunciation.

1……2……3…….4……5

13. If somebody teases me for my pronunciation, it won't discourage me from learning Spanish; on the contrary, I will do my best to improve in order to change their opinions.

1……2……3…….4……5

14. I like to imitate how Spanish speakers pronounce Spanish on TV, radio or movies.

1……2……3…….4……5

15. I would like to have Spanish speaking friends.

1……2……3…….4……5

16. I don't like to speak Spanish.

1……2……3…….4……5

17. I don't care about my pronunciation.

1……2……3…….4……5

18. I like to speak Spanish.

1……2……3…….4……5

APPENDIX E: ORAL MIMICRY TASK (adapted from Lord 2006)

Sentences (invented words in italics)

1- Se podía oír a través de la ventana el violento *pristonar* contra los cristales.
2- El pequeño *fatusal* desapareció rápidamente detrás de los árboles
3- Me gustaría ver el *blasofón* que tienes en la cocina después de la cena
4- Para ser un buen *blugón*, se necesita estudiar mucho
5- El *crestalar* hacía muy difícil que pudiéramos concentrarnos
6- Me dijo que el *sobredino* estaba a punto de llegar
7- Tuvo que devolver a la policía el *plautón* que encontró cerca de mi casa
8- Me dijo que tuviera cuidado con el *blaito* ya que era peligroso
9- El clarucio era lo único que los mantenía vivos
10- Abre el estubre que está sobre la mesa

APPENDIX F: SINGLE WORD INTELLIGIBILITY TEST

Target	Foil	V/C[1]	Minimal pair	Type[2]
piso	peso	VS	/i/–/e/	M
mimo	memo	VS	/i/–/e/	M
tila	tela	VS	/i/–/e/	M
misa	masa	VS	/i/–/a/	M/P
pila	pala	VS	/i/–/a/	M/P
pita	pata	VS	/i/–/a/	M/P
pico	poco	VS	/i/–/o/	M/P
tiro	toro	VS	/i/–/o/	M/P
mito	moto	VS	/i/–/o/	M/P
nido	nudo	VS	/i/–/u/	M
tina	tuna	VS	/i/–/u/	M
pipa	pupa	VS	/i/–/u/	M
pesa	pasa	VS	/e/–/a/	M/P
pelo	palo	VS	/e/–/a/	M/P
vela	bala	VS	/e/–/a/	M/P
pelo	polo	VS	/e/–/o/	P
seso	soso	VS	/e/–/o/	P
veto	voto	VS	/e/–/o/	P
mesa	musa	VS	/e/–/u/	M/P
dedo	dudo	VS	/e/–/u/	M/P
reto	ruta	VS	/e/–/u/	M/P
palo	polo	VS	/a/–/o/	M/P
sala	sola	VS	/a/–/o/	M/P
paso	poso	VS	/a/–/o/	M/P
cana	cuna	VS	/a/–/u/	M/P
mala	mula	VS	/a/–/u/	M/P
malta	multa	VS	/a/–/u/	M/P
oso	uso	VS	/o/–/u/	M
lona	luna	VS	/o/–/u/	M
bola	bula	VS	/o/–/u/	M
pisar	pesar	VU	/i/–/e/	M
pinada	penada	VU	/i/–/e/	M
pilón	pelón	VU	/i/–/e/	M
casi	casa	VU	/i/–/a/	M/P
mitad	matad	VU	/i/–/a/	M/P
literal	lateral	VU	/i/–/a/	M/P
timar	tomar	VU	/i/–/o/	M/P
mirar	morar	VU	/i/–/o/	M/P
firmado	formado	VU	/i/–/o/	M/P
ligar	lugar	VU	/i/–/u/	P

pintada	puntada	VU	/i/–/u/	P
anidar	anudar	VU	/i/–/u/	P
meses	mesas	VU	/e/–/a/	M/P
besar	basar	VU	/e/–/a/	M/P
alemanes	alemanas	VU	/e/–/a/	M/P
pesar	posar	VU	/e/–/o/	P
ternero	tornero	VU	/e/–/o/	P
pesado	posado	VU	/e/–/o/	P
temor	tumor	VU	/e/–/u/	M/P
retina	rutino	VU	/e/–/u/	M/P
perita	purita	VU	/e/–/u/	M/P
calor	color	VU	/a/–/o/	M/P
esposas	esposos	VU	/a/–/o/	M/P
pasaron	posaron	VU	/a/–/o/	M/P
amor	humor	VU	/a/–/u/	M/P
maleta	muleta	VU	/a/–/u/	M/P
sabido	subido	VU	/a/–/u/	M/P
osar	usar	VU	/o/–/u/	M
tornar	turnar	VU	/o/–/u/	M
romano	rumano	VU	/o/–/u/	M
vaina	vana	VC	/ai/–/a/	Dip/D
paisaje	pasaje	VC	/ai/–/a/	Dip/D
baile	vale	VC	/ai/–/a/	Dip/D
peina	pena	VC	/ei/–/e/	Dip/D
veinte	vente	VC	/ei/–/e/	Dip/D
reino	reno	VC	/ei/–/e/	Dip/D
maula	mala	VC	/au/–/a/	Dip/D
aula	ala	VC	/au/–/a/	Dip/D
augita	agita	VC	/au/–/a/	Dip/D
deudo	dedo	VC	/eu/–/e/	Dip/D
ceuta	ceta	VC	/eu/–/e/	Dip/D
neutro	netro	VC	/eu/–/e/	Dip/D
par	bar	C	/p/–/b/	V
peso	beso	C	/p/–/b/	V
pata	bata	C	/p/–/b/	V
pavor	favor	C	/p/–/f/	M
pez	fez	C	/p/–/f/	M
pino	fino	C	/p/–/f/	M
poda	moda	C	/p/–/m/	N
capa	cama	C	/p/–/m/	N
trapo	tramo	C	/p/–/m/	N
capa	cata	C	/p/–/t/	P
copo	coto	C	/p/–/t/	P

trapo	trato	C	/p/–/t/	P
peso	queso	C	/p/–/k/	P
poso	coso	C	/p/–/k/	P
paso	caso	C	/p/–/k/	P
alba	alma	C	/b/–/m/	N
vano	mano	C	/b/–/m/	N
bar	mar	C	/b/–/m/	N
brio	frío	C	/b/–/f/	V/M
vino	fino	C	/b/–/f/	V/M
vibra	fibra	C	/b/–/f/	V/M
vía	día	C	/b/–/d/	P
salvar	saldar	C	/b/–/d/	P
cava	cada	C	/b/–/d/	P
libar	ligar	C	/b/–/g/	P
robar	rogar	C	/b/–/g/	P
bruta	gruta	C	/b/–/g/	P
soldado	soltado	C	/d/–/t/	V
domar	tomar	C	/d/–/t/	V
boda	bota	C	/d/–/t/	V
cada	caza	C	/d/–/ɵ/	V
lado	lazo	C	/d/–/ɵ/	V
rada	raza	C	/d/–/ɵ/	V
toda	toga	C	/d/–/g/	P
dama	gama	C	/d/–/g/	P
vado	vago	C	/d/–/g/	P
torre	corre	C	/t/–/k/	P
toser	coser	C	/t/–/k/	P
tarro	carro	C	/t/–/k/	P
rata	raza	C	/t/–/ɵ/	M
tinta	cinta	C	/t/–/ɵ/	M
moto	mozo	C	/t/–/ɵ/	M
casa	gasa	C	/k/–/g/	V
coma	goma	C	/k/–/g/	V
vaca	vaga	C	/k/–/g/	V
roca	roja	C	/k/–/x/	M
sacar	sajar	C	/k/–/x/	M
coco	cojo	C	/k/–/x/	M
gota	jota	C	/g/–/x/	V
vago	bajo	C	/g/–/x/	V
hago	ajo	C	/g/–/x/	V
café	cacé	C	/f/–/ɵ/	P
forro	zorro	C	/f/–/ɵ/	P
rifar	rizar	C	/f/–/ɵ/	P

infecto	insecto	C	/f/–/s/	P
rifa	risa	C	/f/–/s/	P
gafas	gasas	C	/f/–/s/	P
mofa	moja	C	/f/–/x/	P
fusta	justa	C	/f/–/x/	P
rifa	rija	C	/f/–/x/	P
mazo	majo	C	/ɵ/–/x/	P
cocer	coger	C	/ɵ/–/x/	P
raza	raja	C	/ɵ/–/x/	P
coser	coger	C	/s/–/x/	P
casa	caja	C	/s/–/x/	P
masa	maja	C	/s/–/x/	P
oso	ocho	C	/s/–/tʃ/	M
viso	bicho	C	/s/–/tʃ/	M
asa	hacha	C	/s/–/tʃ/	M
mayo	macho	C	/ʝ/–/tʃ/	M
leyes	leches	C	/ʝ/–/tʃ/	M
haya	hacha	C	/ʝ/–/tʃ/	M
cayada	cañada	C	/ʝ/–/ ɲ/	N
ayo	año	C	/ʝ/–/ ɲ/	N
maya	maña	C	/ʝ/–/ ɲ/	N
chapa	papa	C	/tʃ/–/p/	M/P
choza	poza	C	/tʃ/–/p/	M/P
chico	pico	C	/tʃ/–/p/	M/P
pecho	peto	C	/tʃ/–/t/	M/P
chapa	tapa	C	/tʃ/–/t/	M/P
racha	rata	C	/tʃ/–/t/	M/P
choto	coto	C	/tʃ/–/k/	M/P
pecho	peto	C	/tʃ/–/k/	M/P
marcha	marca	C	/tʃ/–/k/	M/P
hucha	uña	C	/tʃ/–/ ɲ/	N
cacho	caño	C	/tʃ/–/ ɲ/	N
lecho	leño	C	/tʃ/–/ ɲ/	N
loma	lona	C	/m/–/n/	P
rama	rana	C	/m/–/n/	P
como	cono	C	/m/–/n/	P
amo	año	C	/m/–/ɲ/	P
lema	leña	C	/m/–/ɲ/	P
dama	daña	C	/m/–/ɲ/	P
sonar	soñar	C	/n/–/ɲ/	P
pena	peña	C	/n/–/ɲ/	P
una	uña	C	/n/–/ɲ/	P
bello	velo	C	/ʝ/–/l/	M/P

malla	mala	C	/ʝ/–/l/	M/P
llave	lave	C	/ʝ/–/l/	M/P
milla	mira	C	/ʝ/–/ɾ/	M/P
ralla	rara	C	/ʝ/–/ɾ/	M/P
molla	mora	C	/ʝ/–/ɾ/	M/P
valla	barra	C	/ʝ/–/r/	M/P
milla	mirra	C	/ʝ/–/r/	M/P
valle	barre	C	/ʝ/–/r/	M/P
pilla	piña	C	/ʝ/–/ɲ/	M
pella	peña	C	/ʝ/–/ɲ/	M
calla	caña	C	/ʝ/–/ɲ/	M
pelo	pero	C	/l/–/ ɾ/	M
tila	tira	C	/l/–/ ɾ/	M
tala	tara	C	/l/–/ ɾ/	M
pelo	perro	C	/l/–/r/	M
calo	carro	C	/l/–/r/	M
pala	parra	C	/l/–/r/	M
coro	corro	C	/r/–/ɾ/	M
caro	carro	C	/r/–/ɾ/	M
coral	corral	C	/r/–/ɾ/	M
cara	cada	C	/r/–/d/	M
toro	todo	C	/r/–/d/	M/P
poro	podo	C	/r/–/d/	M/P
barba	baba	C	/r/–Ø	D
sarna	sana	C	/r/–Ø	D
barbero	babero	C	/r/–Ø	D

1 Stressed vowel (VS) , unstressed vowel (VU), Vowel combination (VC) and consonant (C) error

2 Manner of articulation (M), point of articulation (P), voicing (V), deletion (D), diphthong (Di) and nasality (N)

APPENDIX G: SPANISH VERSION OF THE HARVARD PSYCHOACOUSTIC SENTENCES (Egan 1948; Valero 1991)

Series 1

La oscuridad me da mucho miedo.
Podéis venir con nosostros al cine.
Se ha roto la tapa del bote de mermelada.
Ganará lo que ganaba en su anterior trabajo.
Carlos es capaz de no saludarnos.
Se tiene por forzuda, pero no resiste nada.
Las llaves están en mi bolsa.
Pili quiere que tú y yo cojamos zarzamoras.
Le apatece un poco de vino blanco.
Las notas de diseño que te faltan no las sabrás por mí.

Series 2

He comido carne con patatas fritas.
El balazo se desvió de la diana.
Se pegaron en medio de la curva.
Esta sustancia nos prolonga la memoria.
Recoge su llavero y pónselo dentro del cesto.
Las oscuras pestañas enmarcan su dulce mirar.
Una de las redes ya está en mi barco pesquero.
Te vencí y di la copa al orfelinato.
Beber anís puede quemarte el gaznate.
Pon mis discos y cállate.

Series 3

Las hormigas se han comido las rosas.
Le envié un dinero por carta certificada.
Las vallas de metal destrozan el puente.
Yo quiero un poco más de jabón.
El duro capataz nos riñó bastante.
Carlos no ha pagado el mes de enero.
Lucía nos critica siempre a David.
Los techos blancos suelen necesitar pintura.
Me enfadé porque llegó mal vestido.
Lázaro casi anda más pronto que gatea.

Series 4

Cuando te vaya bien, llámame.
Los roperos son pequeños.
La tila te calma y duermes mejor.
Si no cenas tendrás dolor de barriga.
La pieza que no funciona es del carburador.
No tienes que soportarle sus caprichos.
Las tazas de café no están dónde siempre.
Olga no tiene valor para quedarse.
Si la luz te incomoda, te vas a dormir.
Su nieta Lola también sabe polaco.

Series 5

Puedes fumar, pero vete al balcón.
La tía Carmen quería guisar el conejo.
No des patadas a los rosales.
Necesito un centímetro para medir las dos telas.
Con un quilo de boniatos ya tengo de sobra.
Quizás pueda volver a llamar mañana por la noche.
Si quieres rezar, te dejo solo.
Nunca debí dinero ni lo tomé prestado.
Sus cacerías acabaron con la fauna de la zona.
Me diste la paga de este mes.

APPENDIX H: PHONETICALLY-BALANCED TEXT (Ortega, González and Marrero 2000)

Hay algo ahí, en el aire, que cambia el sentido de las cosas. Ese viento suave vuela, te toca la cara, mientras cuentas las hojas de los árboles. El agua corre buscando los campos. Al abrir las puertas de mi casa pienso: este país, una mañana más.

A mi edad, comienzan a faltarme las fuerzas, ya casi no soy joven, y la muerte de mi mujer en la guerra me pesa mucho. Cuando el cuerpo llega a esa hora, la ciencia de los doctores no logra detener el paso del tiempo.

De niño, allá en mi tierra, solía pasarme los días revolviendo de un lado a otro. Poco a poco, los coches de la ciudad fueron llamando mi atención. Mi madre decía que tuviera cuidado, pero yo me creía muy mayor, así que no tenía ni interés ni tiempo para mi propio signo.

Pero sigo, es cierto, cuántas cosas buenas encontré entre su gente. Si cuento los queridos veranos de entonces, no son siete, ni nueve, ni veinte. Debe ser que soy niño de nuevo en este cuerpo triste.

APPENDIX I: SEMI-SPONTANEOUS PRODUCTION TASK (sentences taken from the Spanish version of the Harvard Psychoacoustic Sentences, Valero 1991)

Transform the underlined subjects and verbs from singular into plural or vice versa and utter the result out loud.

Series 1

Nunca se interesó por su salud.
Tiende esos pantalones en el balcón.
Mis zapatos rojos llevan cordones.
Mi carta debería llegar el día diez.
La mamá de Laura corta el pescado.
La fábrica cerró los domingos y festivos.
No les permita cazar codornices.
Mañana taparé los agujeros con yeso.
Todavía se pelea mucho con Toni.
La dorada es un pescado bastante meloso.

Series 2

En mayo te daré el resultado
Regresó cinco años más tarde.
Nunca bebió más de la cuenta.
Tenía el pelo liso como la madre
Percibí todo pero callé.
Son las tres de la tarde.
Fuertes sonidos empezaron a asustarnos.
Cavé en el jardín al amanecer.
Cociné el pavo con dátiles.

Series 3

Recoge tus cosas y vete.
Andrea no puede venir esta mañana.
Los lazos largos me parecen ridículos.
Toda la tarde permaneció con Silvia.
Falló en cada una de las pruebas.
Pedro mató el toro sin el estoque.
Mi novio caza pero no pesca.
La instalación del gas se ha estropeado.
Pilar dirige la limpieza de la casa.
Ya no me darás nunca más la lata.

Series 4

Ella se fiaba de él, no le daba ningún miedo.
A los niños hay que contestarles a lo que te preguntan.
Su madre parecía muy despistada.
Se hizo esta peluca de rizos caoba.
Los críos llamaban Nela a la tejedora.
Los premios tocaron en Zamora.
Todos cantaron por soleares.
Le daré un ron doble.
Quizás ya te permitan cavar donde estuvo Félix.
Las ovejas pacen mientras el sol quema.

Series 5

La lluvia, gris, constante, no paró en todo el día.
El cachorro despedazó las cortinas de la casa.
Humedece los pantalones con este material.
Esta carta va dirigida al juez supremo.
La frecuencia del sondeo nos informa acerca de la opinión global.
Tenías sesenta años cuando yo llegué.
El crimen parece debido a un ataque de celos.
Sabía sobre mi padre cosas prometedoras.
La prisión rozaba los muros del castillo.
Nunca me presentaste a Berta.

APPENDIX J: GROUP DIFFERENCES (GENDER, L1 AND NATIVE VERSUS NON-NATIVE SPANISH SPEAKERS) IN INTELLIGIBILITY SCORES

Figure J.1: Gender differences in intelligibility scores

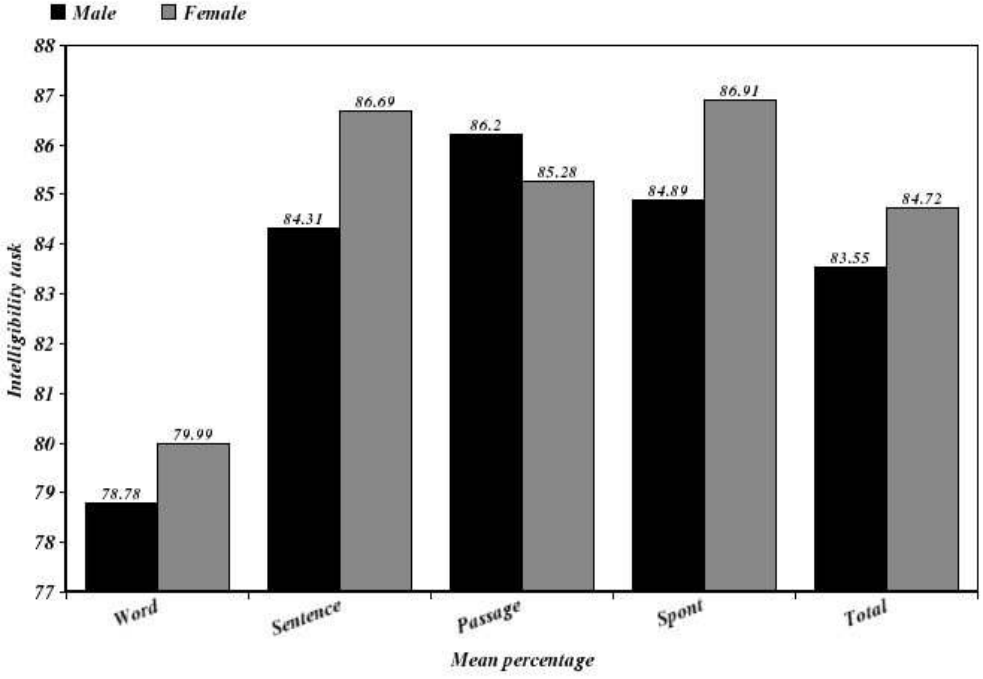


Figure J.2: L1 differences in intelligibility scores

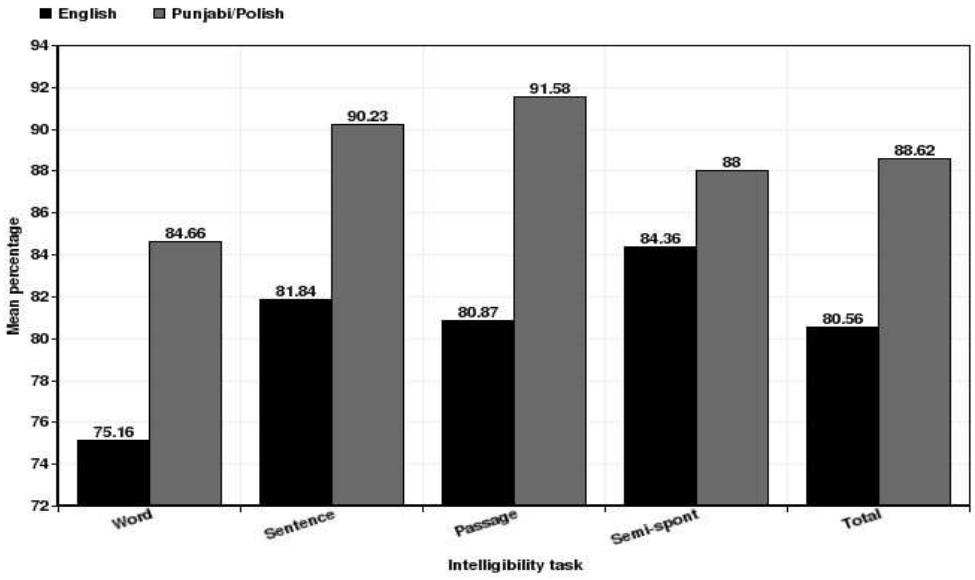
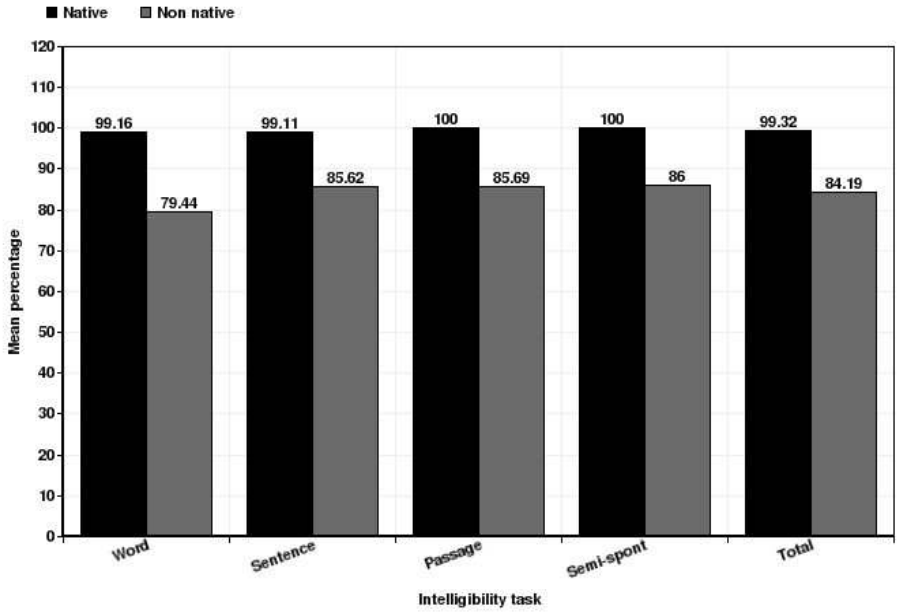


Figure J.3: Native versus non native differences in intelligibility scores



APPENDIX K: GROUP DIFFERENCES (GENDER, L1 AND NATIVE VERSUS NON-NATIVE SPANISH SPEAKERS) IN SCORES ON EIGHT ERROR CATEGORIES (SINGLE WORD TEST)

Figure K.1: Gender differences in scores on eight error categories

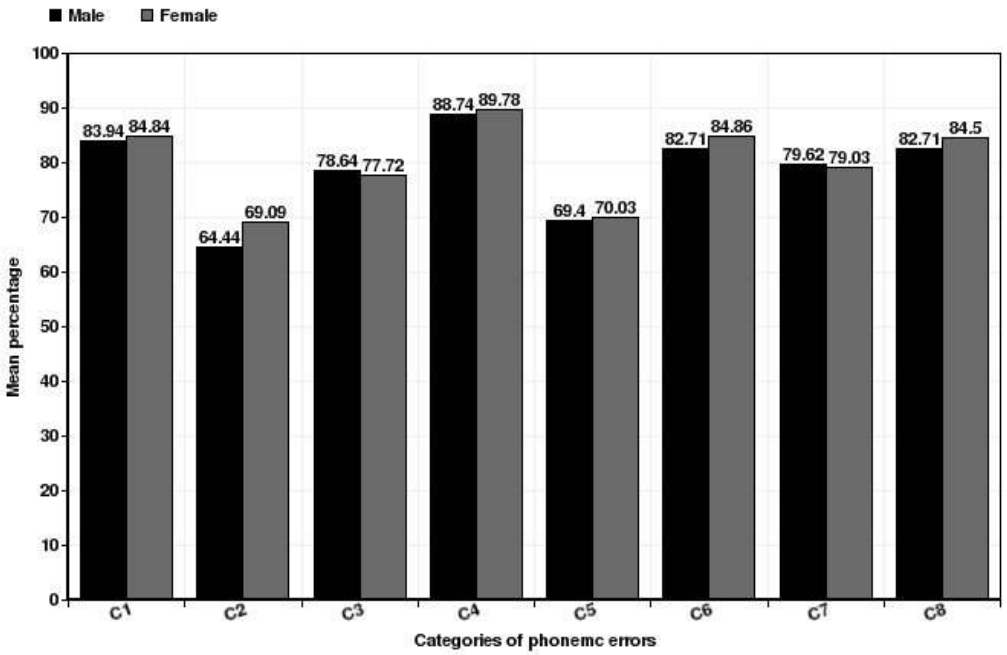


Figure K.2: L1 differences in scores on eight error categories

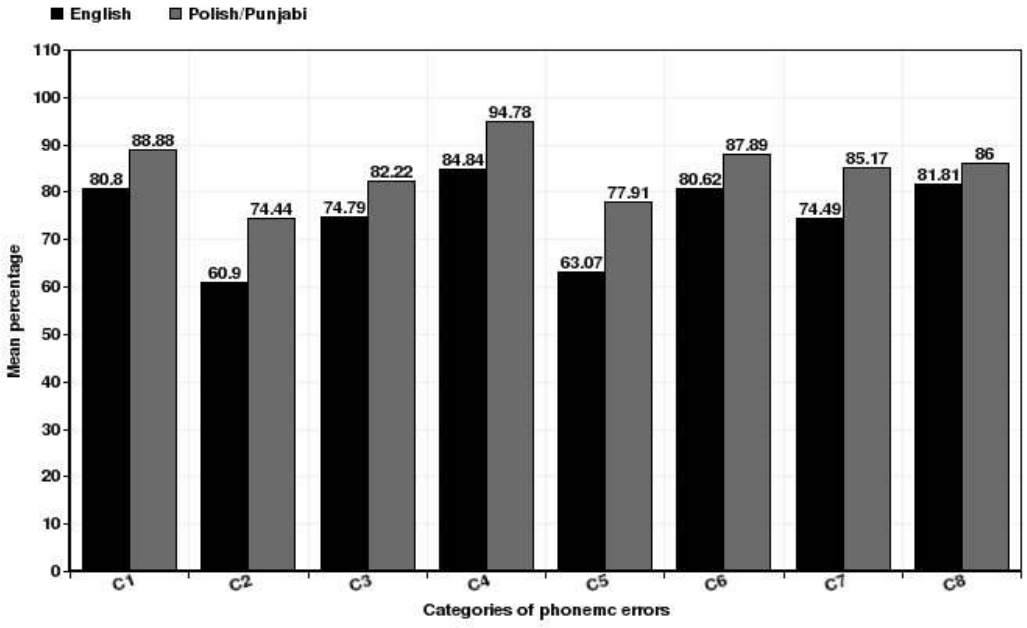
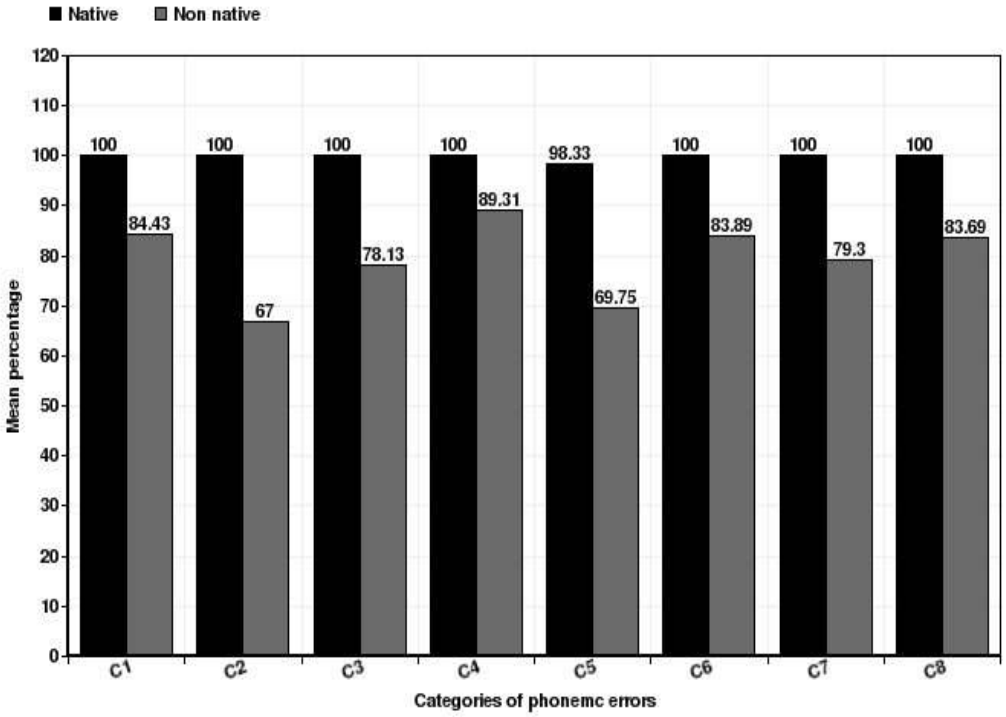


Figure K.3: Native versus non native differences in scores on eight error categories



APPENDIX L: ERROR PROFILE PER SPEAKER (SINGLE WORD INTELLIGIBILITY TEST)

Figure L.1: Error profile (Speaker 1)

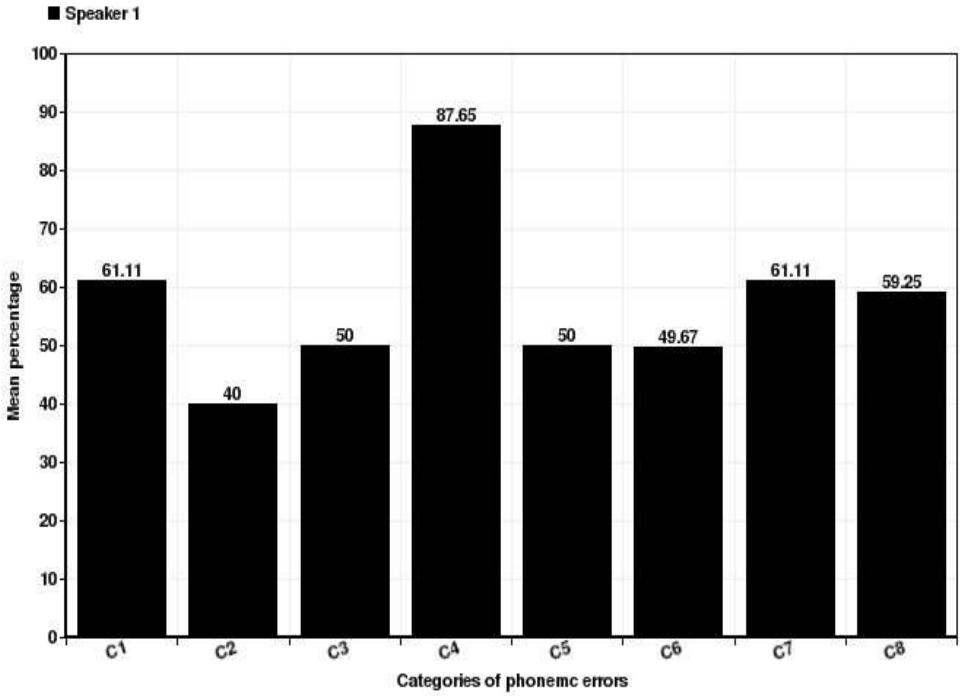


Figure L.2: Error profile (Speaker 2)

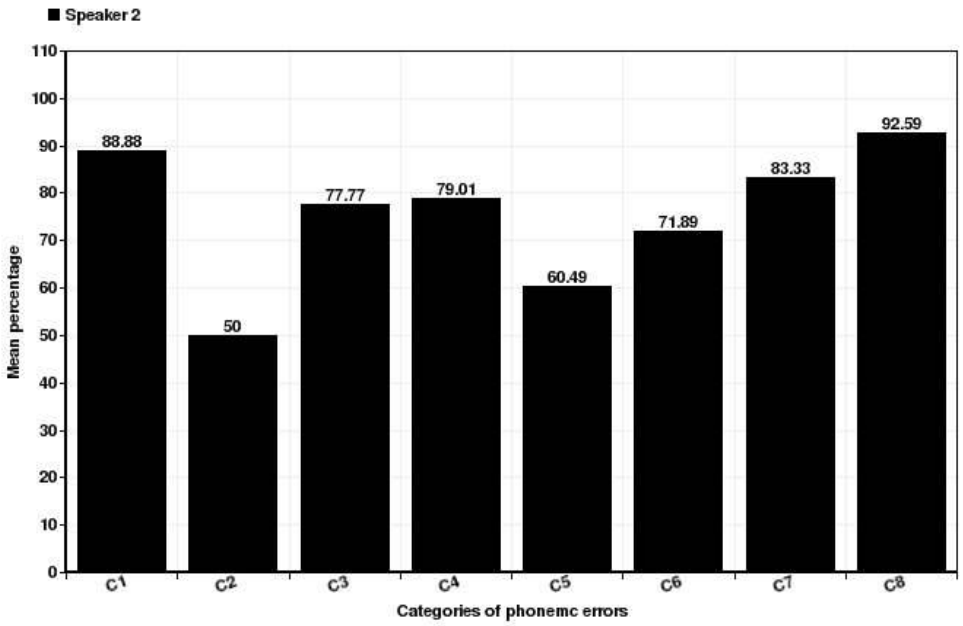


Figure L.3: Error profile (Speaker 3)

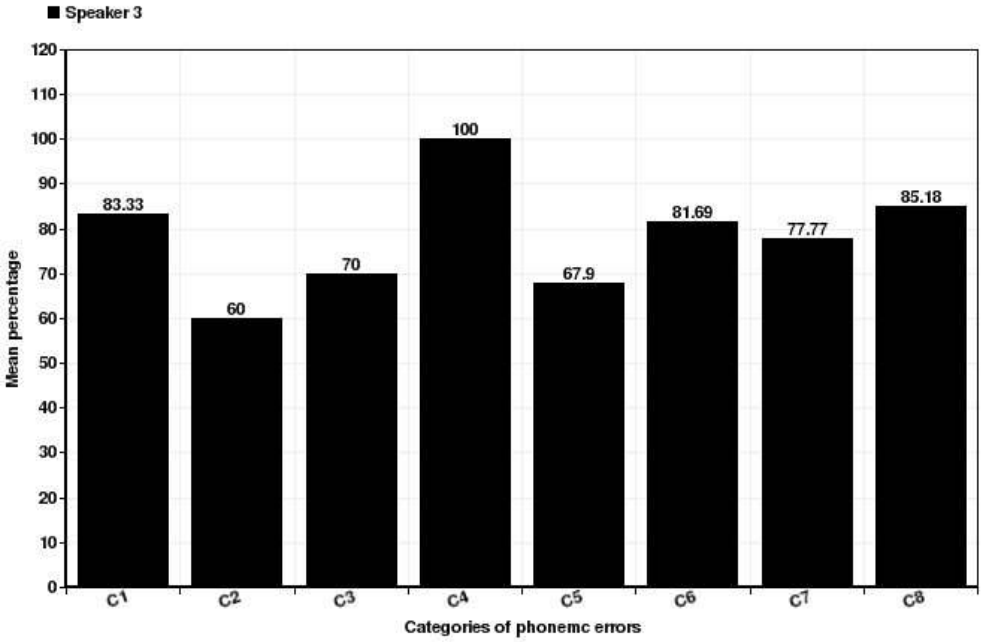


Figure L.4: Error profile (Speaker 4)

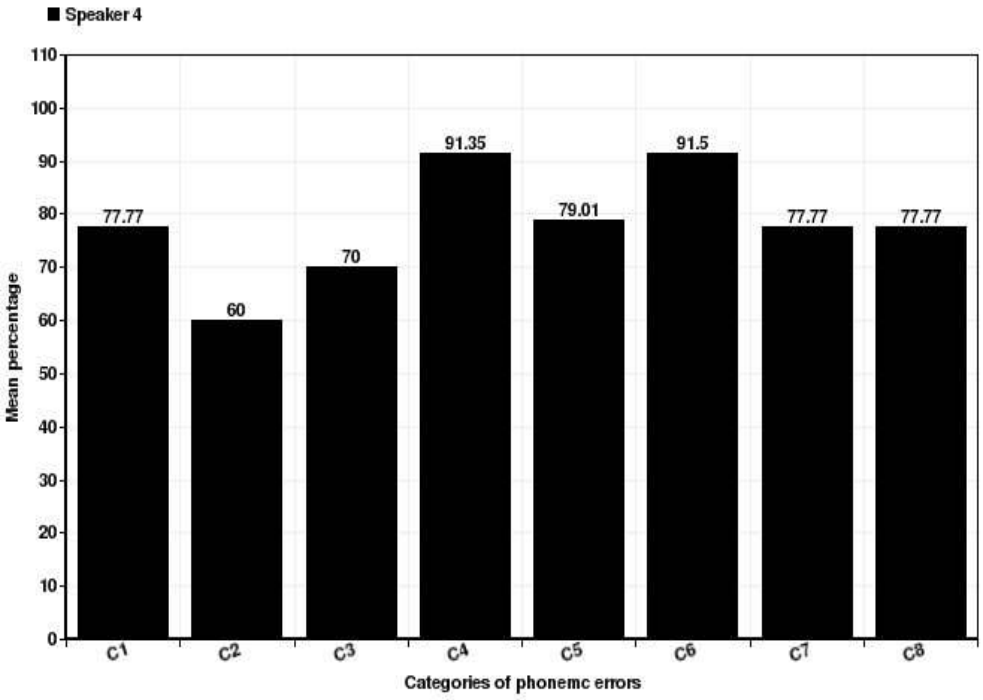


Figure L.5: Error profile (Speaker 5)

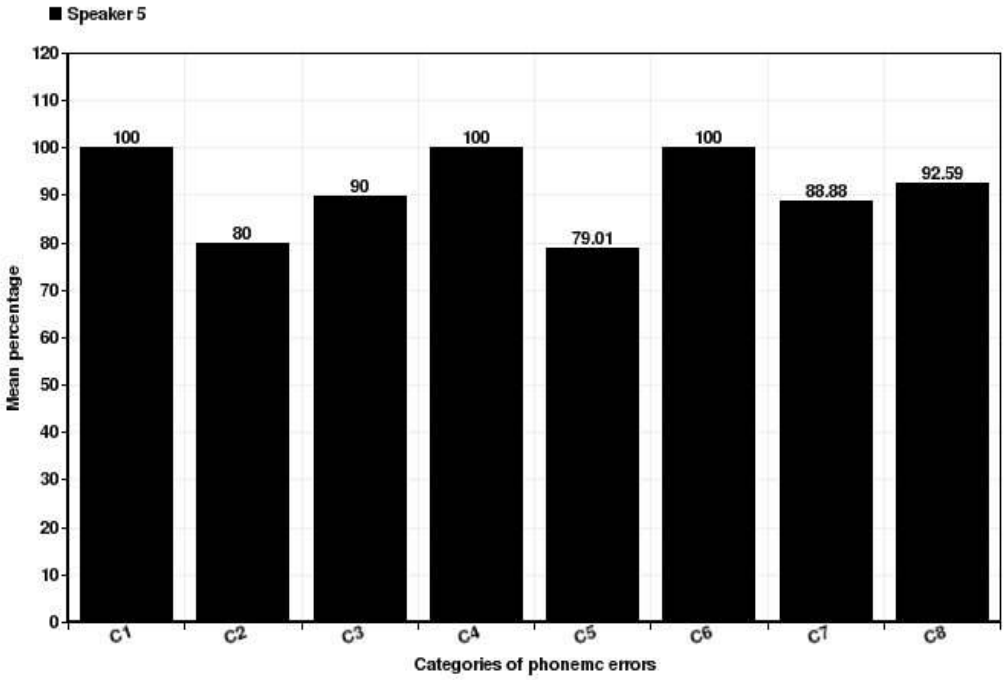


Figure L.6: Error profile (Speaker 6)

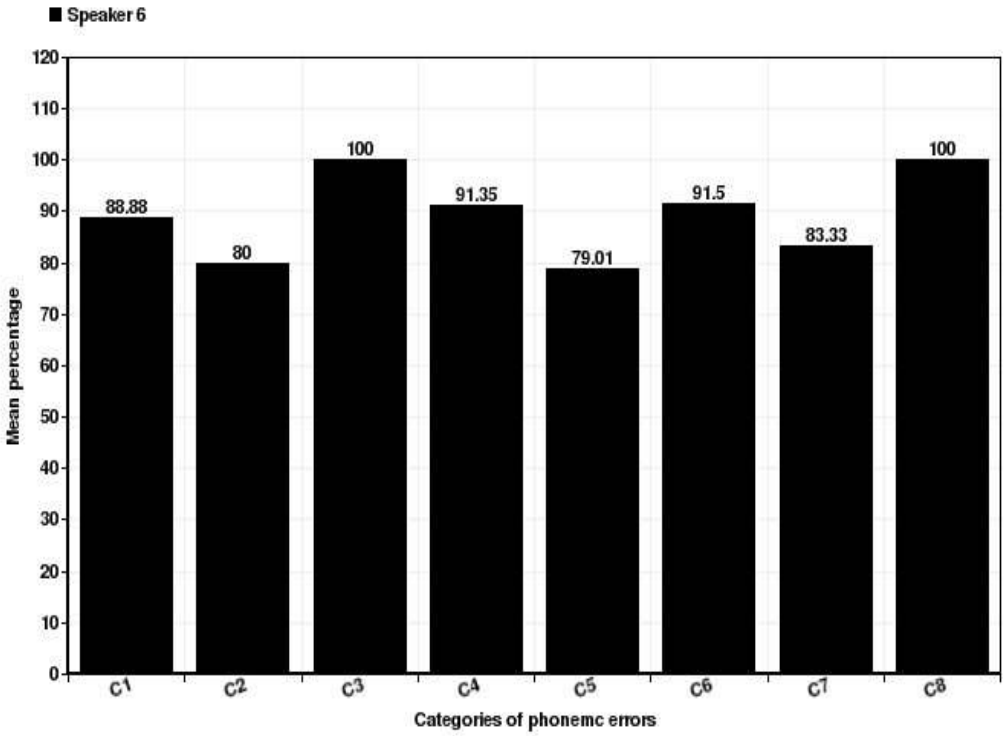


Figure L.7: Error profile (Speaker 7)

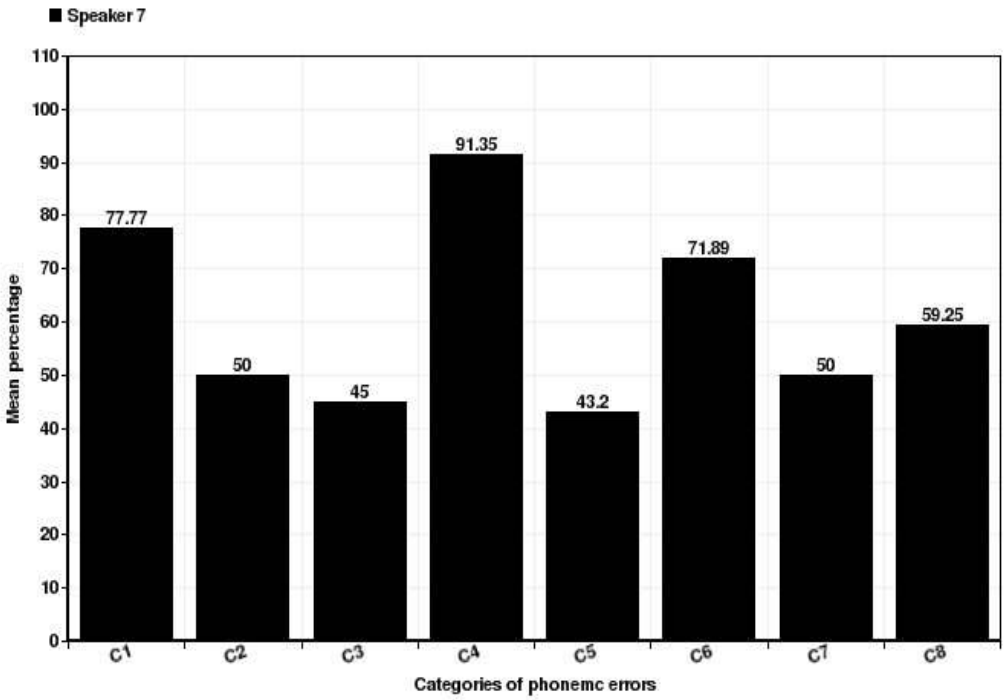


Figure L.8: Error profile (Speaker 8)

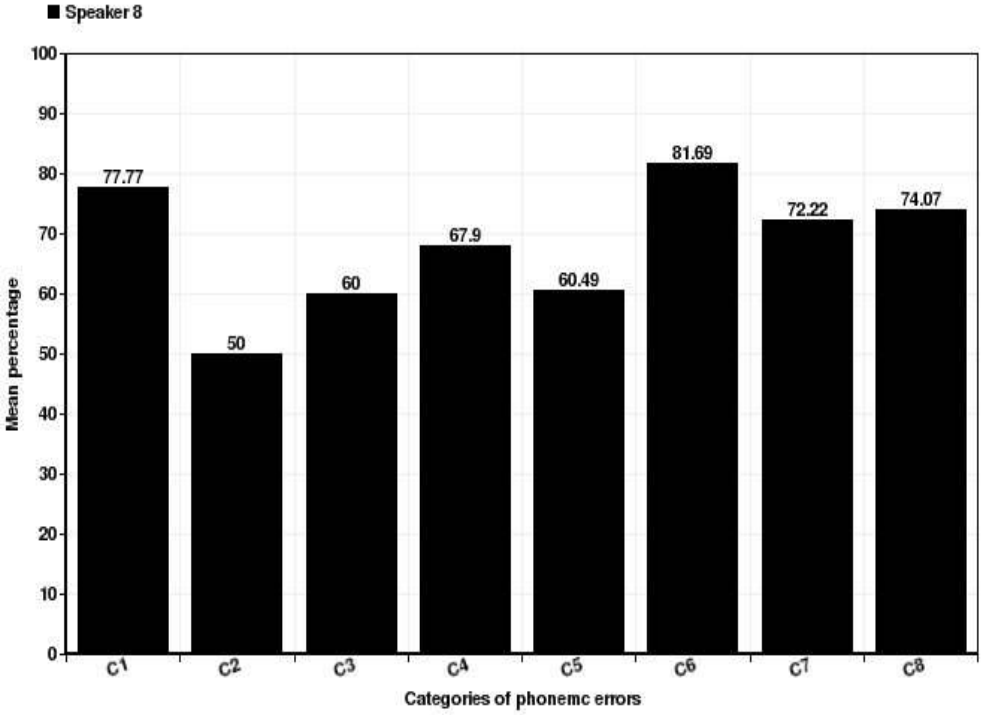


Figure L.9: Error profile (Speaker 9)

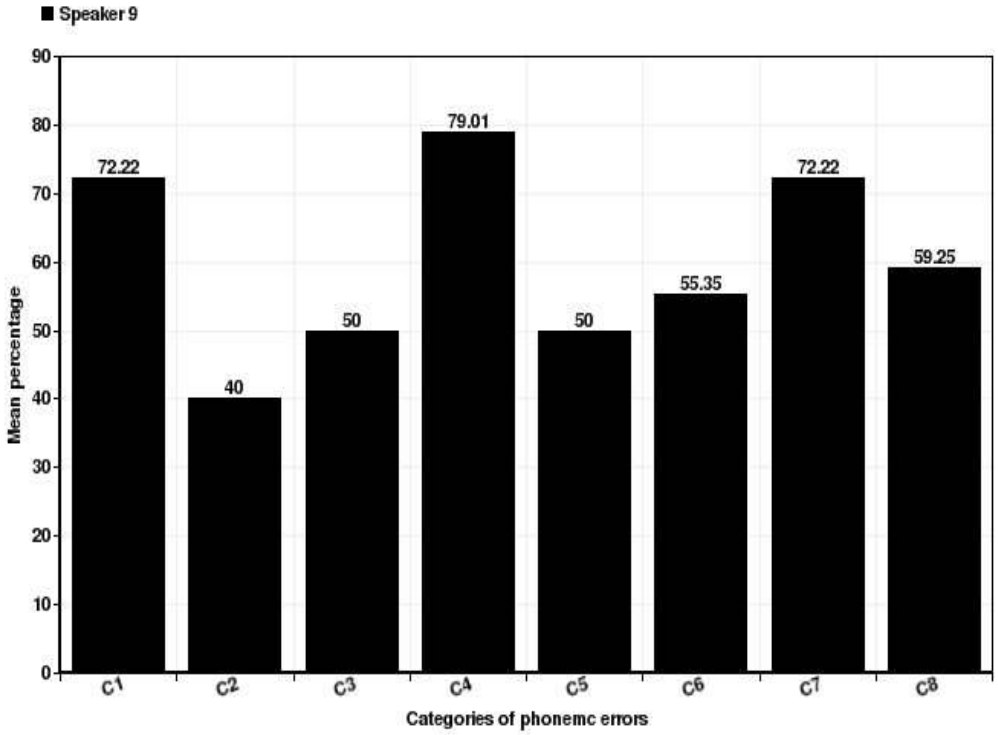


Figure L.10: Error profile (Speaker 10)

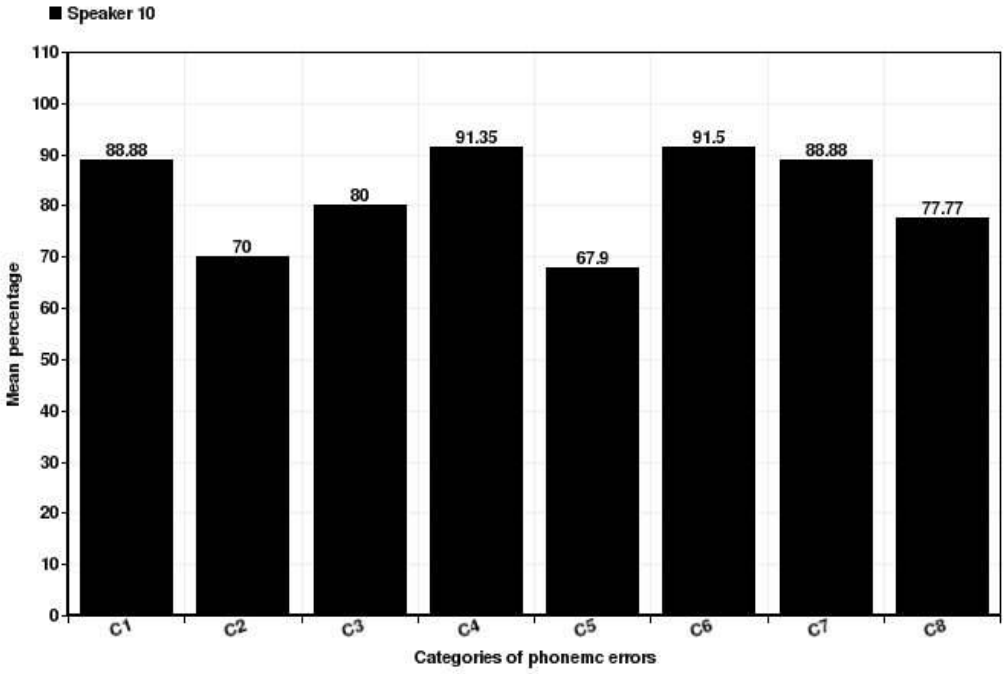


Figure L.11: Error profile (Speaker 11)

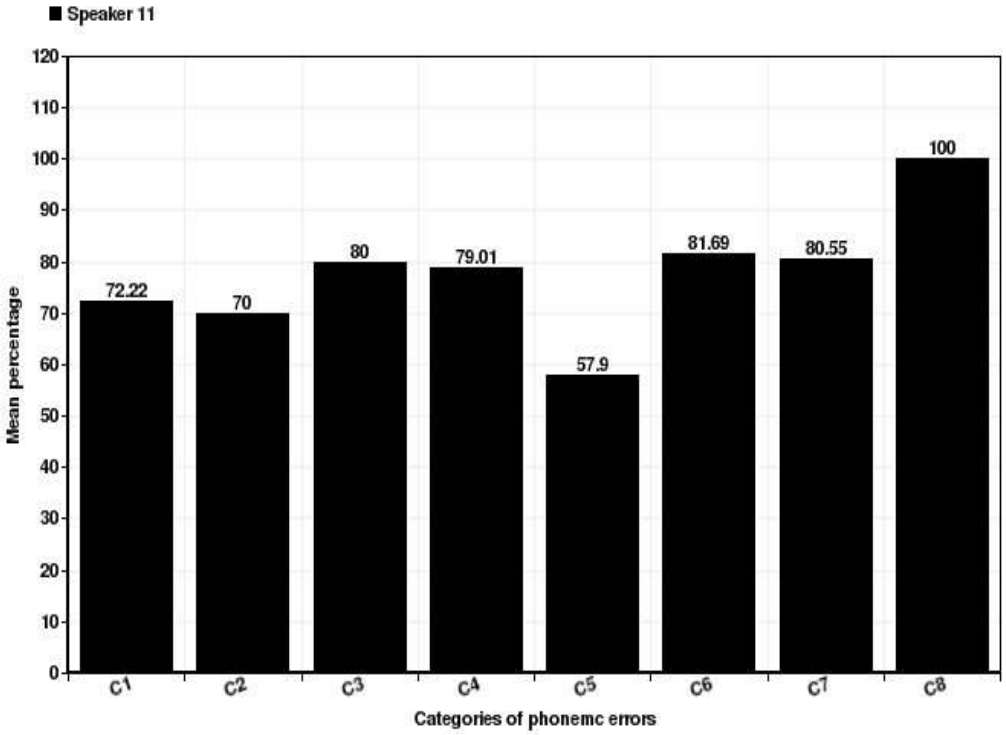


Figure L.12: Error profile (Speaker 12)

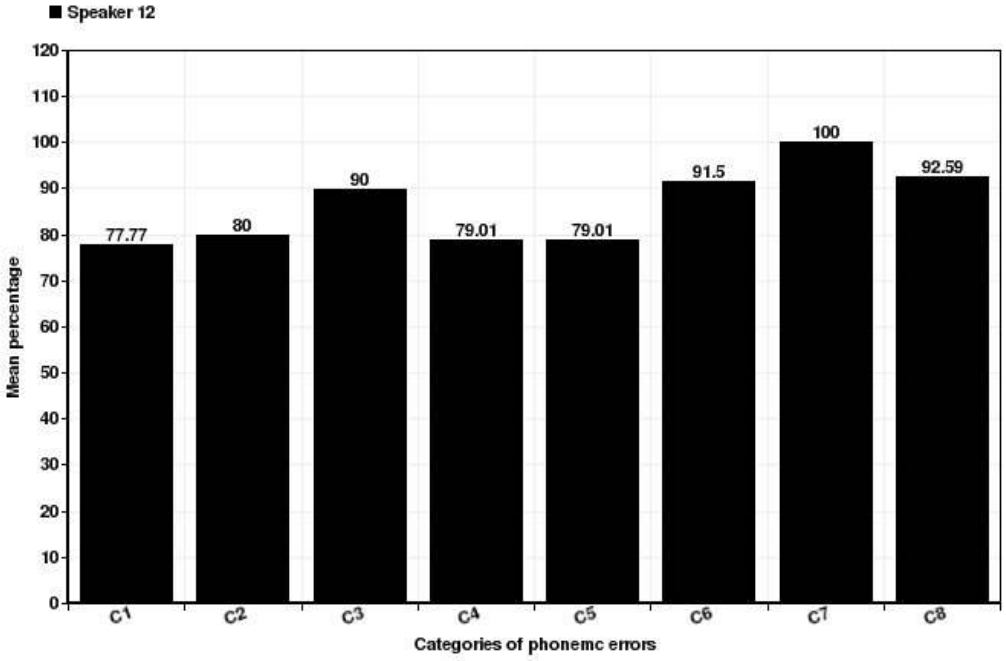


Figure L.13: Error profile (Speaker 13)

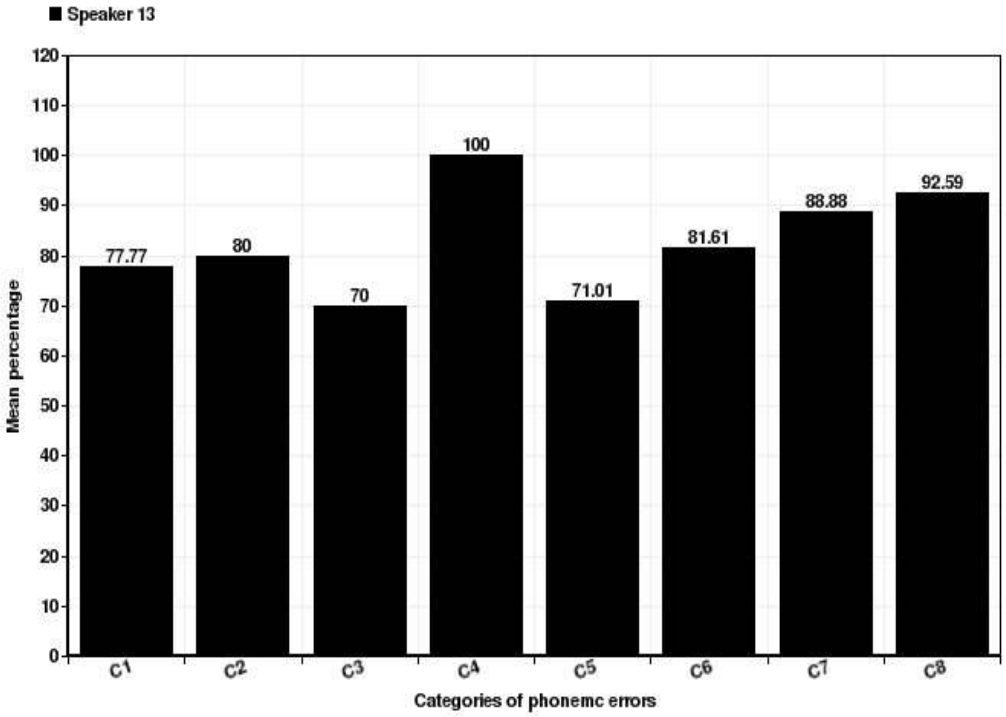


Figure L.14: Error profile (Speaker 14)

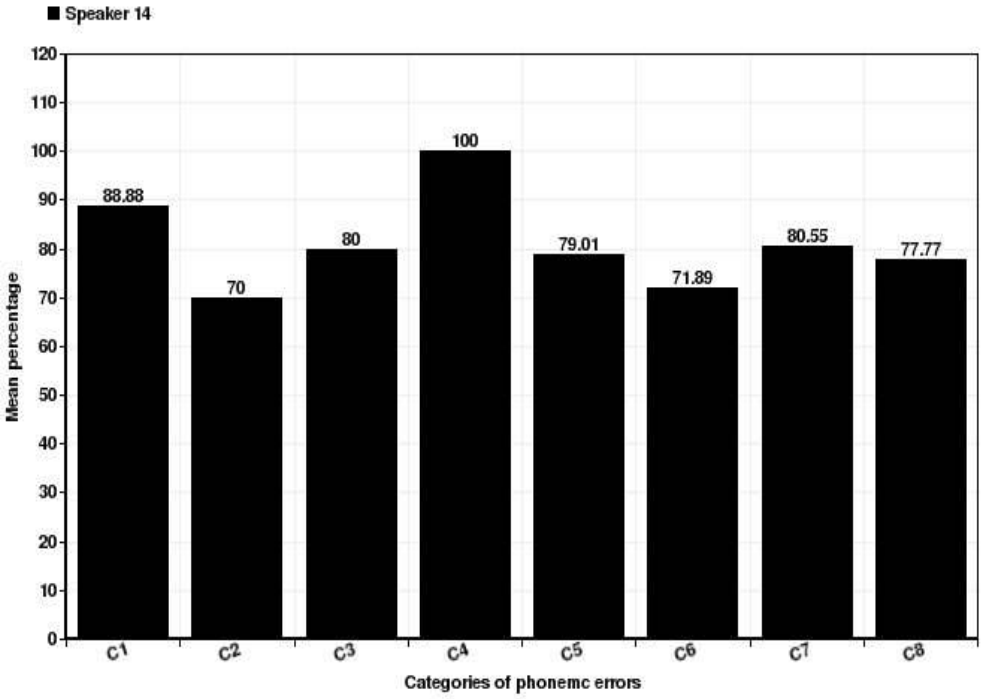


Figure L.15: Error profile (Speaker 15)

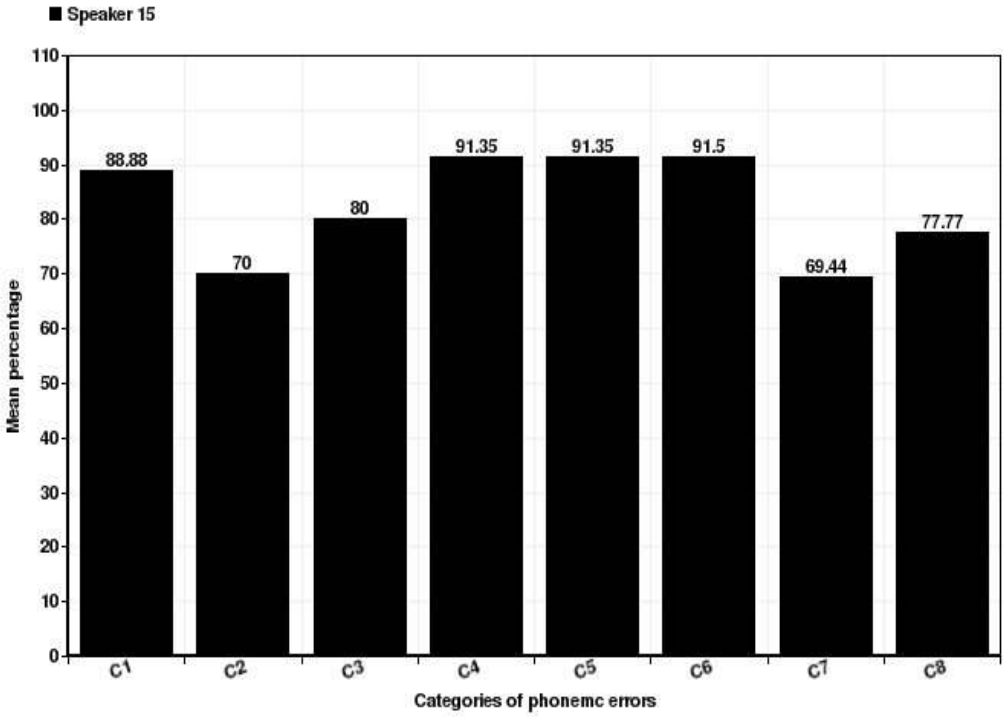


Figure L.16: Error profile (Speaker 16)

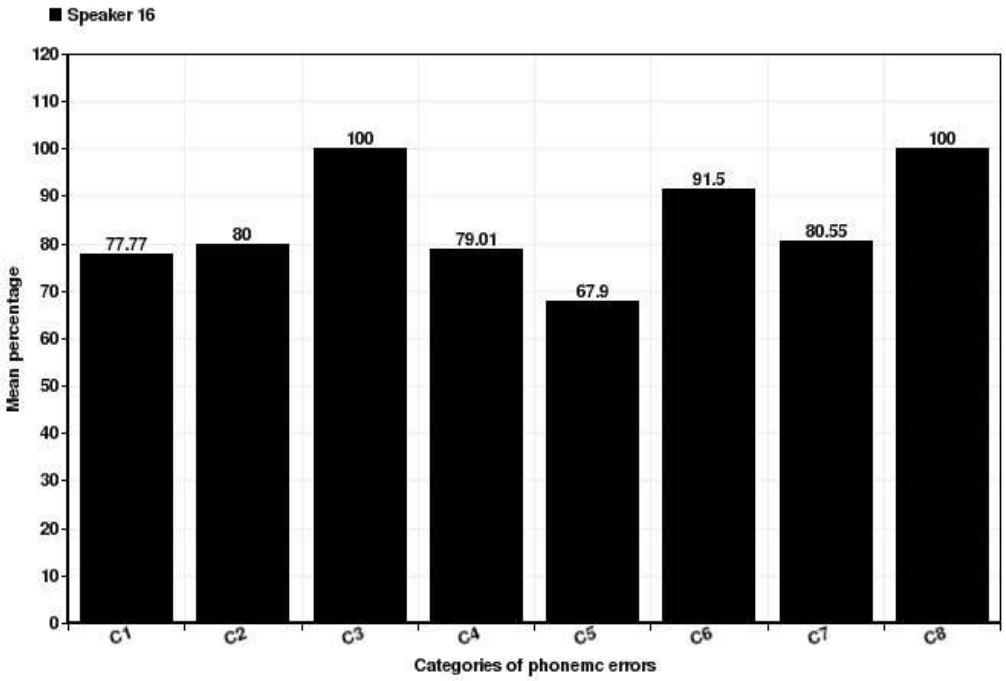


Figure L.17: Error profile (Speaker 17)

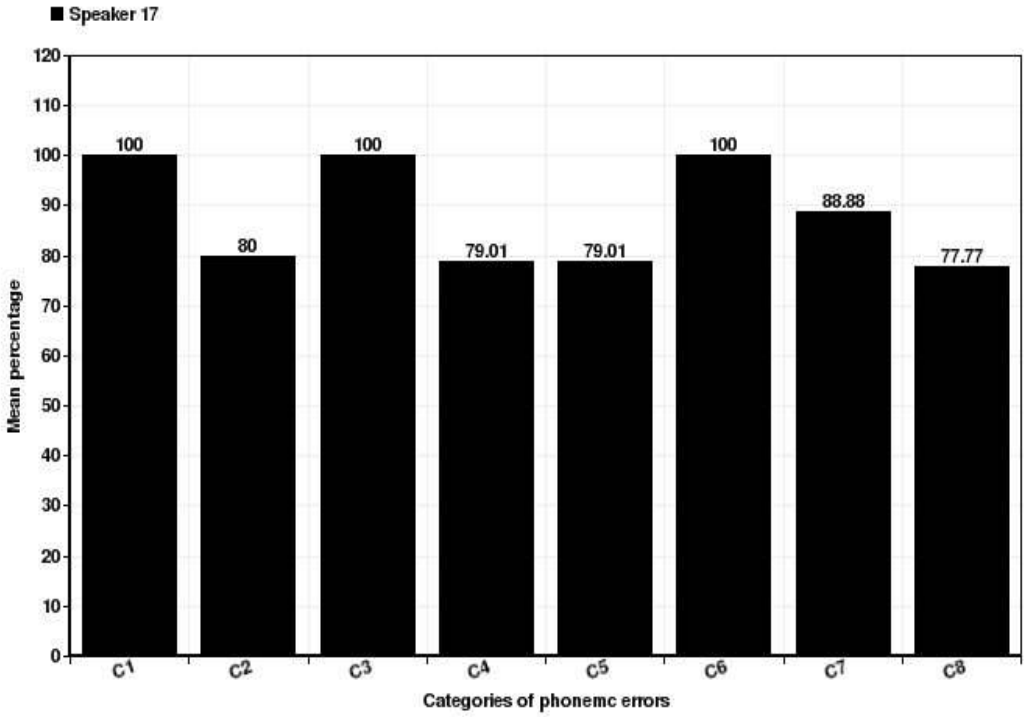


Figure L.18: Error profile (Speaker 18)

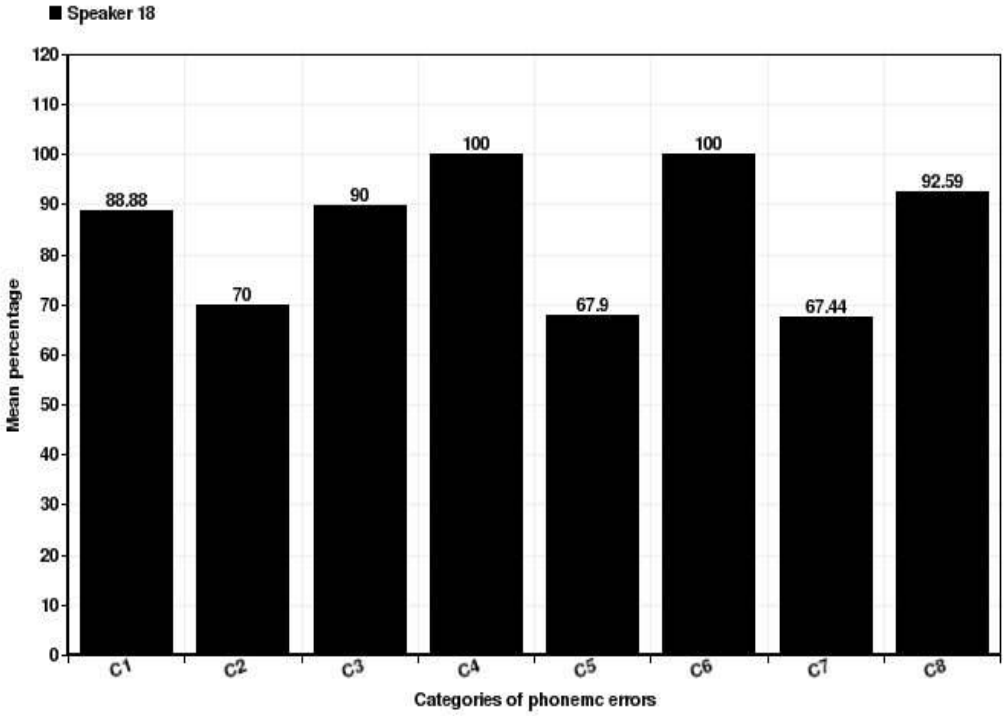


Figure L.19: Error profile (Speaker 19)

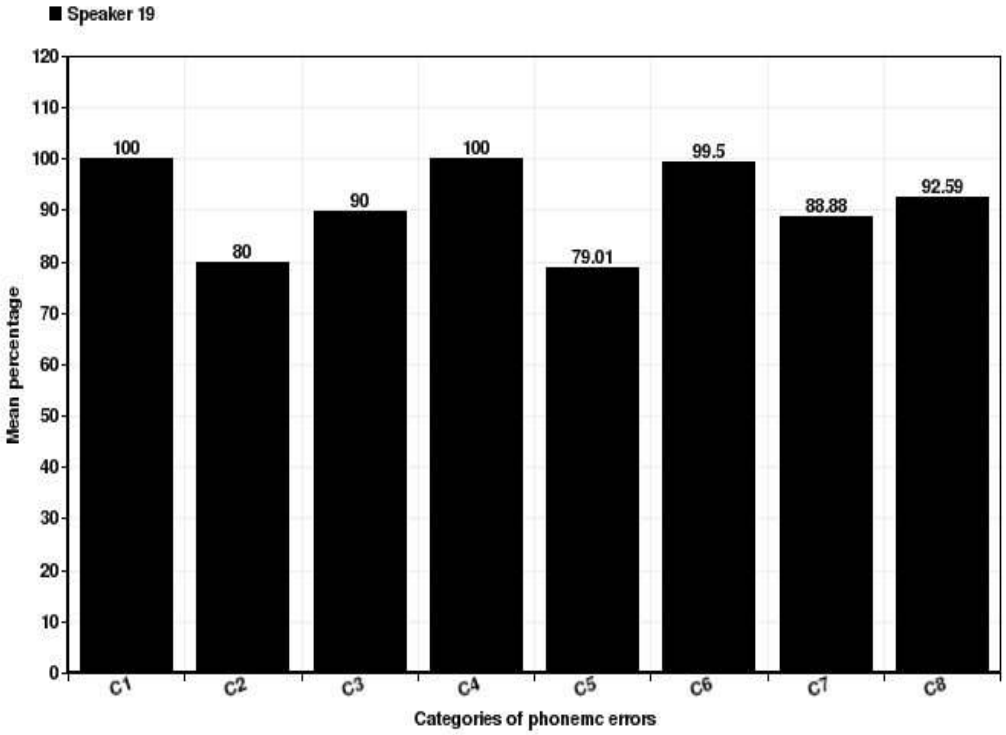
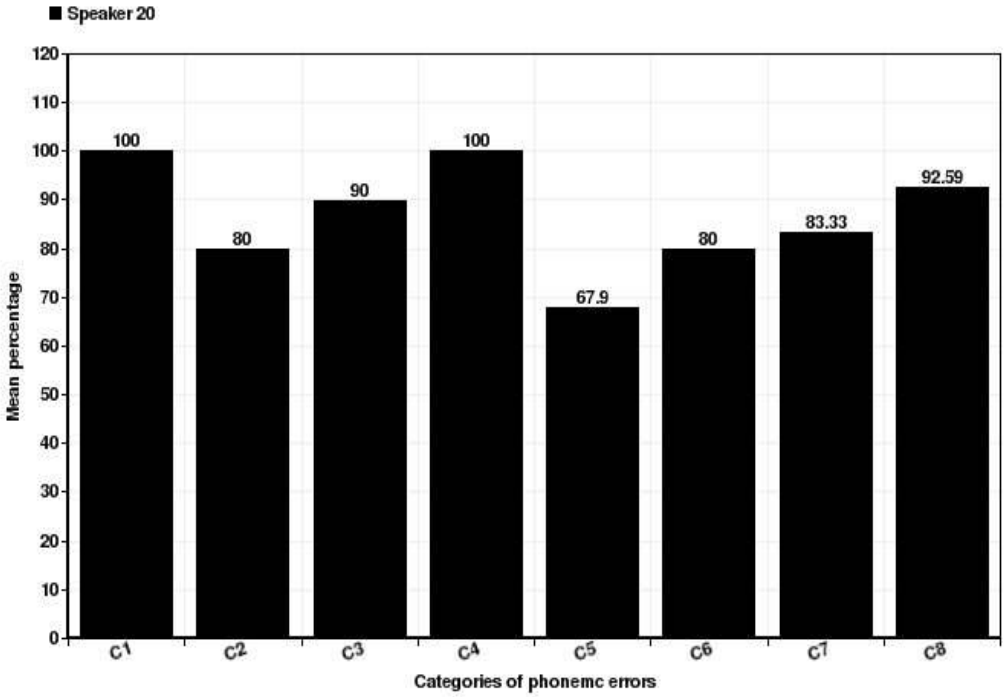
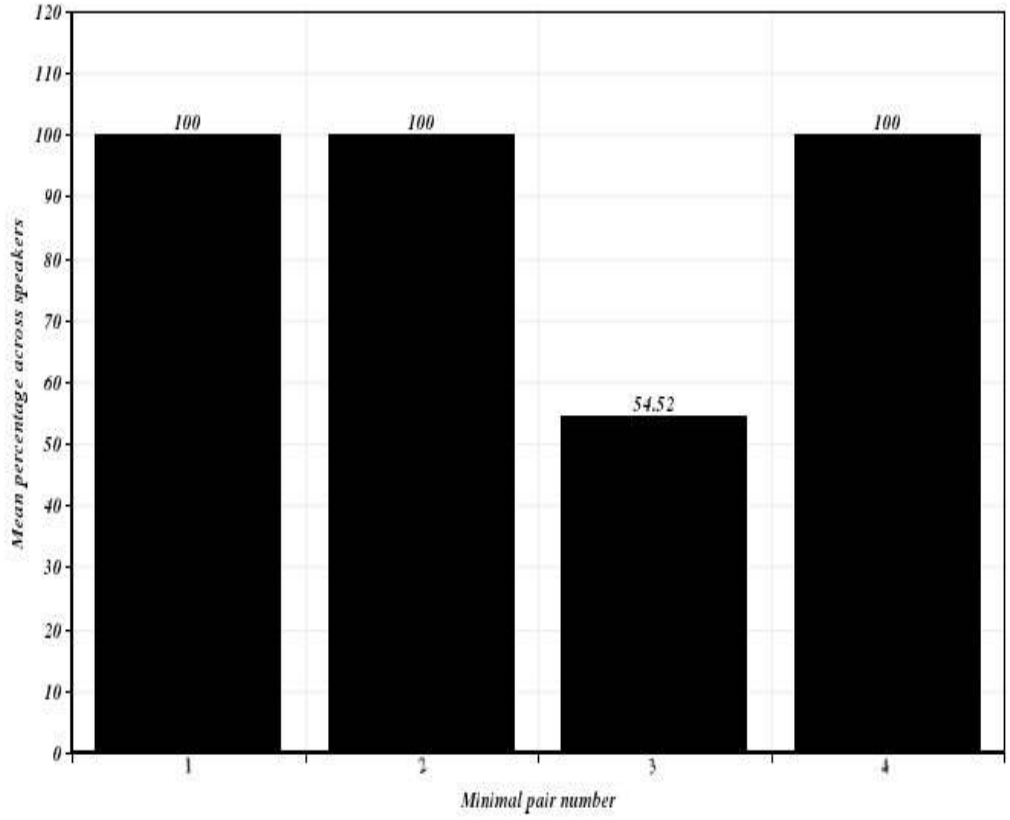


Figure L.20: Error profile (Speaker 20)



APPENDIX M: SCORES ON MINIMAL PAIRS FOR EACH ERROR CATEGORY ACROSS SPEAKERS (SINGLE WORD INTELLIGIBILITY TEST)

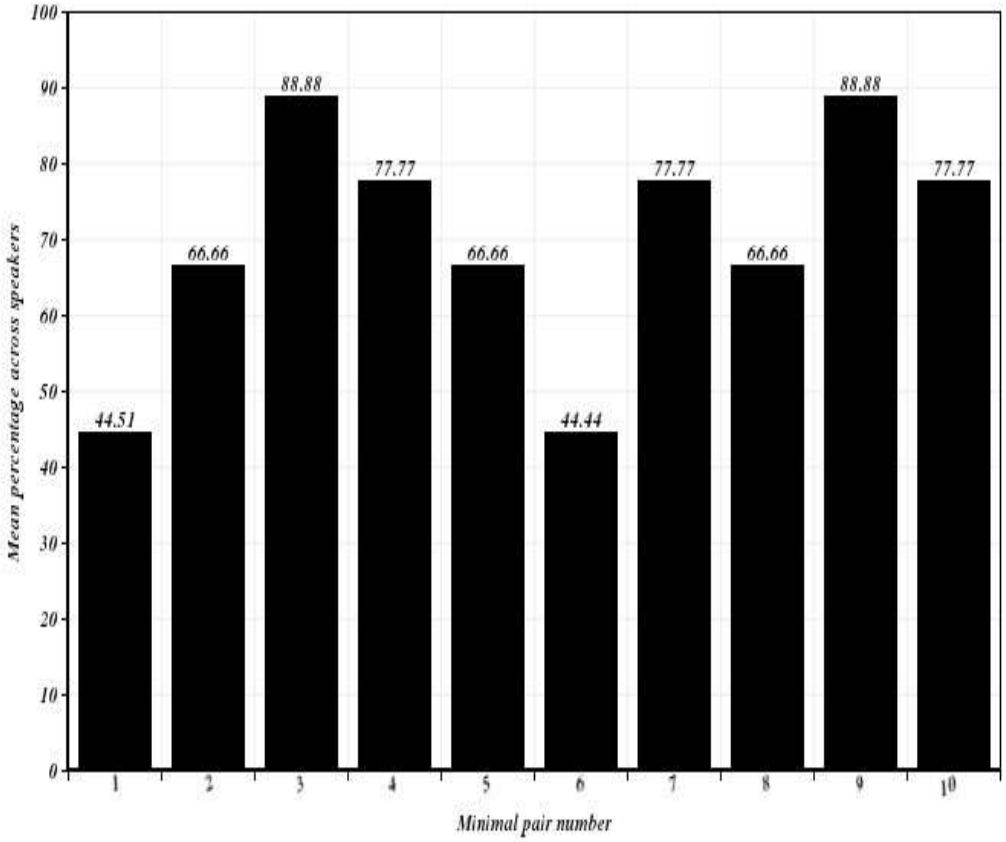
Figure M.1: Scores on minimal pairs across speakers (africates)



Minimal pairs number (affricates):

1-/tʃ/-/p/
2-/tʃ/-/t/
3-/tʃ/-/k/
4-/tʃ/-/ ɲ/

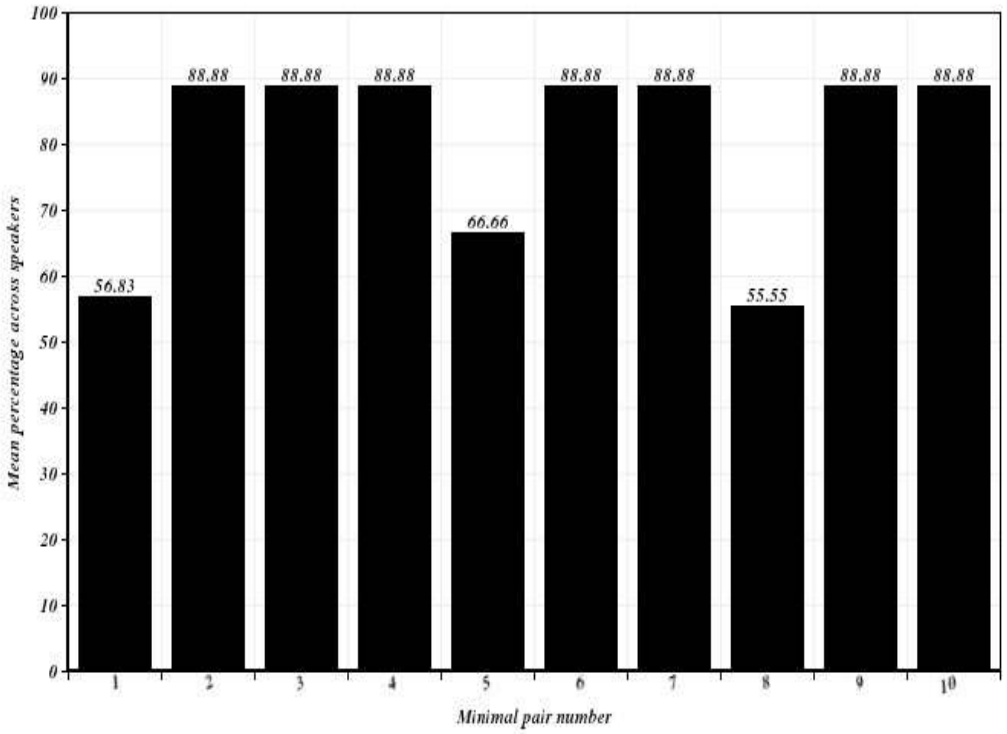
Figure M.2: Scores on minimal pairs across speakers (unstressed vowels)



Minimal pair number (unstressed vowels):

1-/i/-/e/
2-/i/-/a/
3-/i/-/o/
4-/i/-/u/
5-/e/-/a/
6-/e/-/o/
7-/e/-/u/
8-/a/-/o/
9-/a/-/u/
10-/o/-/u/

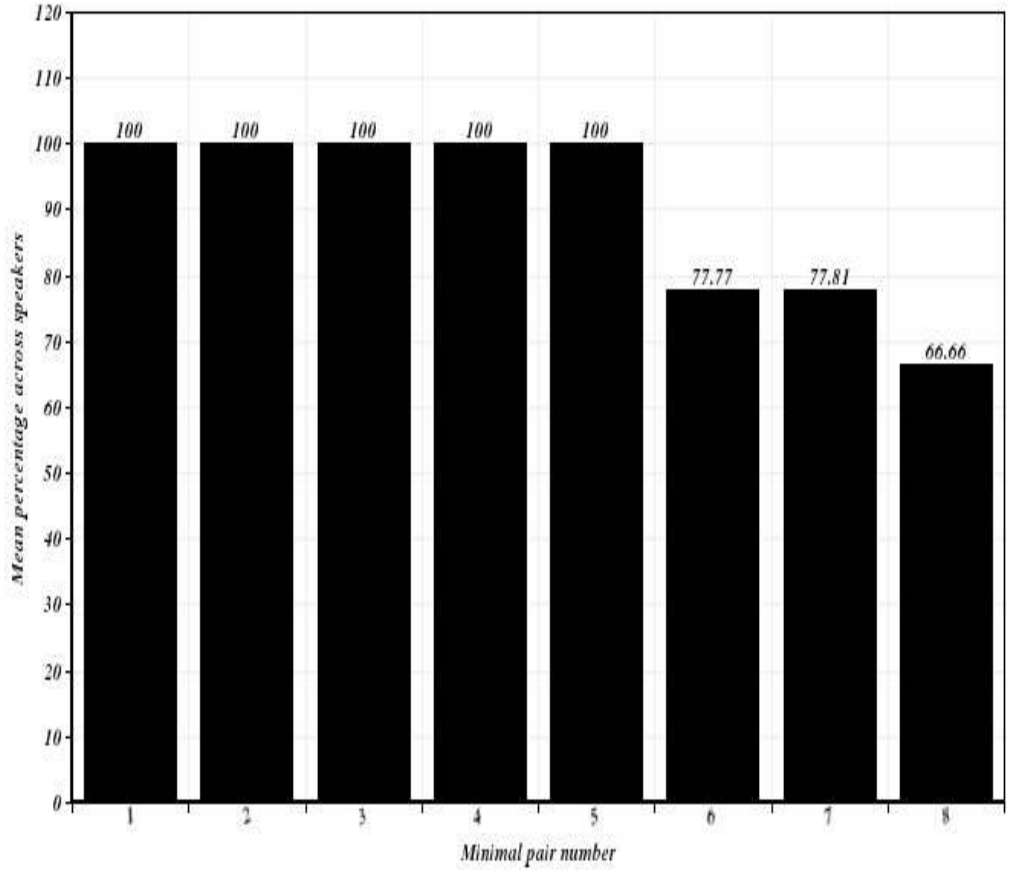
Figure M.3: Scores on minimal pairs across speakers (stressed vowels)



Minimal pair number (stressed vowels):

1-/i/-/e/
2-/i/-/a/
3-/i/-/o/
4-/i/-/u/
5-/e/-/a/
6-/e/-/o/
7-/e/-/u/
8-/a/-/o/
9-/a/-/u/
10-/o/-/u/

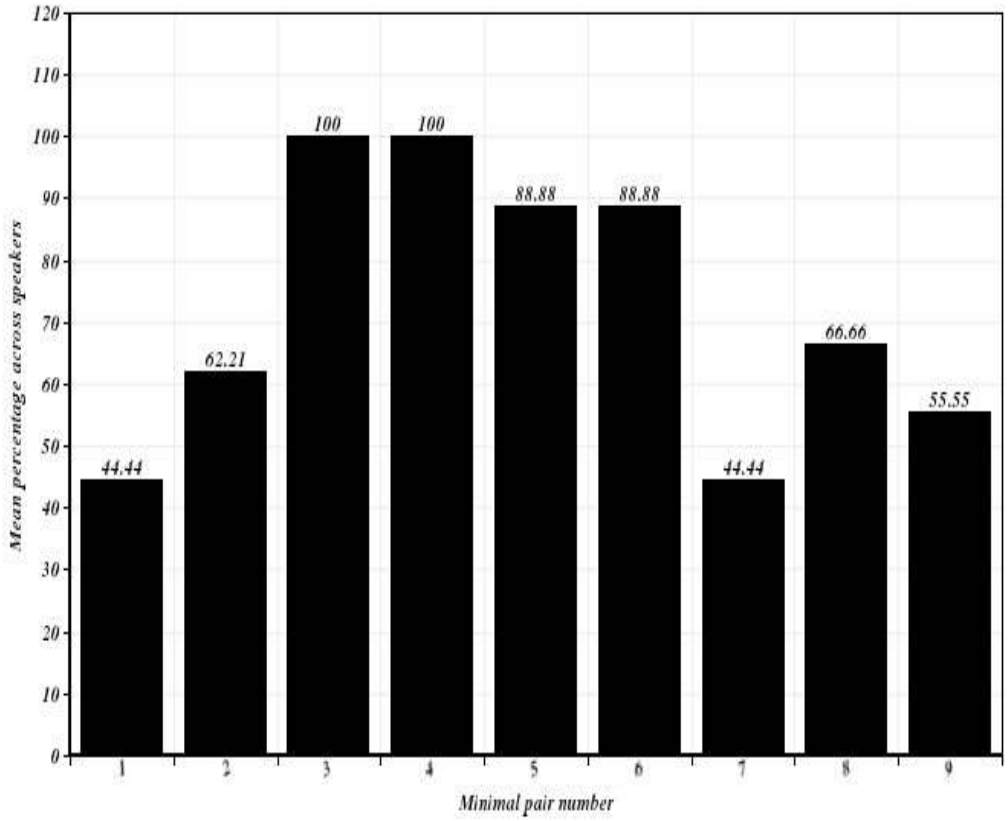
Figure M.4: Scores on minimal pairs across speakers (fricatives)



Minimal pair number (fricatives):

1-/f/-/θ/
2-/f/-/s/
3-/f/-/x/
4-/θ/-/x/
5-/s/-/x/
6-/s/-/tʃ/
7-/ɟ/-/tʃ/
8-/ɟ/-/ ɲ/

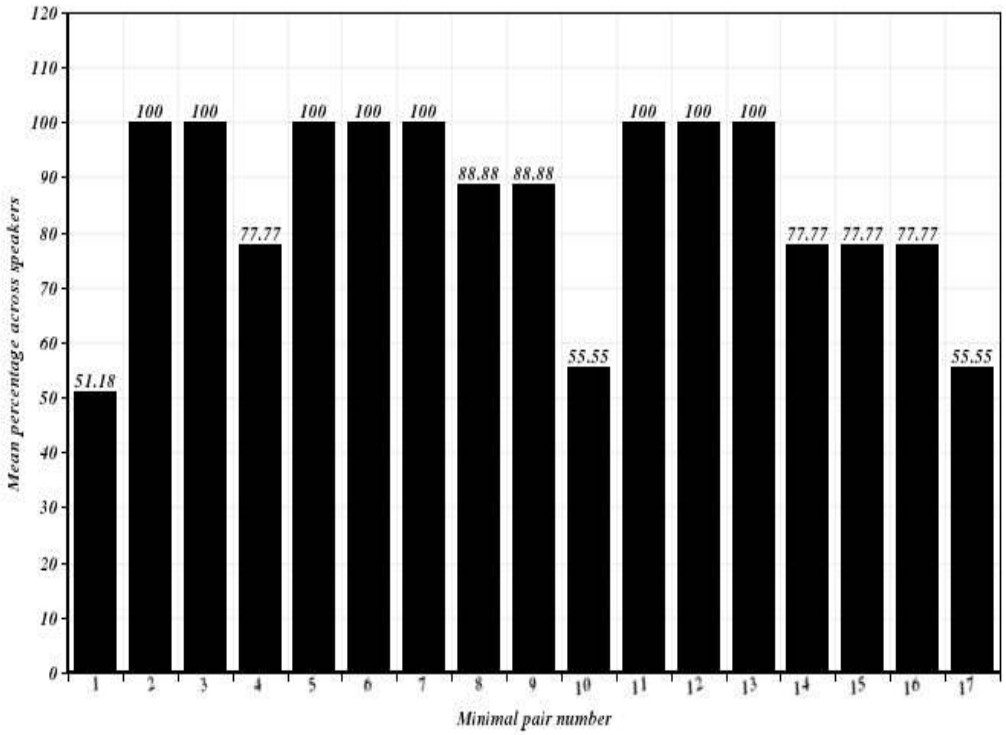
Figure M.5: Scores on minimal pairs across speakers (liquids)



Minimal pair number (liquids):

1-/ɻ/-/l/
2-/ɻ/-/r/
3-/ɻ/-/ɾ/
4-/ɻ/-/ɲ/
5-/l/-/ɾ/
6-/l/-/r/
7-/r/-/ɾ/
8-/r/-/d/
9-/r/-Ø

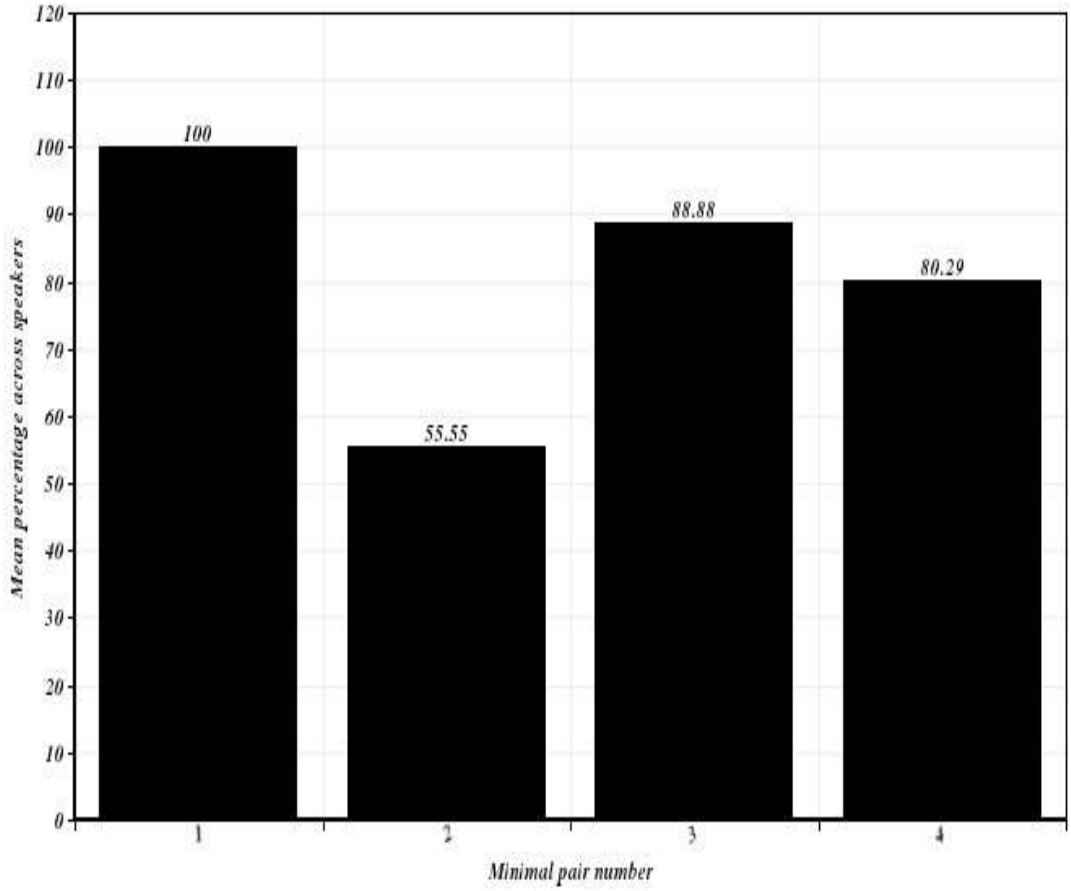
Figure M.6: Scores on minimal pairs across speakers (stops)



Minimal pair number (stops):

1-/p/-/b/
2-/p/-/f/
3-/p/-/m/
4-/p/-/t/
5-/p/-/k/
6-/b/-/m/
7-/b/-/f/
8-/b/-/d/
9-/b/-/g/
10-/d/-/t/
11-/d/-/θ/
12-/d/-/g/
13-/t/-/k/
14-/t/-/θ/
15-/k/-/g/
16-/k/-/x/
17-/g/-/x/

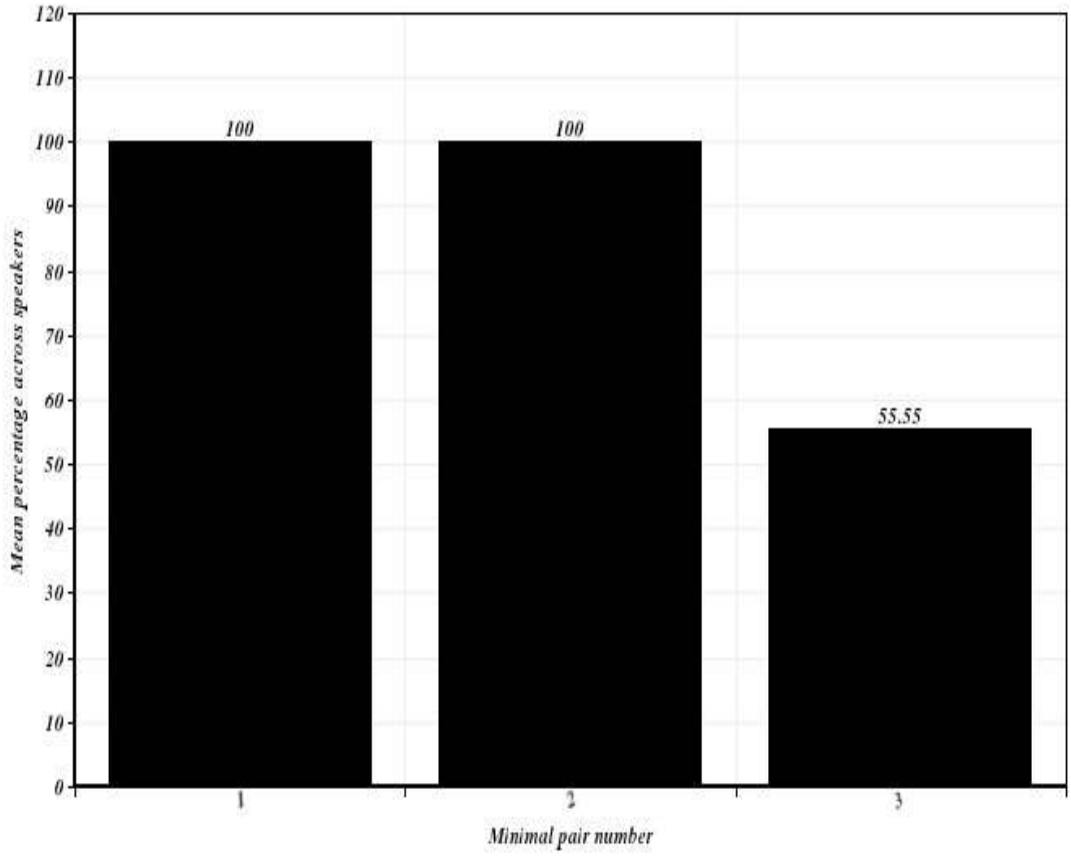
Figure M.7: Scores on minimal pairs across speakers (vowel combinations)



Minimal pair number (vowel combinations)

1-/ai/-/a/
2-/ei/-/e/
3-/au/-/a/
4-/eu/-/e/

Figure M.8: Scores on minimal pairs across speakers (nasals)



Minimal pair number (nasals):

1-/m/-/n/
2-/m/-/ɲ/
3-/n/-/ɲ/

APPENDIX N: SCORES ON CATEGORIES OF PHONEMIC ERRORS BASED ON PHONEMIC ERROR ANALYSIS (SENTENCE, PASSAGE, SEMI-SPONTANEOUS PRODUCTION TASKS)

Table N.1: Percentage of phonemic errors in each category per speaker (sentence level

Student	C1	C2	C3	C4	C5	C6	C7	C8
1	71.15	50.00	65.00	78.65	60.00	59.67	70.15	65.25
2	92.85	60.00	87.70	80.00	70.45	75.80	88.60	95.60
3	87.33	70.00	80.00	100.00	77.90	85.69	87.77	90.15
4	87.77	70.00	75.00	95.35	84.01	94.50	85.77	80.75
5	100.00	90.00	100.00	100.00	80.00	100.00	92.88	94.59
6	88.88	100.00	100.00	95.35	84.01	94.50	87.33	100.00
7	87.77	80.00	65.00	92.35	64.20	88.00	70.00	78.25
8	87.77	75.00	80.00	87.90	80.49	81.69	82.22	85.00
9	82.22	65.00	80.00	79.01	70.00	75.35	84.50	69.25
10	90.88	87.00	90.00	95.35	92.95	94.50	94.85	88.77
11	88.25	85.00	90.00	79.01	87.90	83.69	86.55	100.00
12	87.77	84.00	90.00	88.00	85.00	94.50	100.00	98.59
13	87.75	85.00	80.00	100.00	89.01	85.00	98.88	92.59
14	90.00	85.00	100.00	100.00	89.01	91.89	85.55	84.77
15	90.00	85.00	80.00	92.35	94.35	93.50	79.85	87.70
16	87.77	90.00	100.00	85.01	77.90	95.50	84.55	100.00
17	100.00	90.00	100.00	85.00	88.00	100.00	92.88	81.75
18	88.88	80.00	90.00	100.00	87.90	100.00	75.45	92.59
19	100.00	90.00	90.00	100.00	89.01	92.50	90.00	94.50
20	100.00	80.00	90.00	100.00	67.90	80.00	83.33	92.59
NS1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
NS2	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table N.2: Percentage of phonemic errors in each category per speaker (passage level)

Student	C1	C2	C3	C4	C5	C6	C7	C8
1	75.15	55.00	70.00	83.65	65.00	64.67	75.15	70.25
2	94.85	65.00	92.70	85.00	75.45	79.80	92.60	94.60
3	92.00	79.00	88.00	100.00	82.90	89.00	92.77	93.15
4	92.77	75.00	80.00	94.00	89.01	95.50	89.77	84.75
5	100.00	94.00	100.00	100.00	84.00	100.00	96.88	95.00
6	92.88	100.00	100.00	95.35	88.01	94.50	92.33	100.00
7	89.77	84.00	69.00	92.00	68.20	90.00	79.00	82.25
8	89.77	77.00	84.00	92.90	94.49	88.69	90.22	92.00
9	88.22	69.00	82.00	82.01	75.00	79.35	89.50	72.25
10	88.88	85.00	92.00	94.00	90.00	92.50	92.85	90.77
11	92.25	88.00	94.00	82.01	88.90	86.69	88.55	100.00
12	89.77	84.00	90.00	88.00	85.00	94.50	100.00	98.59
13	87.75	89.00	84.00	100.00	92.01	92.00	96.00	94.50
14	94.00	88.00	100.00	100.00	86.01	94.85	86.50	86.75
15	94.00	90.00	85.00	92.00	96.35	94.50	82.85	90.70
16	92.77	96.00	100.00	88.01	82.90	96.50	88.55	100.00
17	100.00	90.40	100.00	85.00	88.00	100.00	92.88	84.75
18	92.80	86.00	96.00	100.00	88.90	100.00	80.45	92.59
19	100.00	94.00	95.00	100.00	89.01	92.50	90.00	94.50
20	100.00	80.00	90.00	100.00	67.90	80.00	83.33	92.59
NS1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
NS2	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table N.3: Percentage of phonemic errors in each category per speaker (semi-spontaneous production)

Student	C1	C2	C3	C4	C5	C6	C7	C8
1	70.15	50.00	68.00	80.65	61.00	62.67	70.15	72.25
2	88.85	56.00	82.70	82.00	71.45	76.80	87.60	92.60
3	84.00	76.50	82.00	98.00	80.90	83.00	90.77	86.15
4	90.77	73.00	73.00	90.00	86.01	91.50	85.77	82.75
5	96.00	92.00	100.00	95.50	82.00	97.00	93.88	91.00
6	90.88	94.00	100.00	91.25	80.00	92.50	90.33	94.00
7	85.75	82.00	65.00	90.00	62.20	88.00	76.00	80.25
8	80.75	74.00	80.00	86.00	90.50	83.69	85.22	86.00
9	85.25	65.00	80.00	80.00	73.00	77.35	87.50	70.25
10	85.88	83.00	90.00	92.00	90.00	88.50	88.85	85.77
11	88.25	85.00	90.00	80.00	84.90	83.69	84.50	90.00
12	83.75	85.00	88.00	86.00	84.00	90.50	100.00	94.59
13	85.75	84.00	83.00	100.00	90.00	90.00	94.00	92.50
14	92.00	85.00	98.00	98.00	83.00	90.85	84.00	84.75
15	92.00	90.00	83.00	90.00	93.35	92.50	80.85	88.70
16	90.77	94.00	96.00	83.00	80.00	92.50	84.55	100.00
17	94.00	86.50	94.00	80.00	82.00	93.00	86.88	80.75
18	88.80	84.00	92.00	96.00	86.00	95.00	80.50	88.55
19	100.00	92.00	93.00	100.00	86.00	90.50	88.00	92.50
20	100.00	80.00	90.00	100.00	67.90	80.00	83.33	92.59
NS1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
NS2	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

APPENDIX O: SPEECH RATE, PAUSE FREQUENCY AND PAUSE DURATION IN CONNECTED SPEECH

Table O.1: Speech rate, pause frequency and pause duration: mean values per speaker (sentence level)

Student	Speech rate	Pause frequency	Pause duration
1	2.60	3.48	0.49
2	2.70	3.00	0.50
3	2.90	2.60	0.52
4	3.25	1.15	0.25
5	3.85	0.68	0.10
6	3.50	1.10	0.20
7	3.10	1.75	0.35
8	3.00	2.00	0.45
9	2.65	3.50	0.46
10	3.88	0.62	0.15
11	3.20	1.20	0.30
12	4.00	0.60	0.14
13	2.85	2.65	0.50
14	3.02	2.50	0.50
15	3.20	1.50	0.30
16	3.00	1.90	0.40
17	2.80	2.66	0.48
18	3.80	1.00	0.18
19	3.25	1.15	0.25
20	3.83	0.85	0.16
NS1	4.18	0.52	0.09
NS2	4.22	0.48	0.07

Table O.2: Speech rate, pause frequency and pause duration: mean values per speaker (passage task)

Student	Speech rate	Pause frequency	Pause duration
1	2.56	3.55	0.50
2	2.66	3.15	0.60
3	2.80	2.70	0.58
4	3.15	1.20	0.35
5	3.75	0.75	0.20
6	3.44	1.20	0.35
7	3.00	1.80	0.40
8	2.95	2.15	0.50
9	2.60	3.75	0.48
10	3.75	0.68	0.20
11	3.15	1.25	0.38
12	4.00	0.70	0.19
13	2.65	2.80	0.60
14	2.90	2.60	0.54
15	3.15	1.75	0.38
16	2.90	1.95	0.46
17	2.70	2.66	0.52
18	3.65	1.15	0.22
19	3.15	1.25	0.32
20	3.80	0.90	0.44
NS1	4.20	0.50	0.08
NS2	4.25	0.45	0.06

Table O.3: Speech rate, pause frequency and pause duration: mean value per speaker (semi-spontaneous task)

Student	Speech rate	Pause frequency	Pause duration
1	2.50	3.59	0.54
2	2.59	3.12	0.59
3	2.80	2.68	0.60
4	3.10	1.25	0.32
5	3.74	0.75	0.19
6	3.36	1.19	0.28
7	3.00	1.83	0.42
8	2.98	2.12	0.54
9	2.60	3.65	0.60
10	3.76	0.58	0.25
11	3.12	1.32	0.39
12	3.94	0.74	0.20
13	2.70	2.85	0.60
14	2.90	2.68	0.62
15	3.15	1.70	0.40
16	2.89	1.99	0.49
17	2.71	2.80	0.52
18	3.74	1.12	0.25
19	3.12	1.20	0.34
20	3.76	0.90	0.28
NS1	4.25	0.45	0.08
NS2	4.30	0.38	0.09

APPENDIX P: SPEECH RATE, PAUSE FREQUENCY AND PAUSE DURATION: L1, GENDER AND NATIVE VERSUS NON NATIVE DIFFERENCES

Figure P.1: Speech rate (native versus non-native speakers)

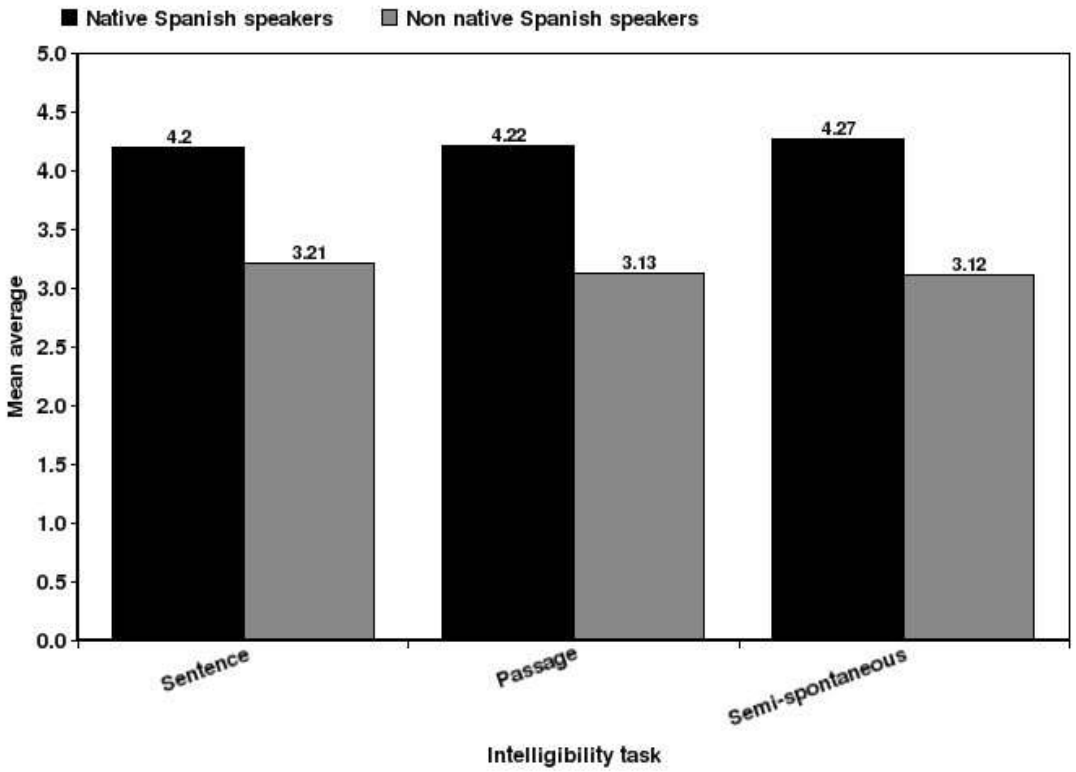


Figure P.2: Pause frequency (native versus non-native speakers)

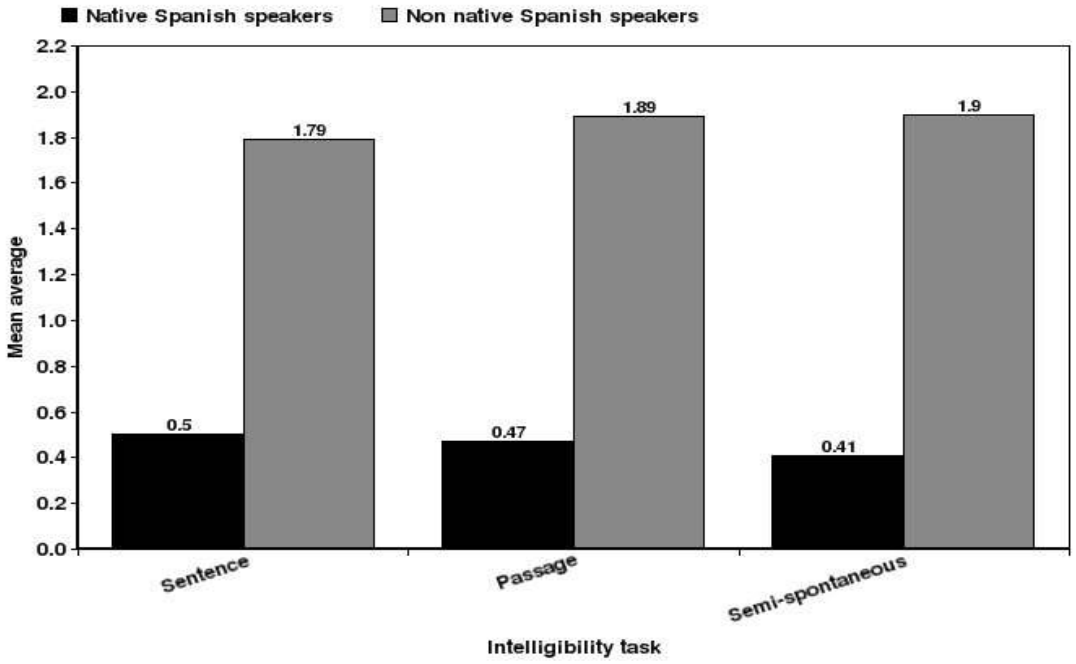


Figure P.3: Pause duration (native versus non-native speakers)

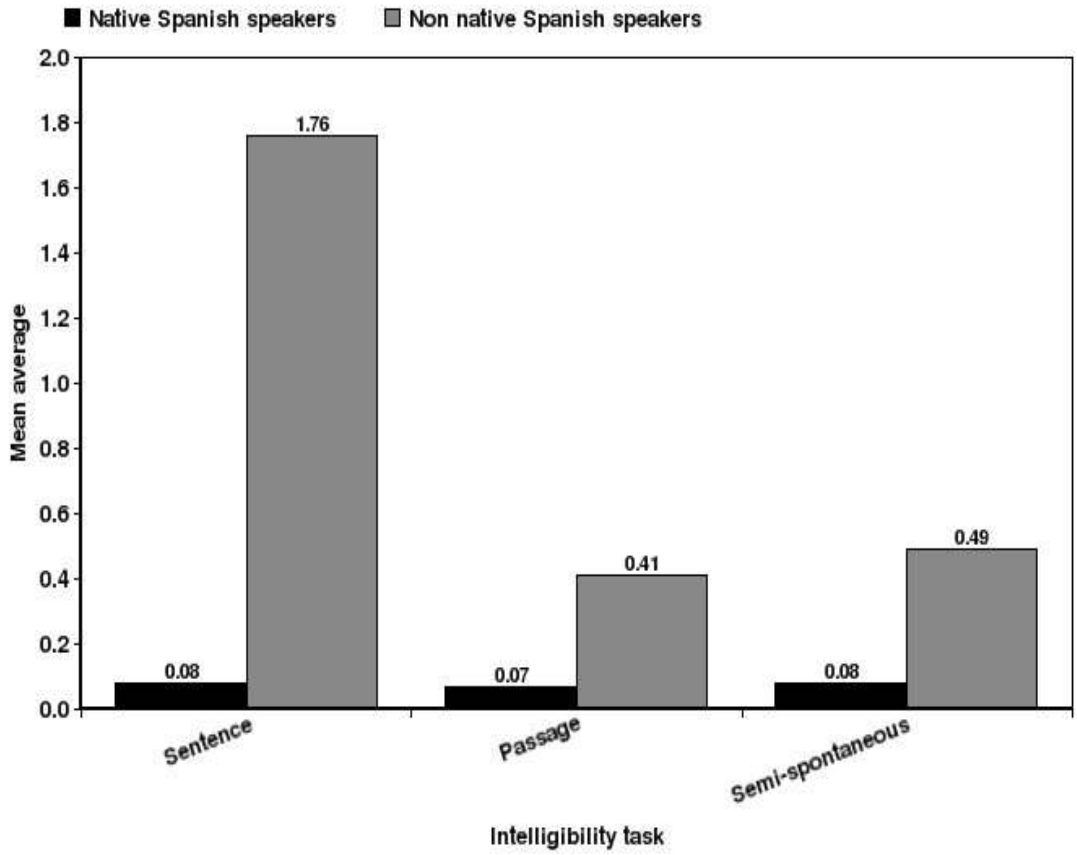


Figure P.4: Speech rate (gender)

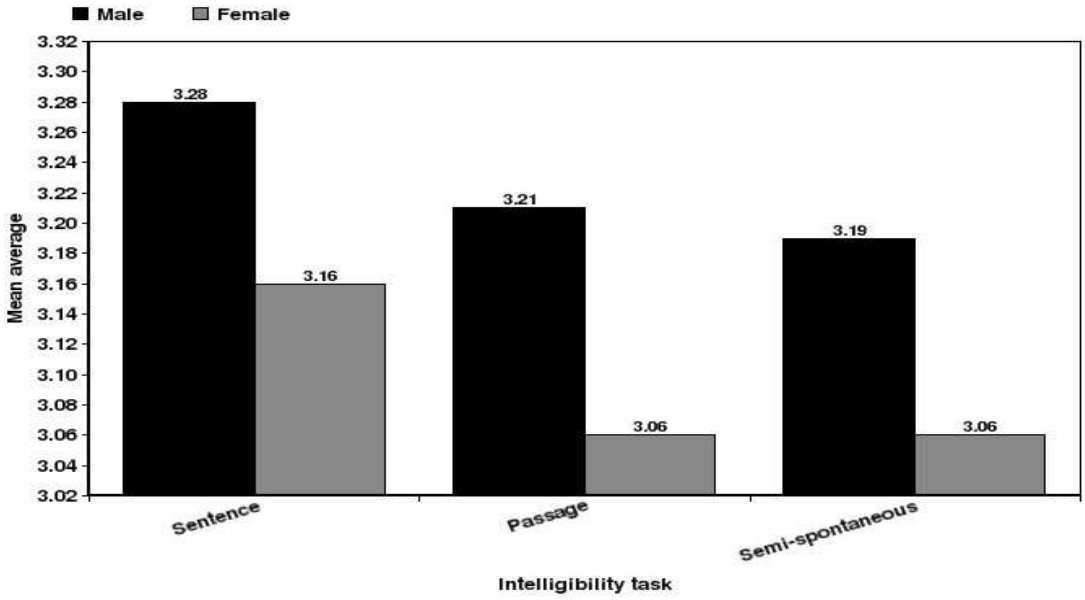


Figure P.5: Pause frequency (gender)

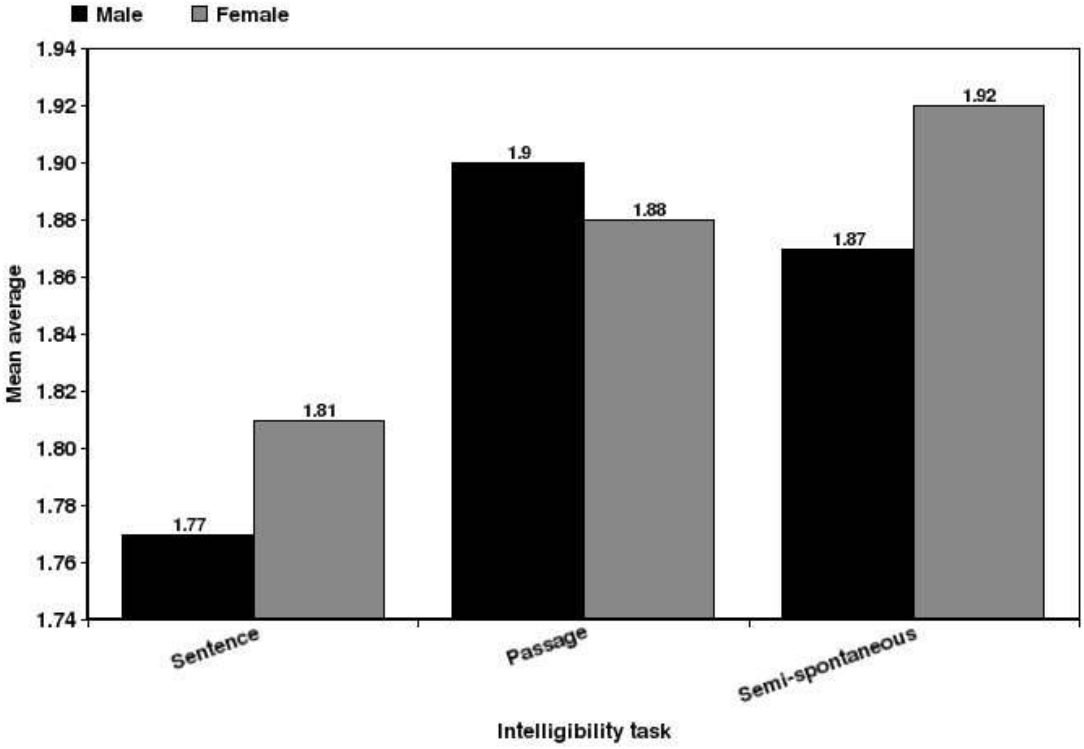


Figure P.6: Pause duration (gender)

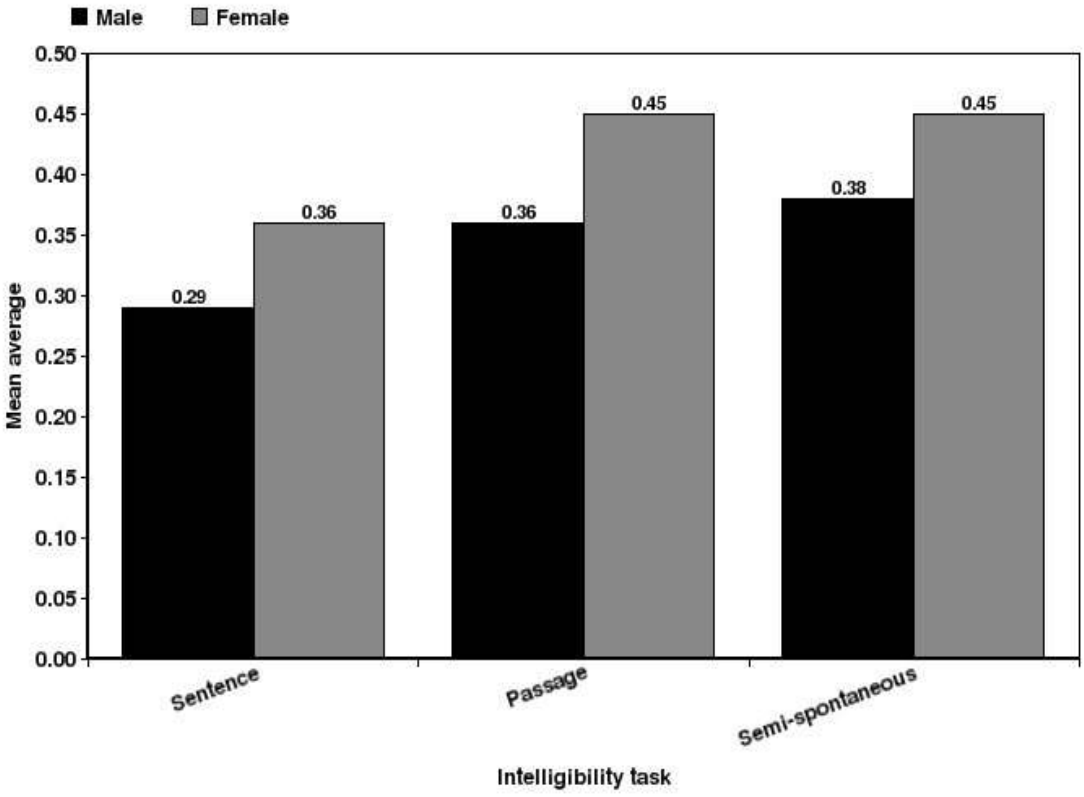


Figure P.7: Speech rate (monolingual versus bilingual)

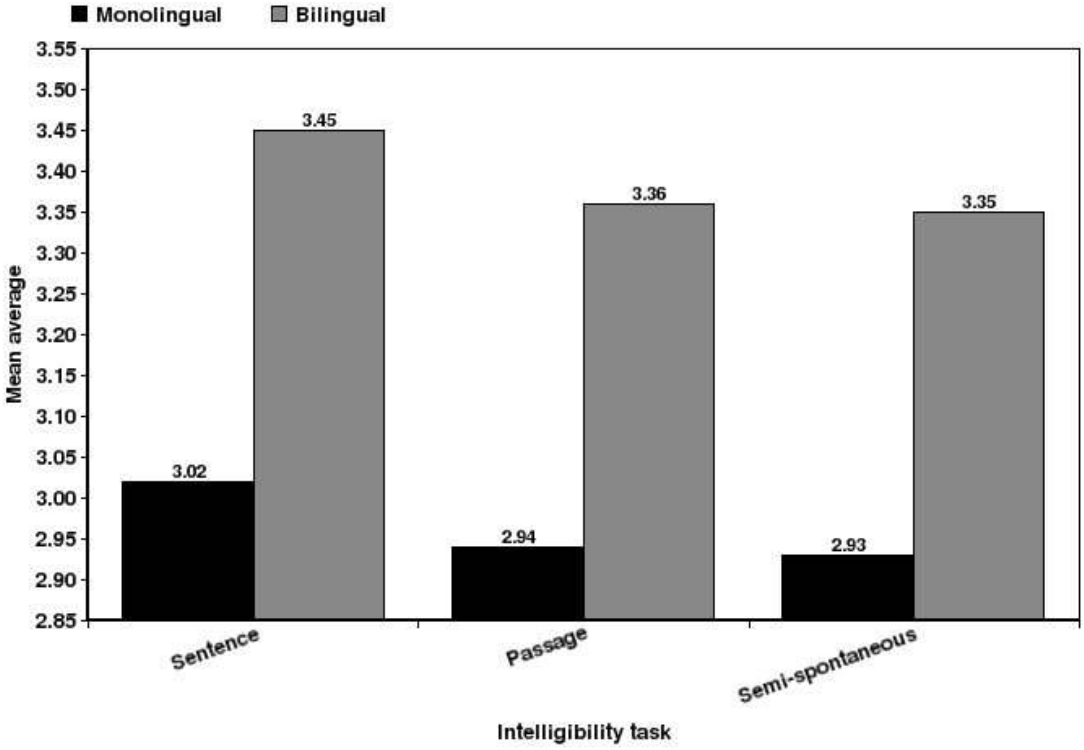


Figure P. 8: Pause frequency (monolingual versus bilingual)

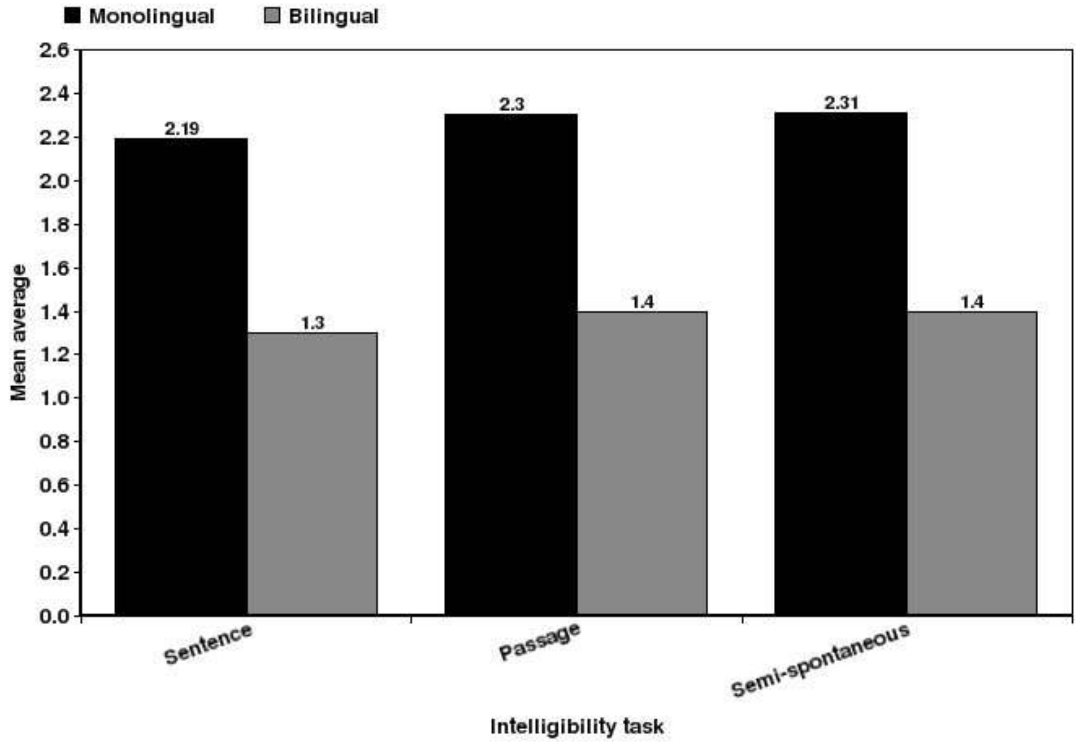
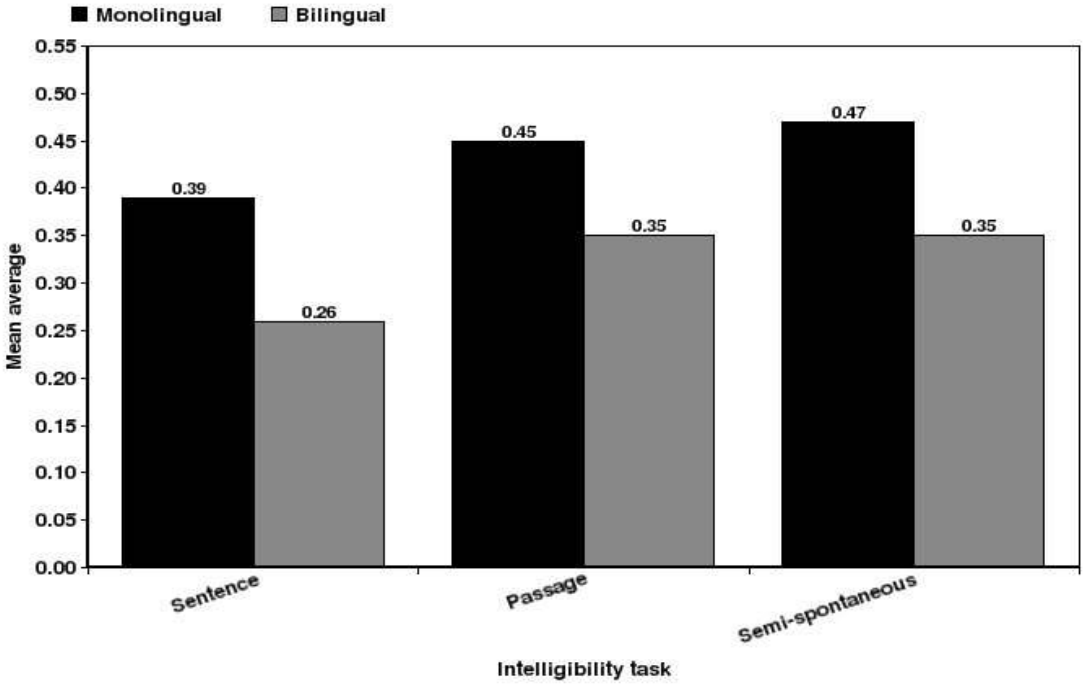


Figure P. 9: Pause duration (monolingual versus bilingual)



APPENDIX Q: CORRELATION ANALYSES (SCATTER PLOTS)

Figure Q.1: Scatter plot of intelligibility scores at word and sentence level

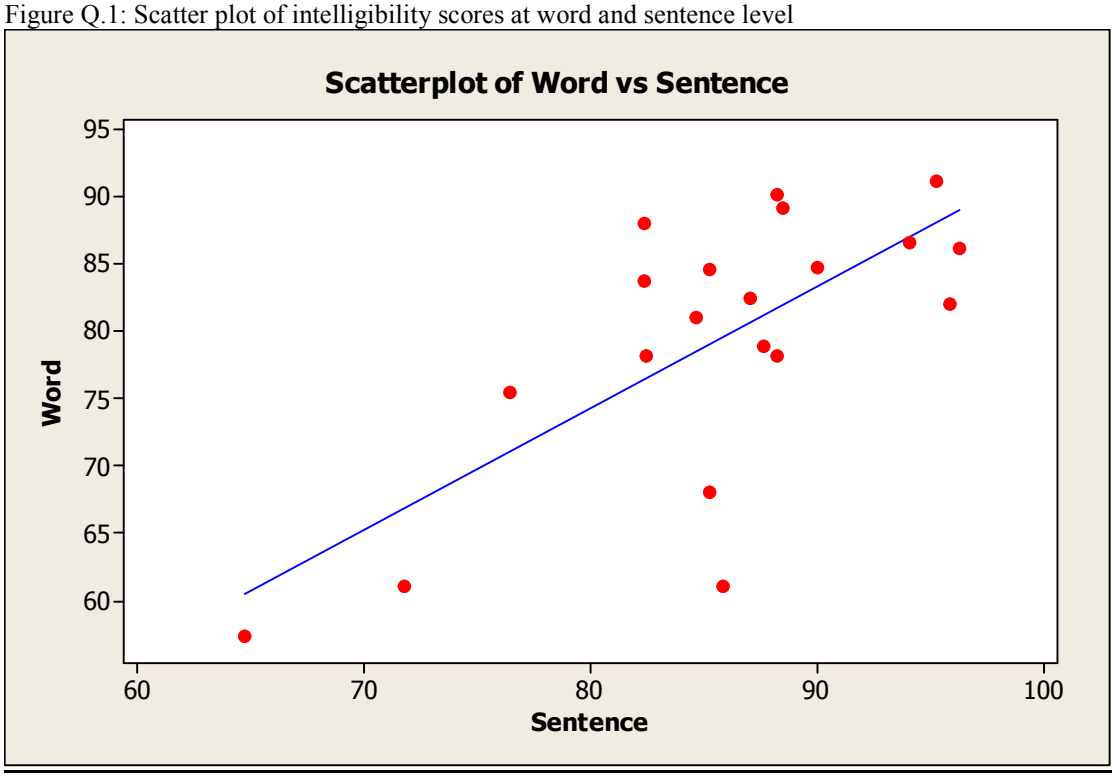


Figure Q.2: Scatter plot of intelligibility scores at word and passage level

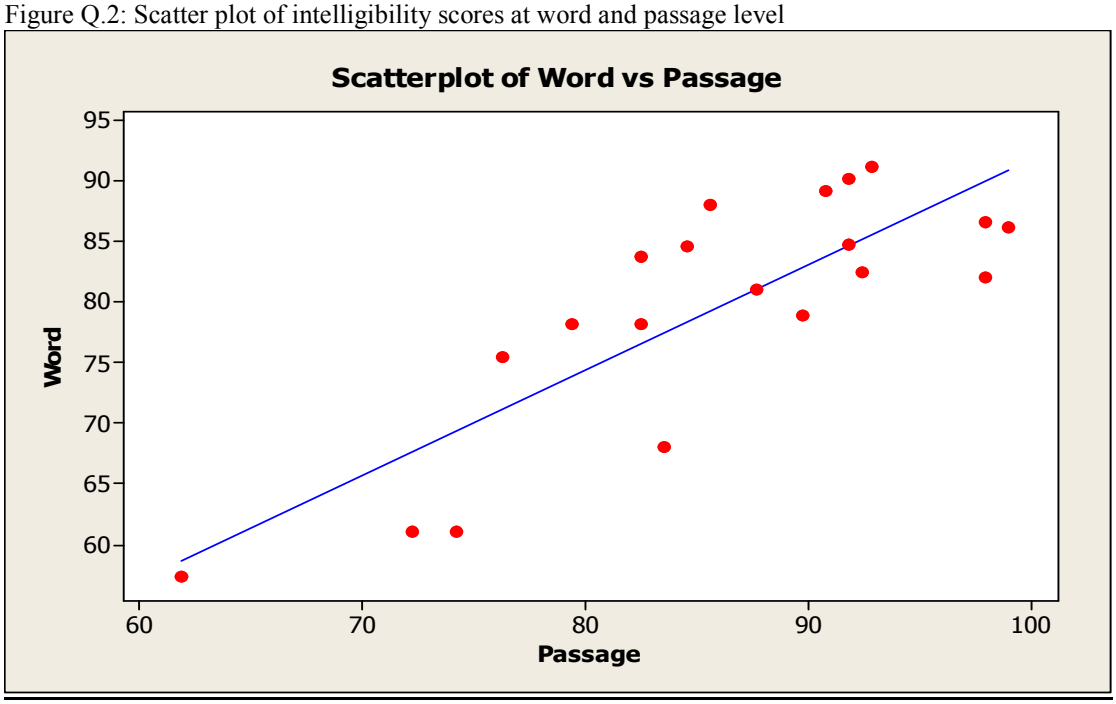


Figure Q.3: Scatter plot of intelligibility scores at word level and in the semi-spontaneous task

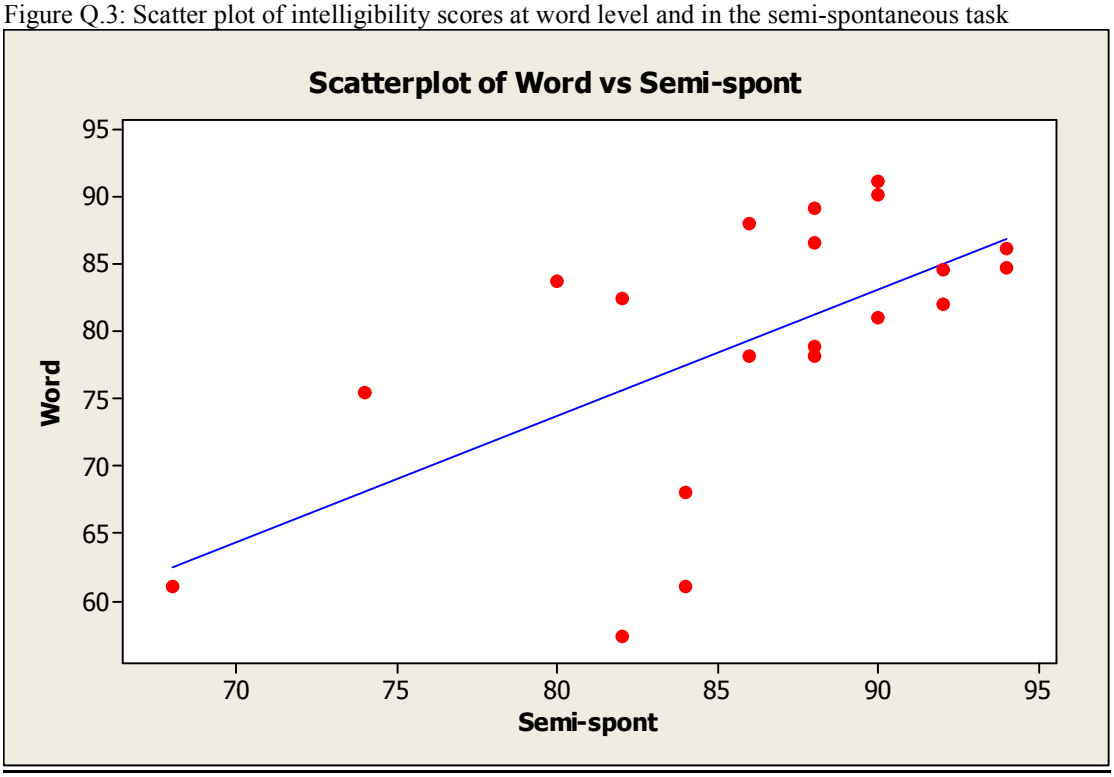


Figure Q.4: Scatter plot of intelligibility scores at sentence and passage levels

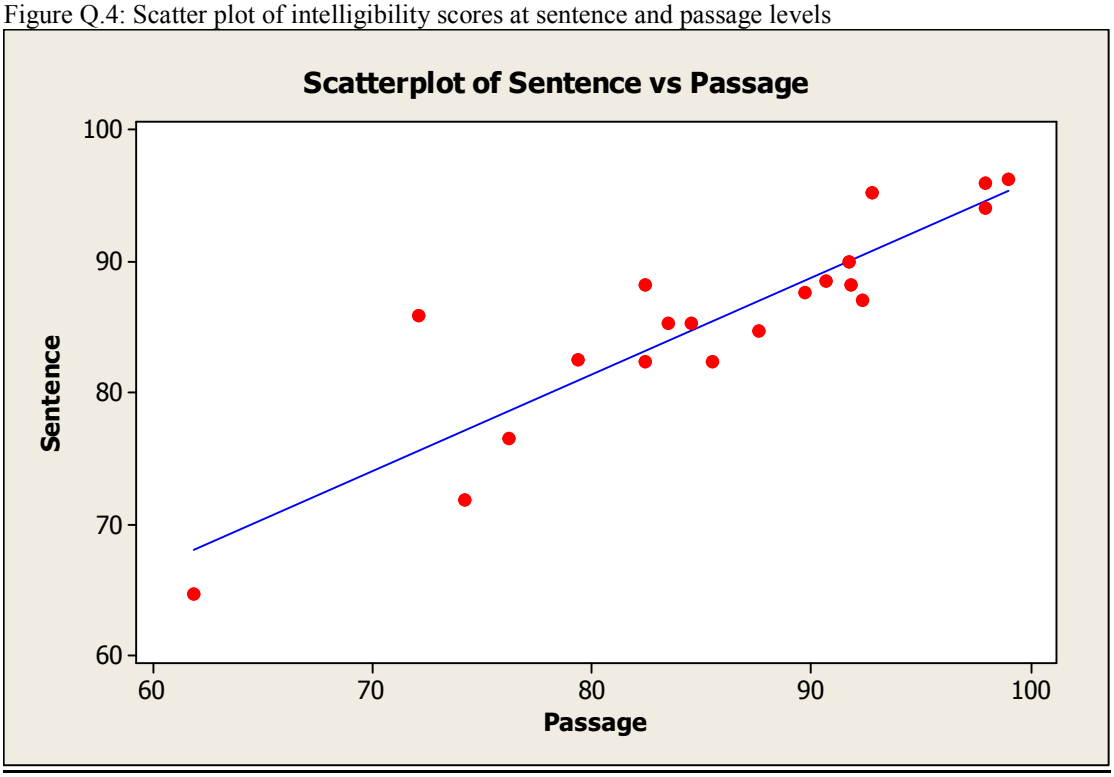


Figure Q.5: Scatter plot of intelligibility scores at sentence level and in the semi-spontaneous task

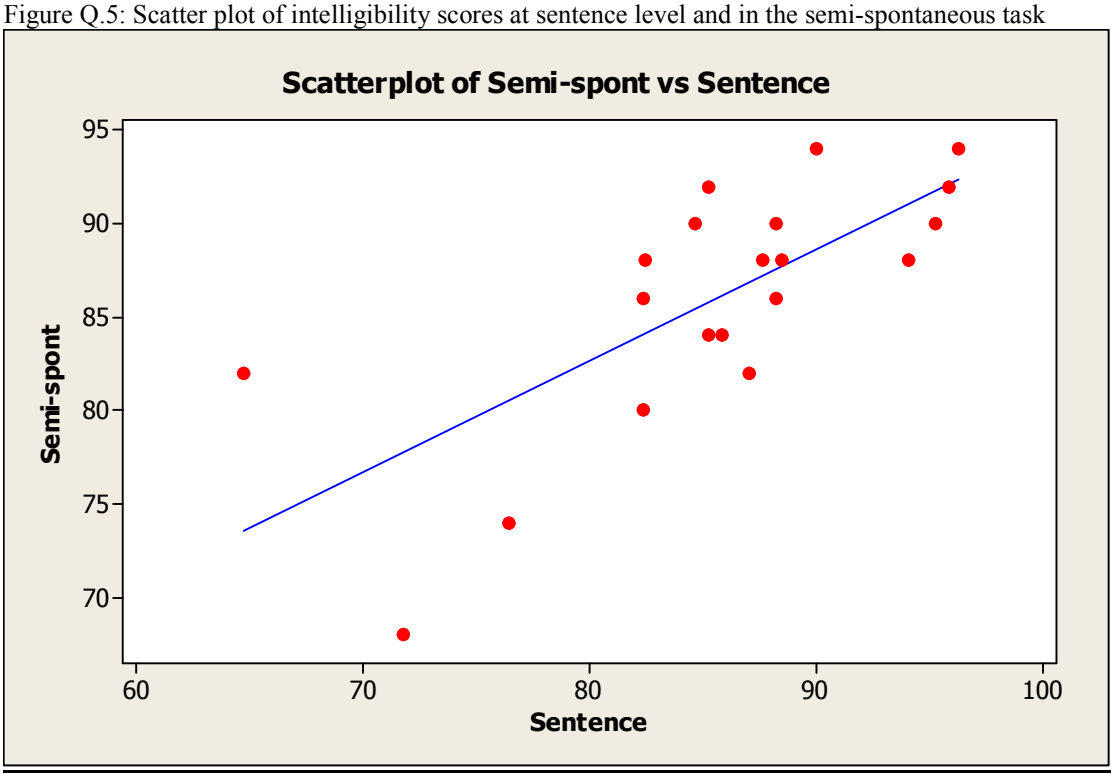


Figure Q.6: Scatter plot of intelligibility scores at passage level and in the semi-spontaneous task

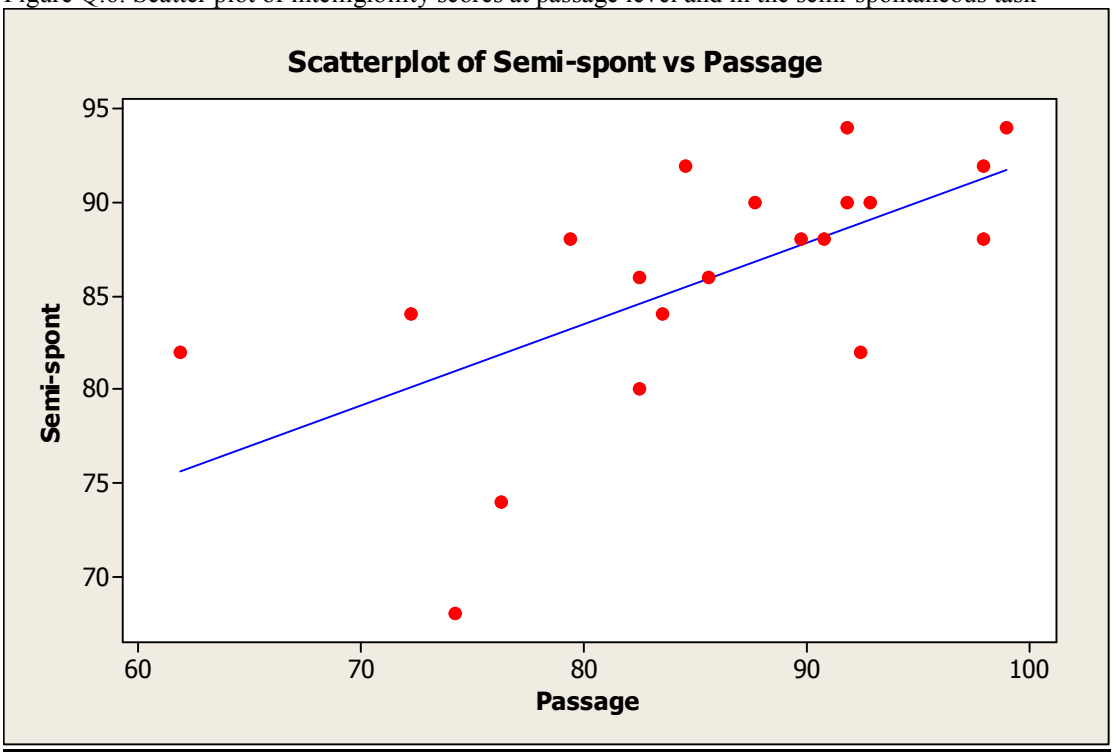


Figure Q.7: Scatter plot of intelligibility scores at sentence level and error category 1 (affricates)

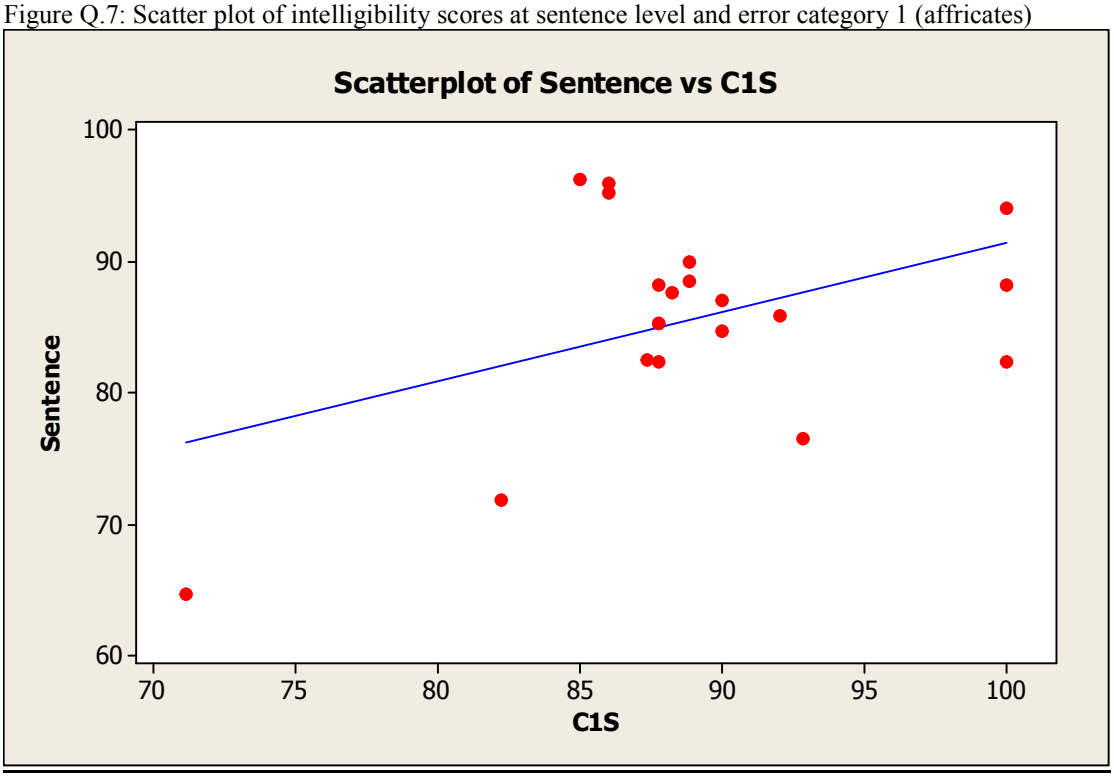


Figure Q.8: Scatter plot of intelligibility scores at sentence level and error category 2 (unstressed vowels)

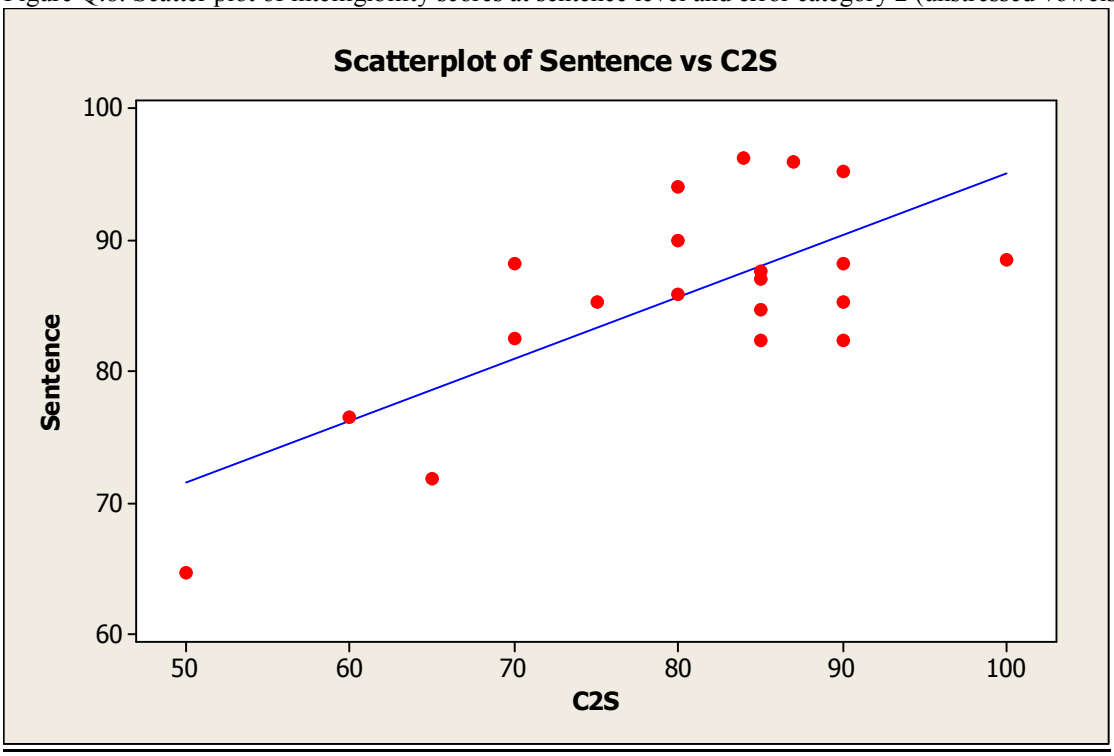


Figure Q. 9: Scatter plot of intelligibility scores at sentence level and error category 3 (stressed vowels)

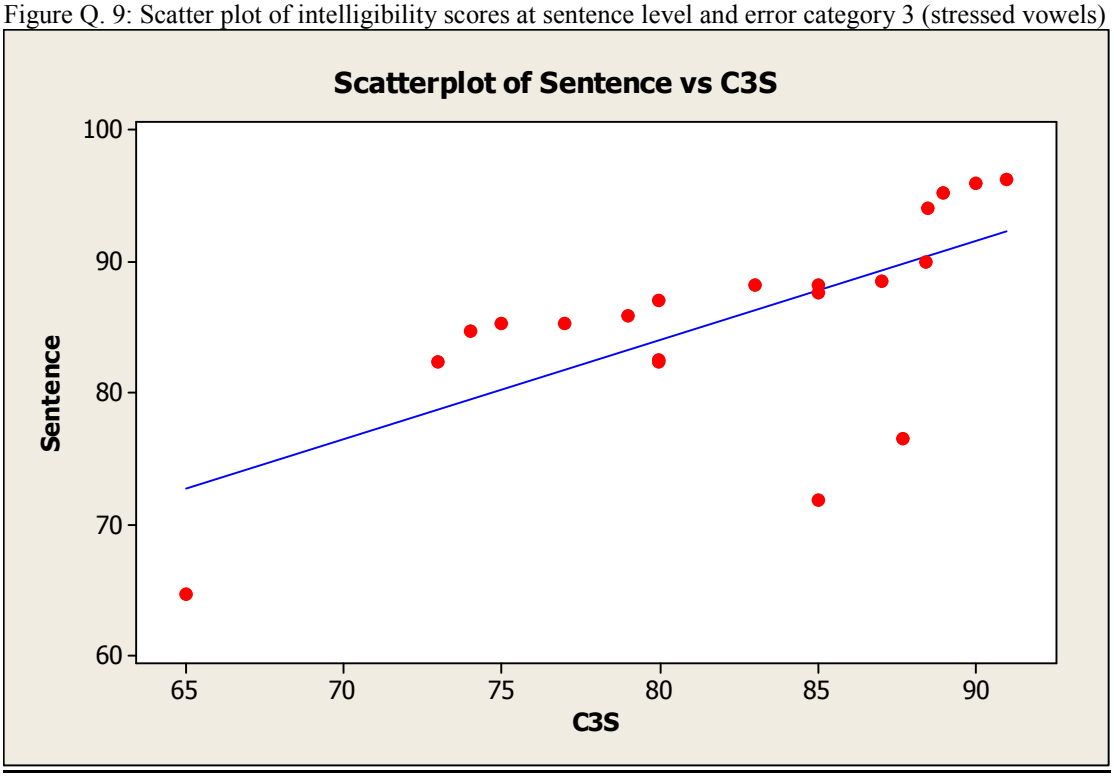


Figure Q.10: Scatter plot of intelligibility scores at sentence level and error category 4 (fricatives)

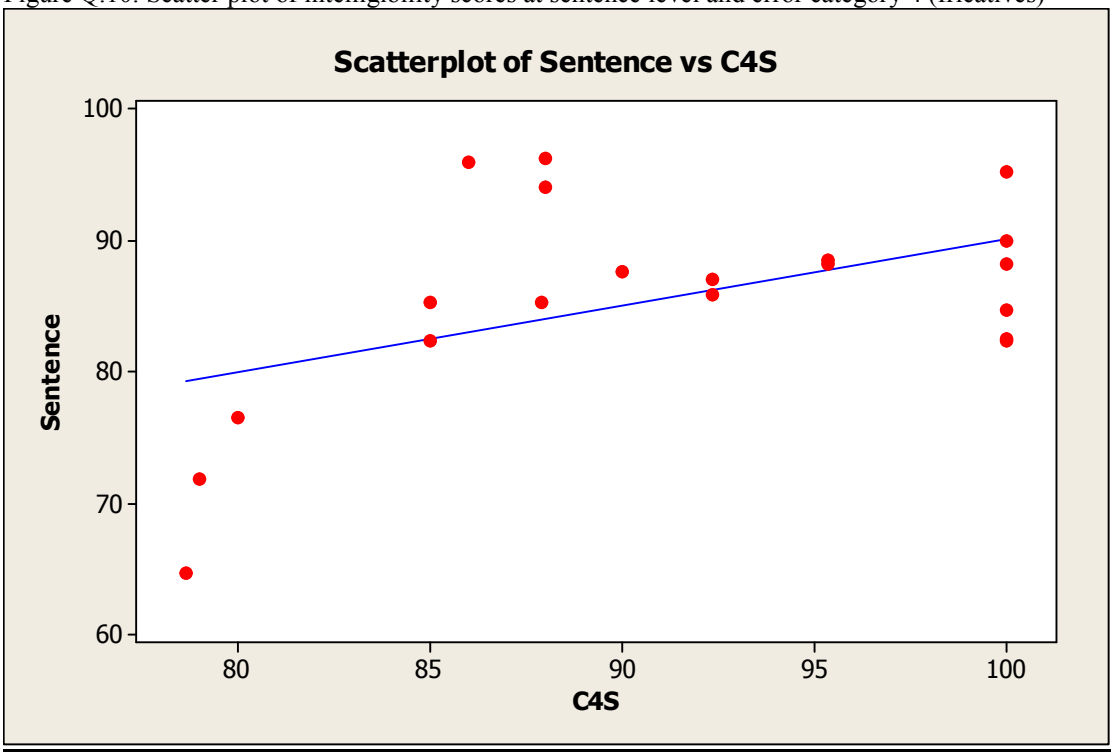


Figure Q.11: Scatter plot of intelligibility scores at sentence level and error category 5 (liquids)

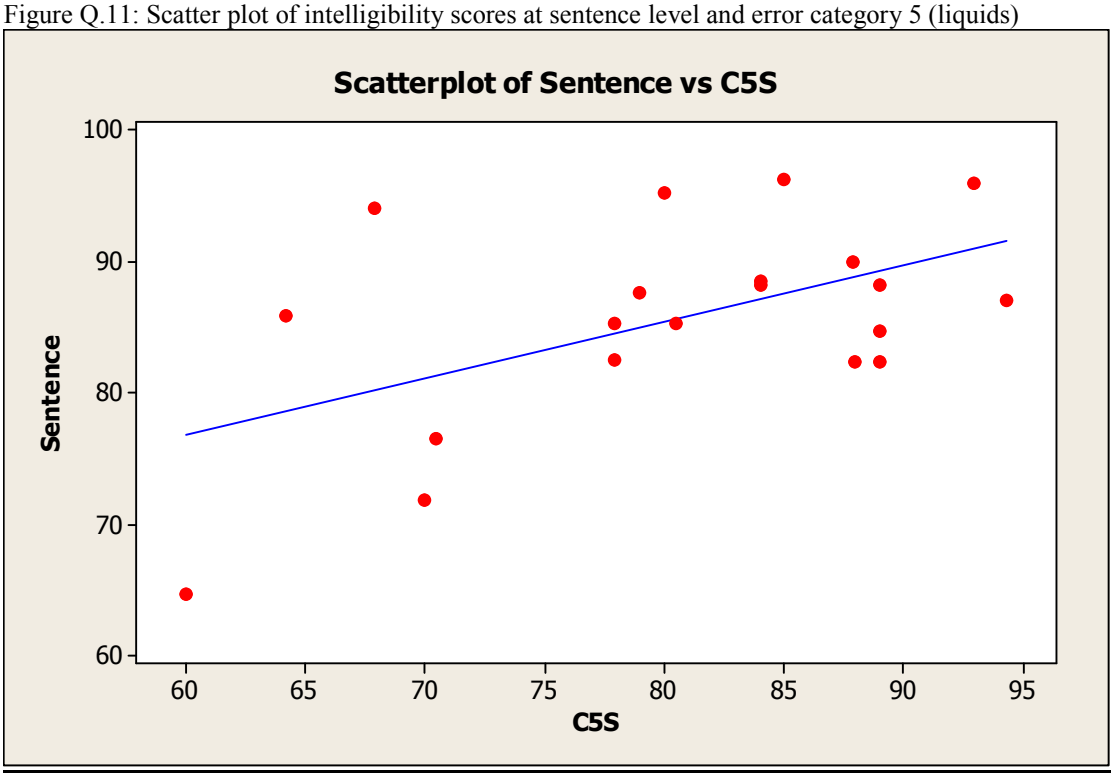


Figure Q.12: Scatter plot of intelligibility scores at sentence level and error category 6 (stops)

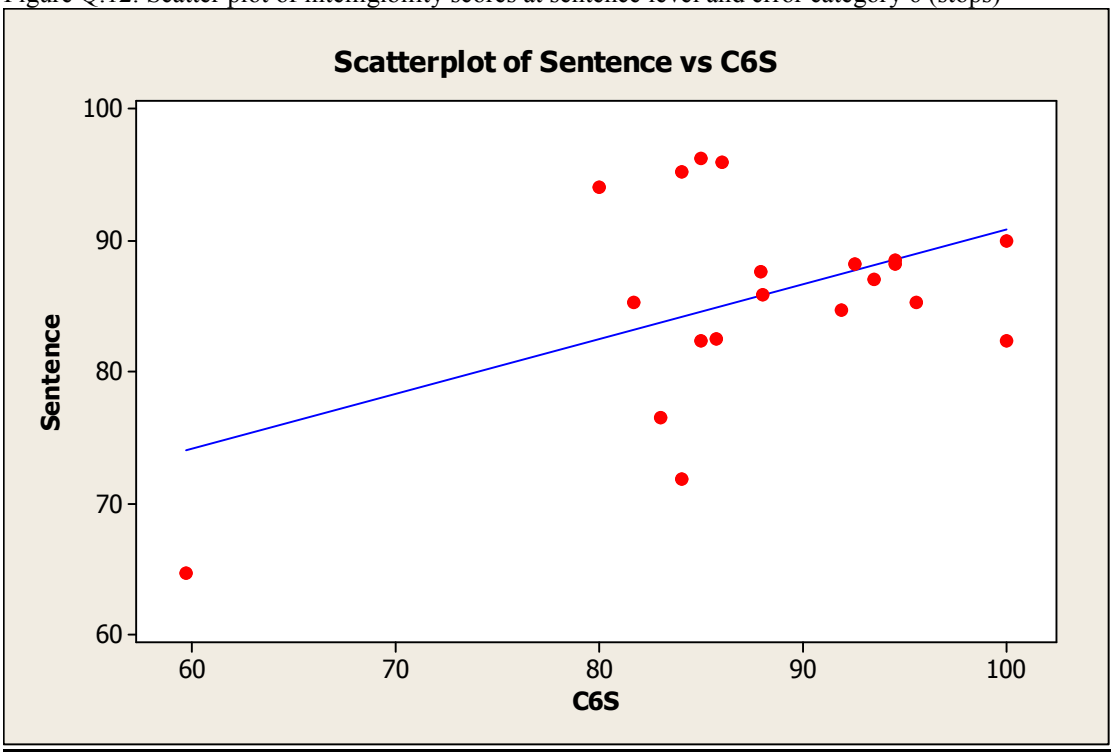


Figure Q.13: Scatter plot of intelligibility scores at sentence level and error category 7 (vowel combinations)

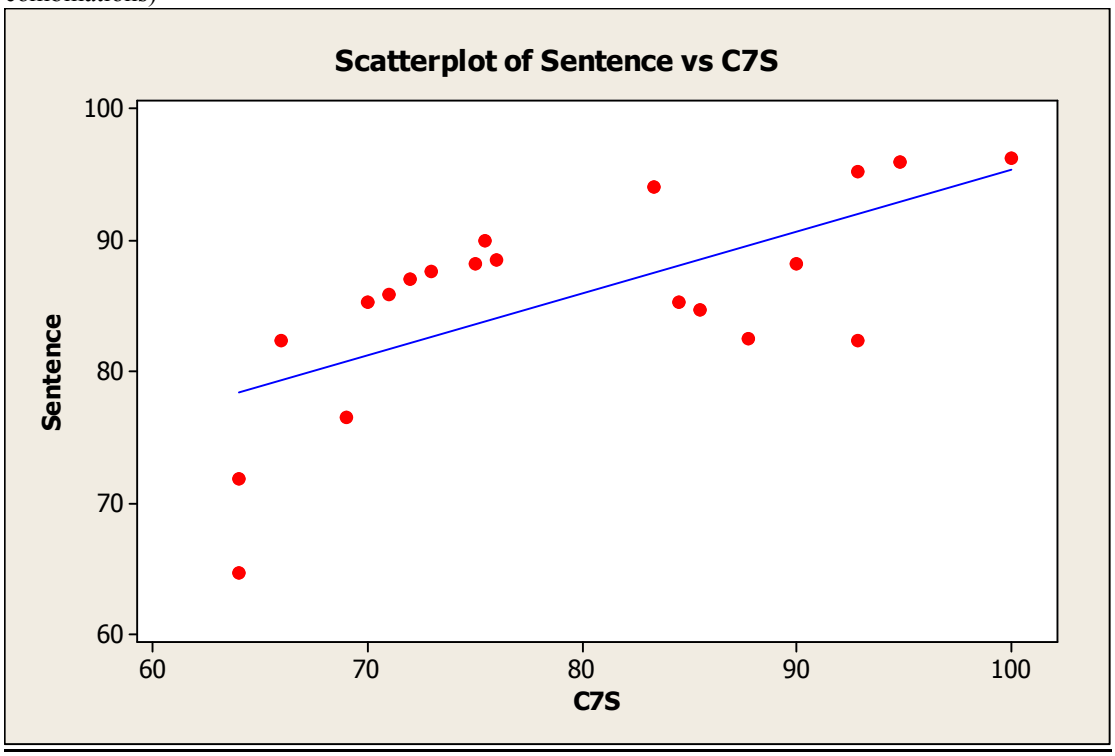


Figure Q.14: Scatter plot of intelligibility scores at sentence level and error category 8 (nasals)

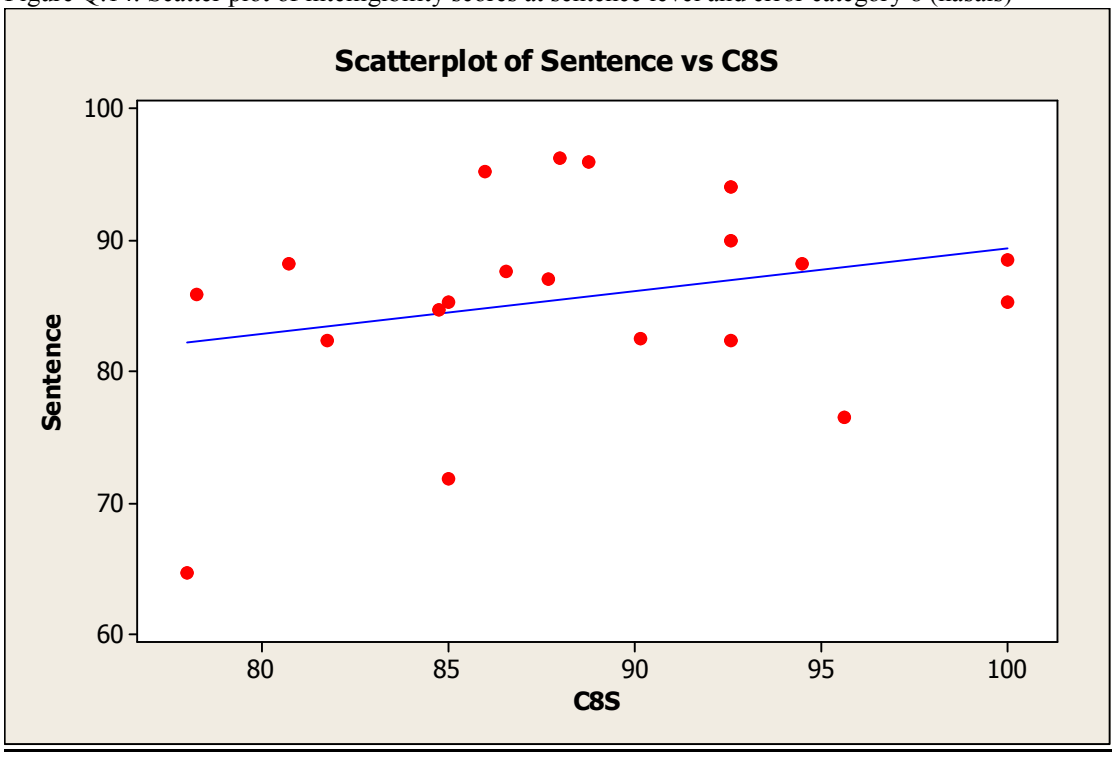


Figure Q.15: Scatter plot of intelligibility scores at passage level and error category 1 (affricates)

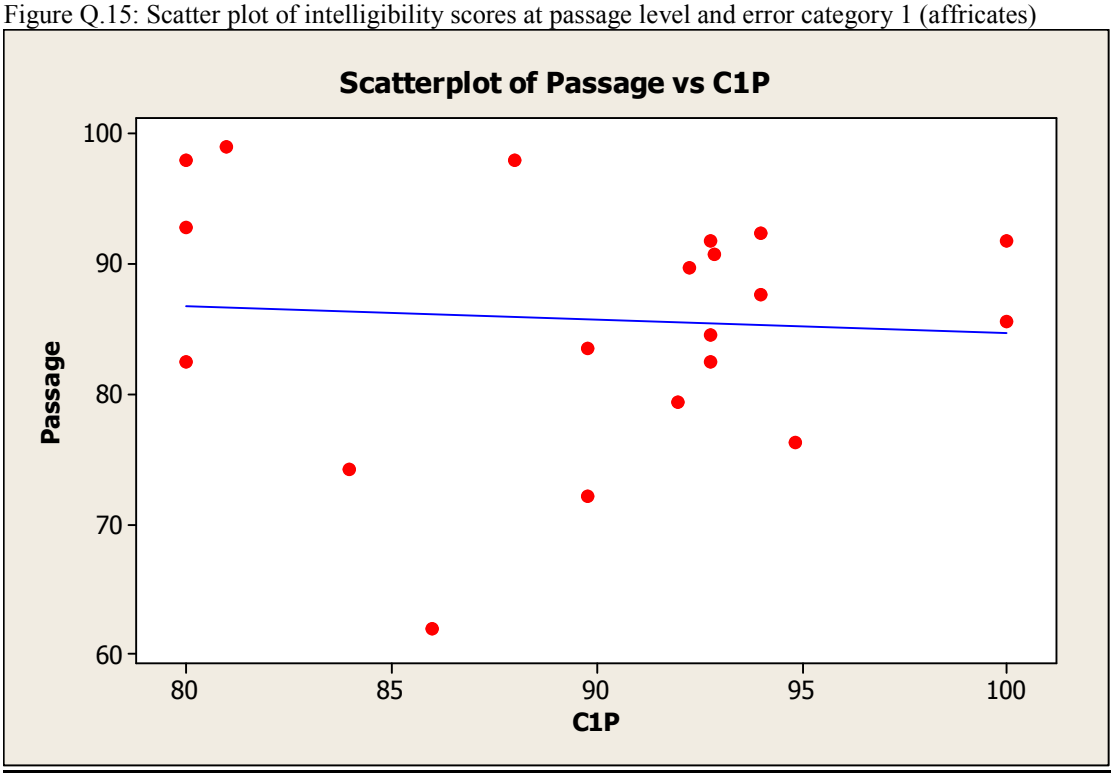


Figure Q.16: Scatter plot of intelligibility scores at passage level and error category 2 (unstressed vowels)

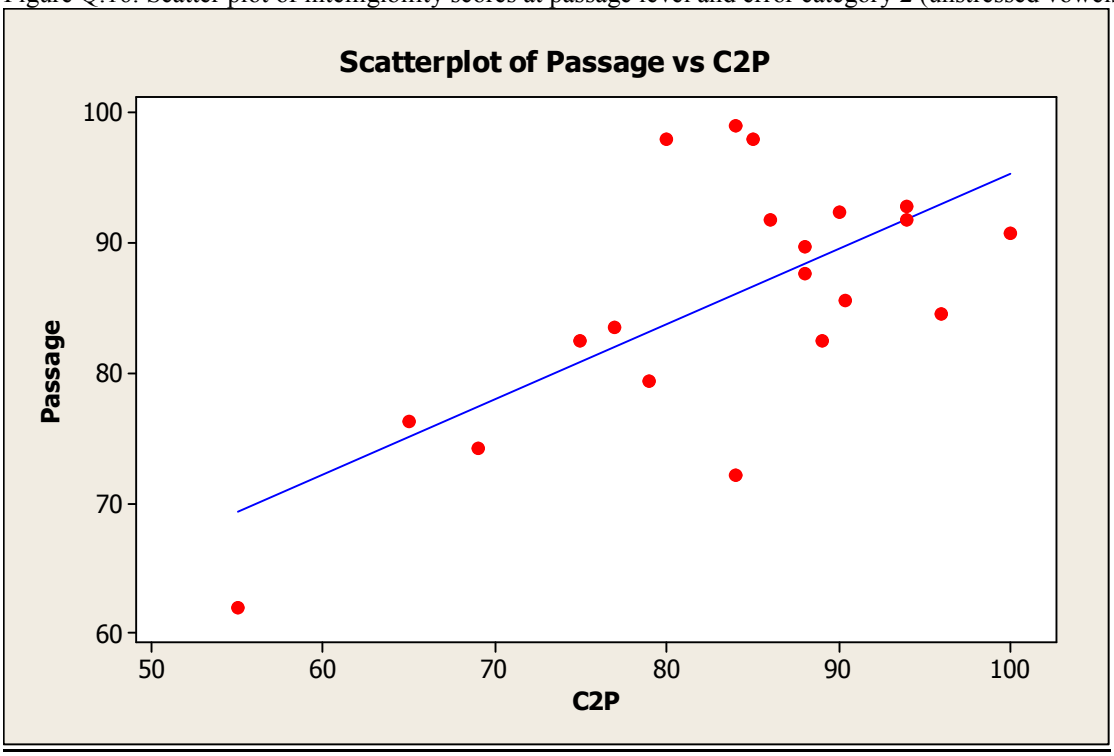


Figure Q.17: Scatter plot of intelligibility scores at passage level and error category 3 (stressed vowels)

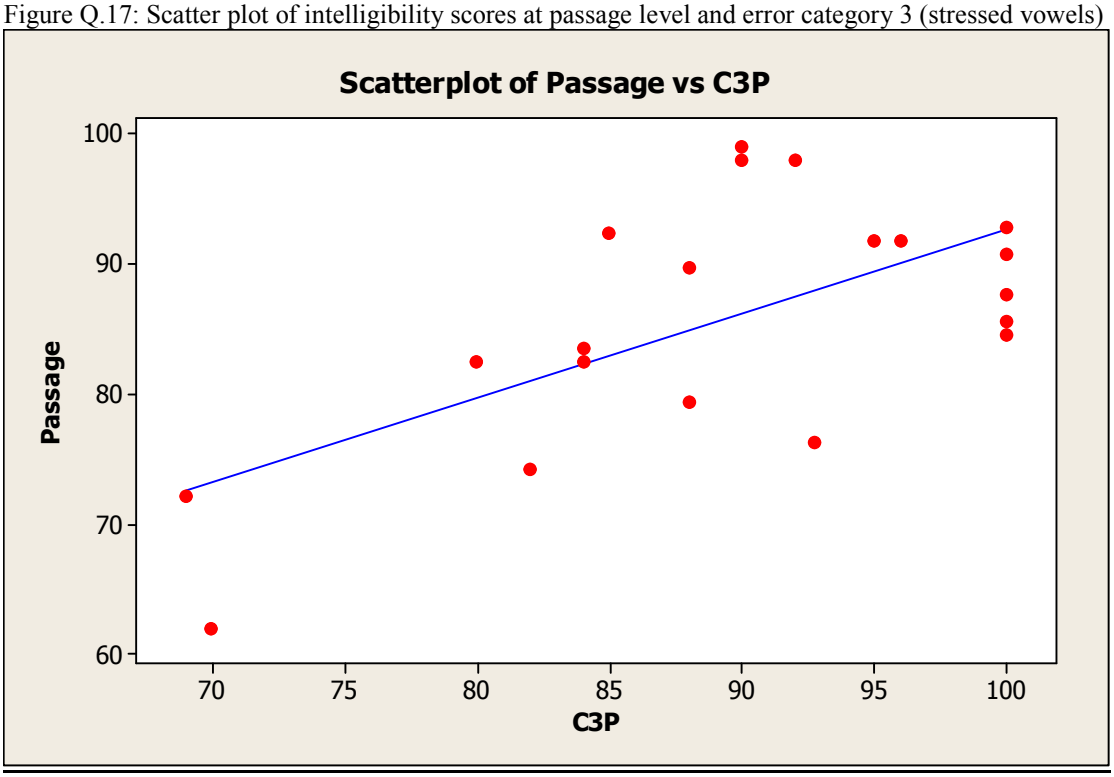


Figure Q.18: Scatter plot of intelligibility scores at passage level and error category 4 (fricatives)

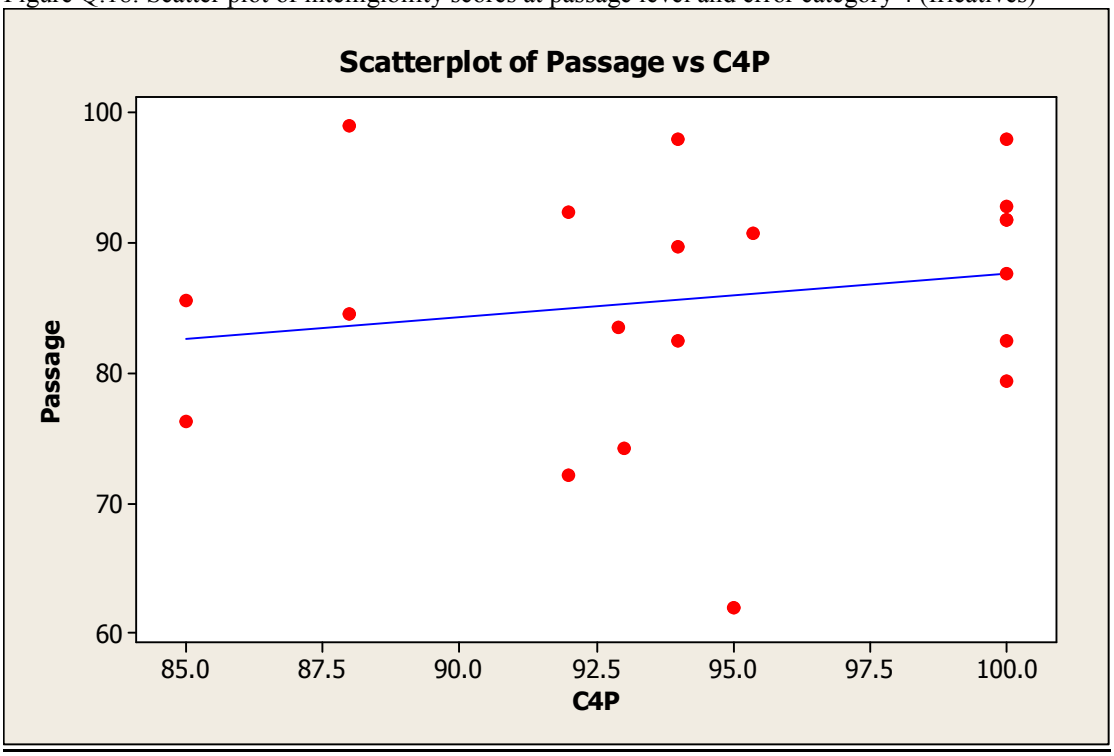


Figure Q.19: Scatter plot of intelligibility scores at passage level and error category 5 (liquids)

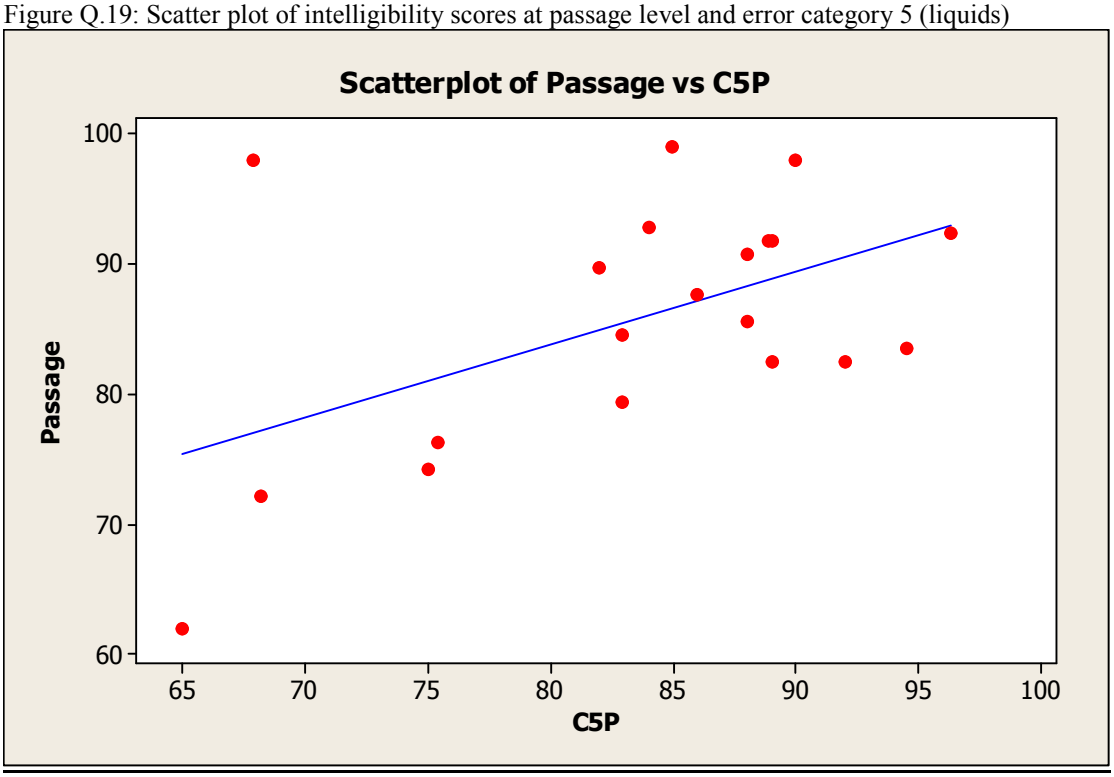


Figure Q.20: Scatter plot of intelligibility scores at passage level and error category 6 (stops)

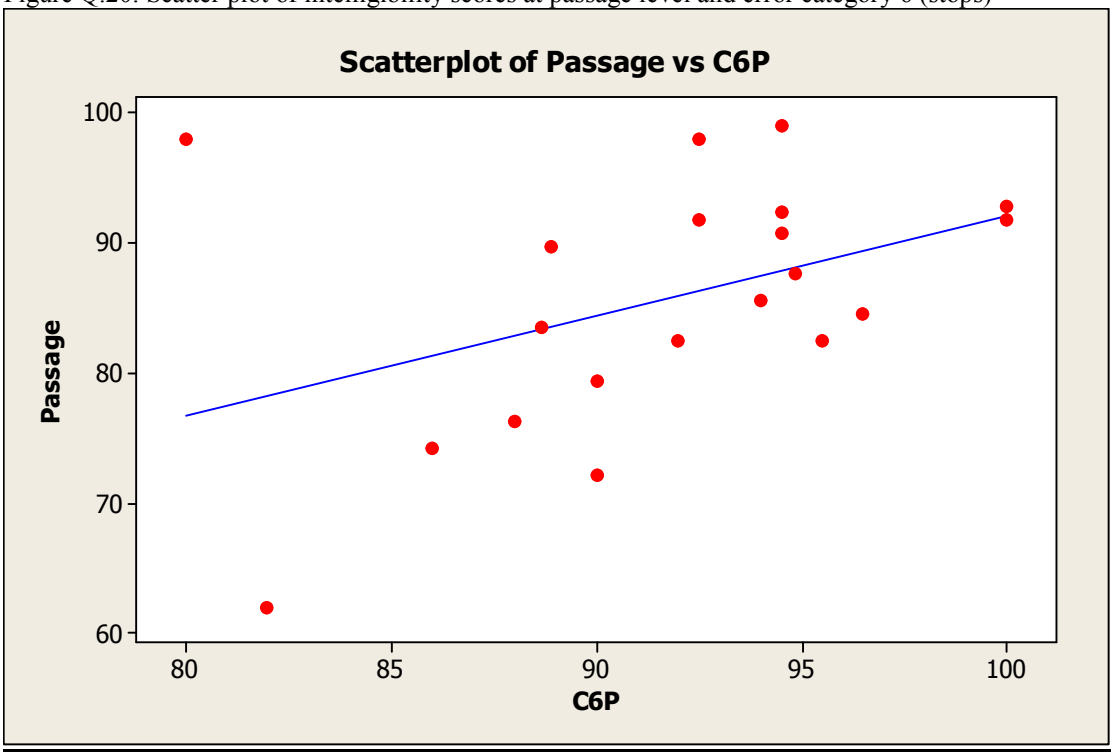


Figure Q.21: Scatter plot of intelligibility scores at passage level and error category 7 (vowel combinations)

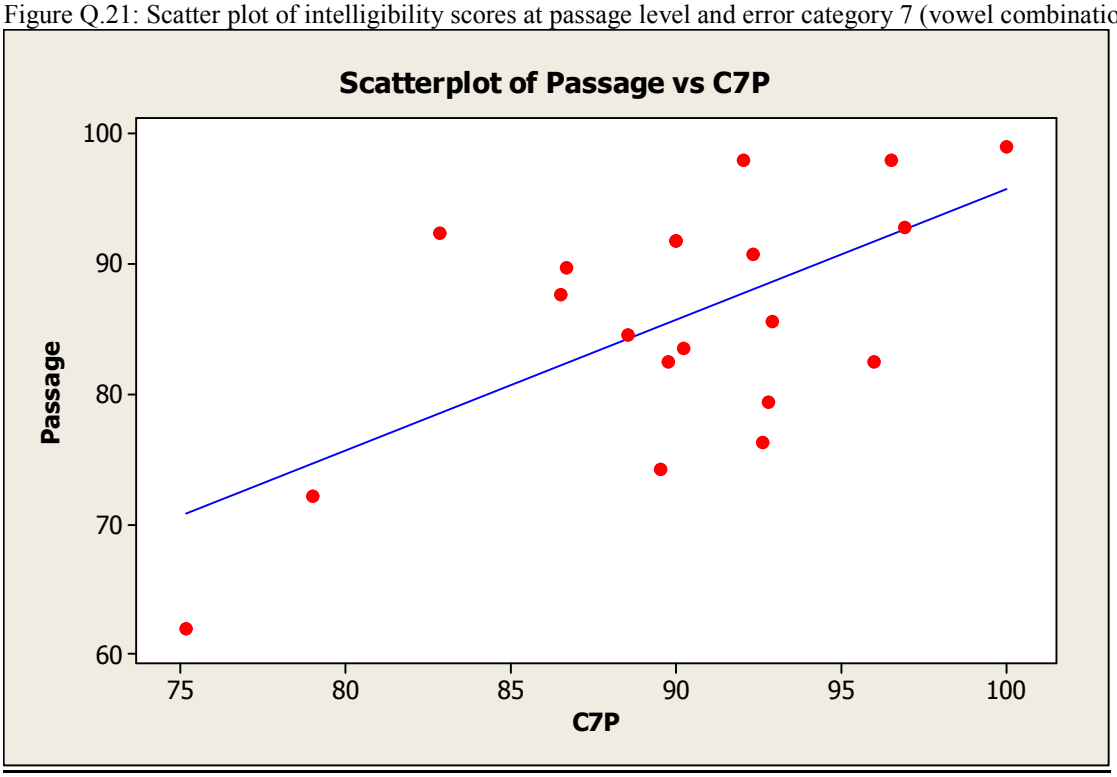
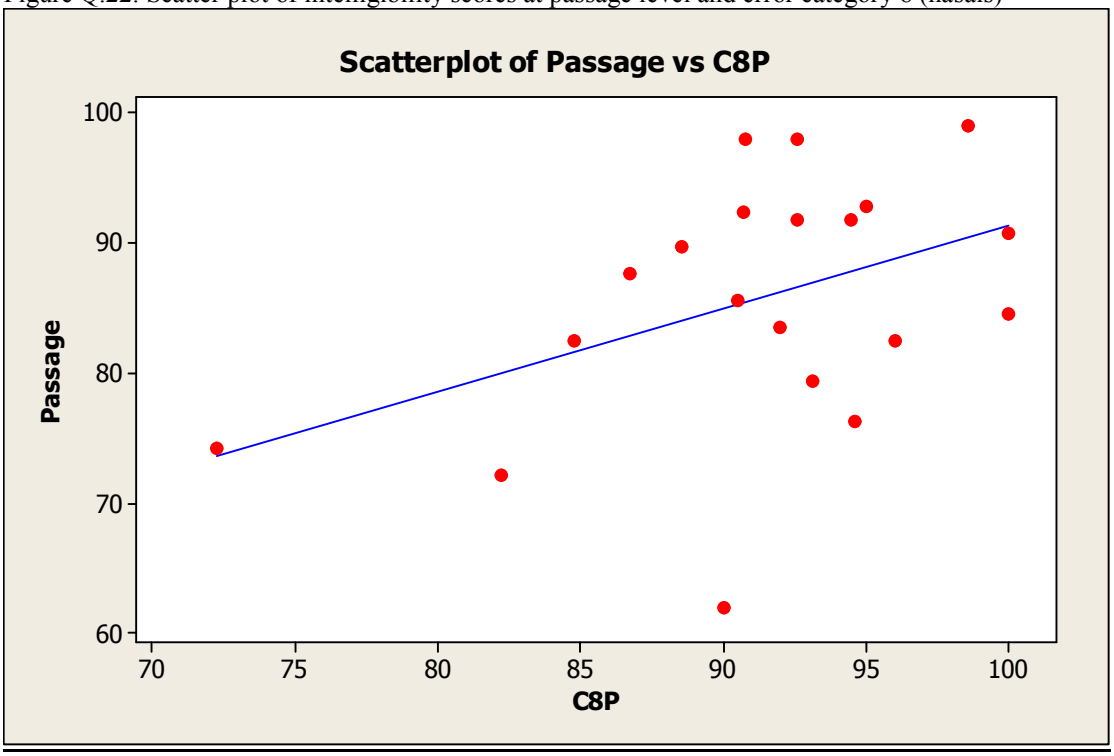


Figure Q.22: Scatter plot of intelligibility scores at passage level and error category 8 (nasals)



Q.23: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 1 (affricates)

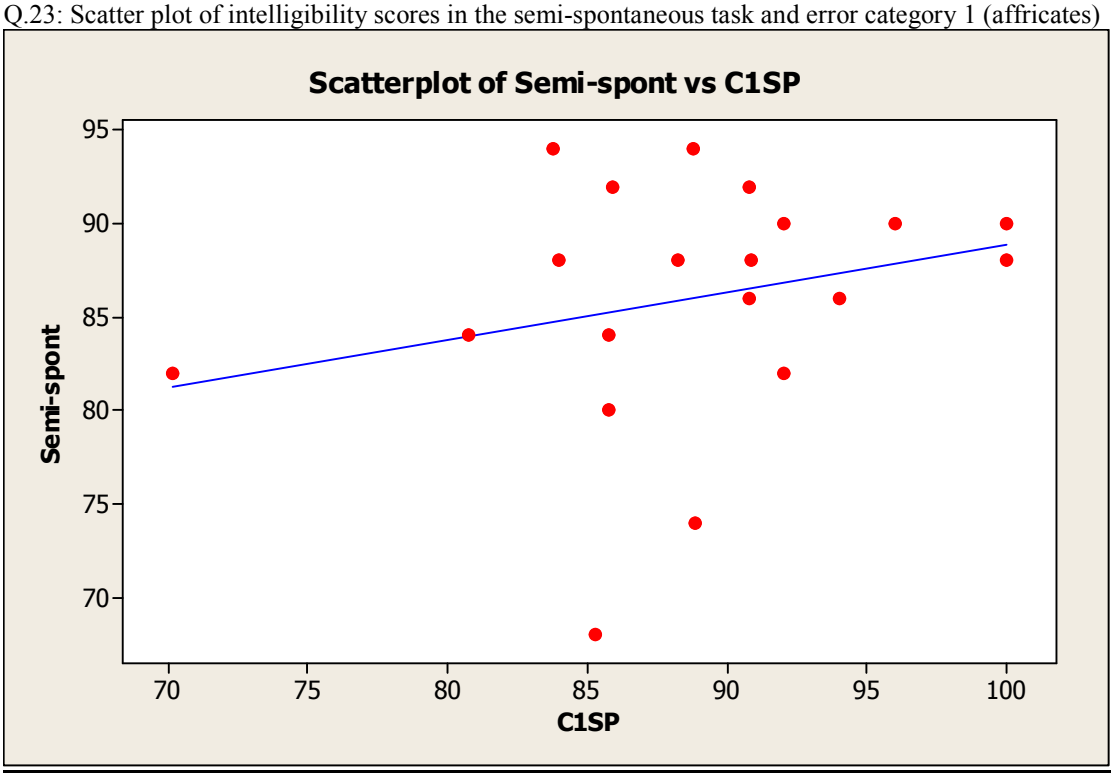


Figure Q.24: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 2 (unstressed vowels)

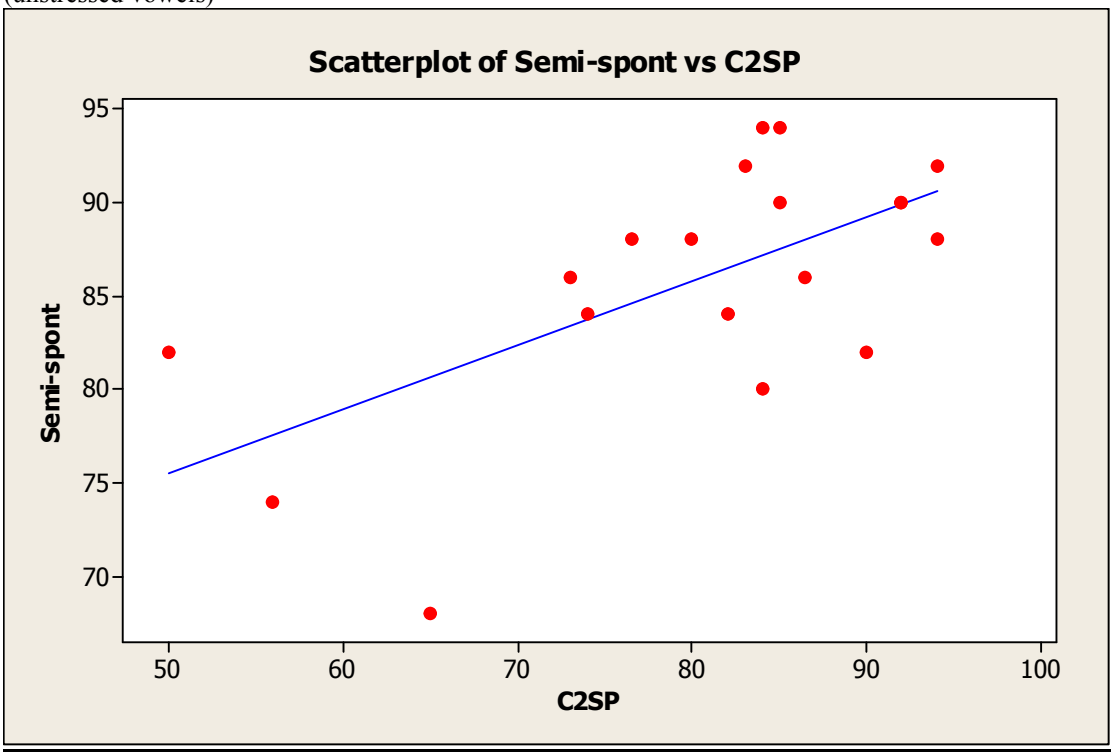


Figure Q.25: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 3 (stressed vowels)

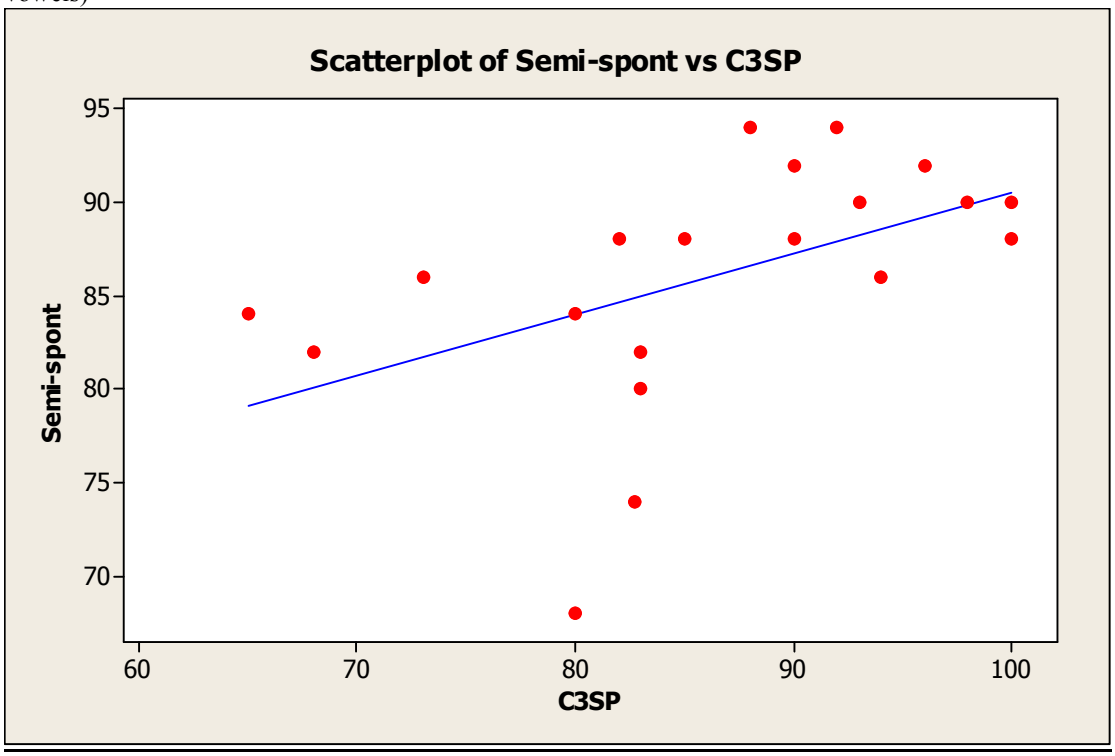


Figure Q.26: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 4 (fricatives)

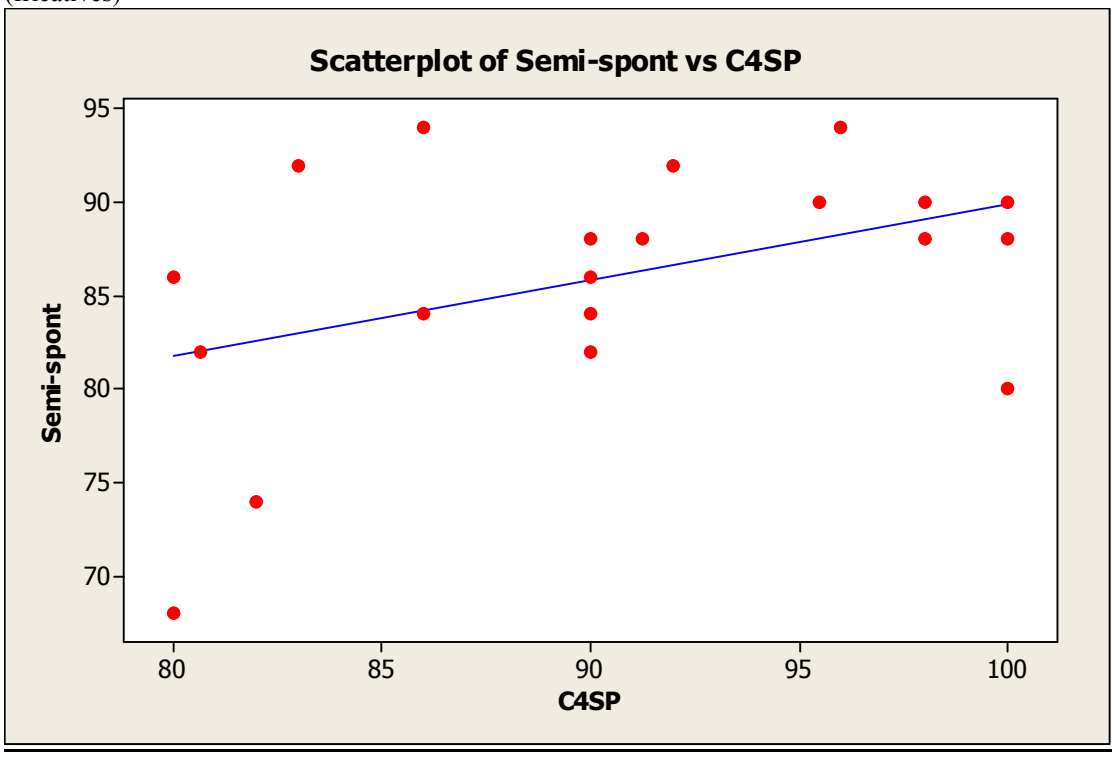


Figure Q.27: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 5 (liquids)

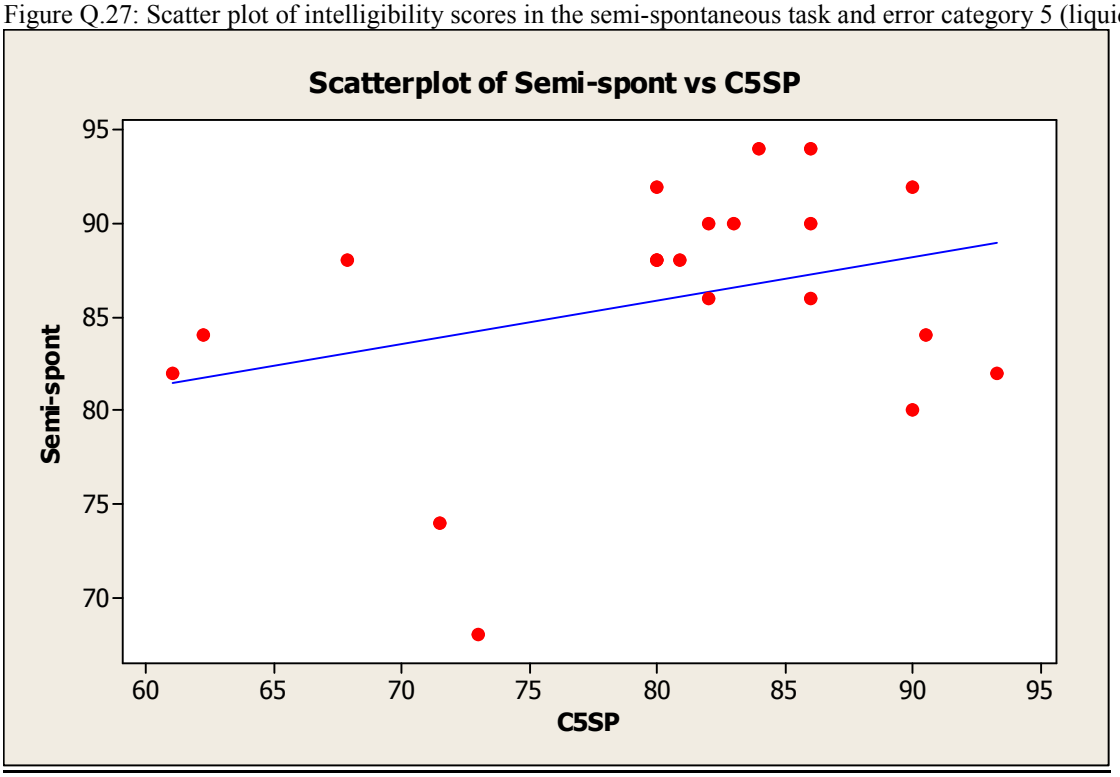


Figure Q.28: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 6 (stops)

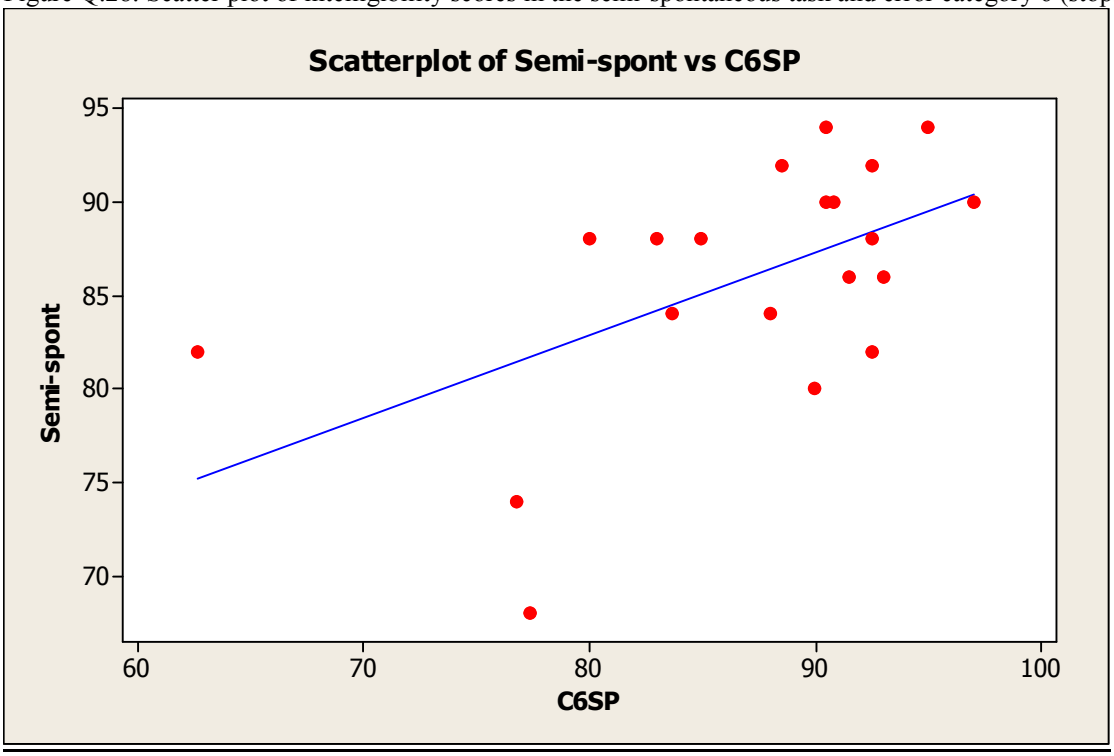


Figure Q.29: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 7 (vowel combinations)

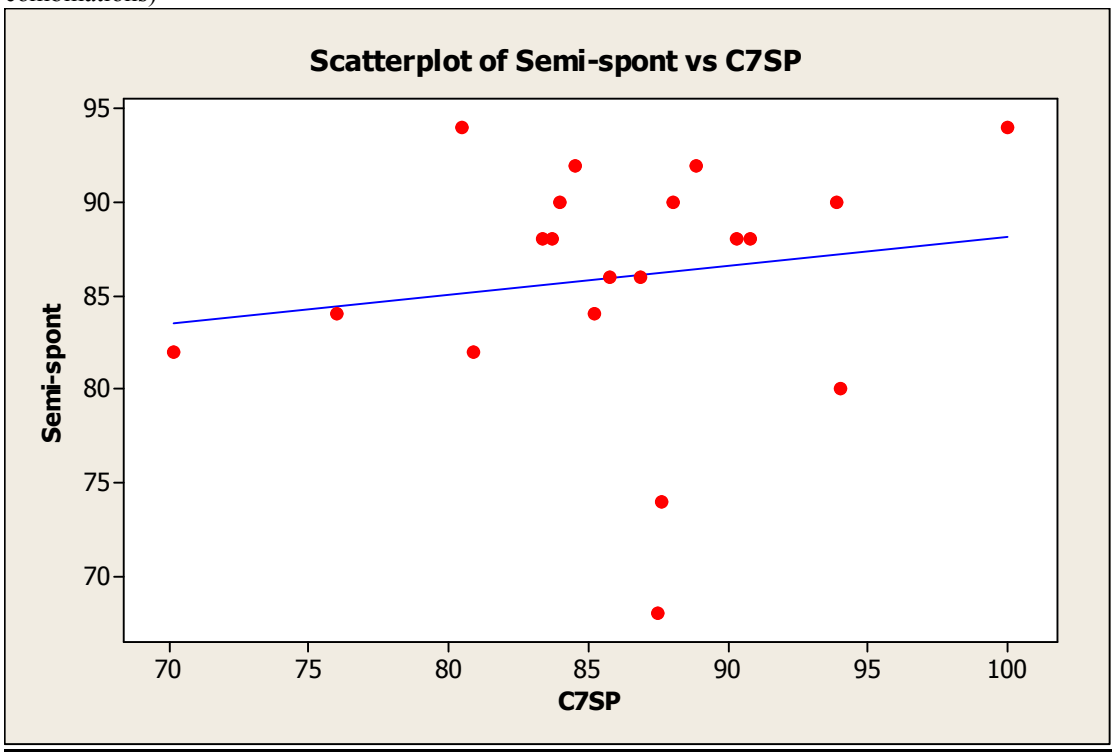


Figure Q.30: Scatter plot of intelligibility scores in the semi-spontaneous task and error category 8 (nasals)

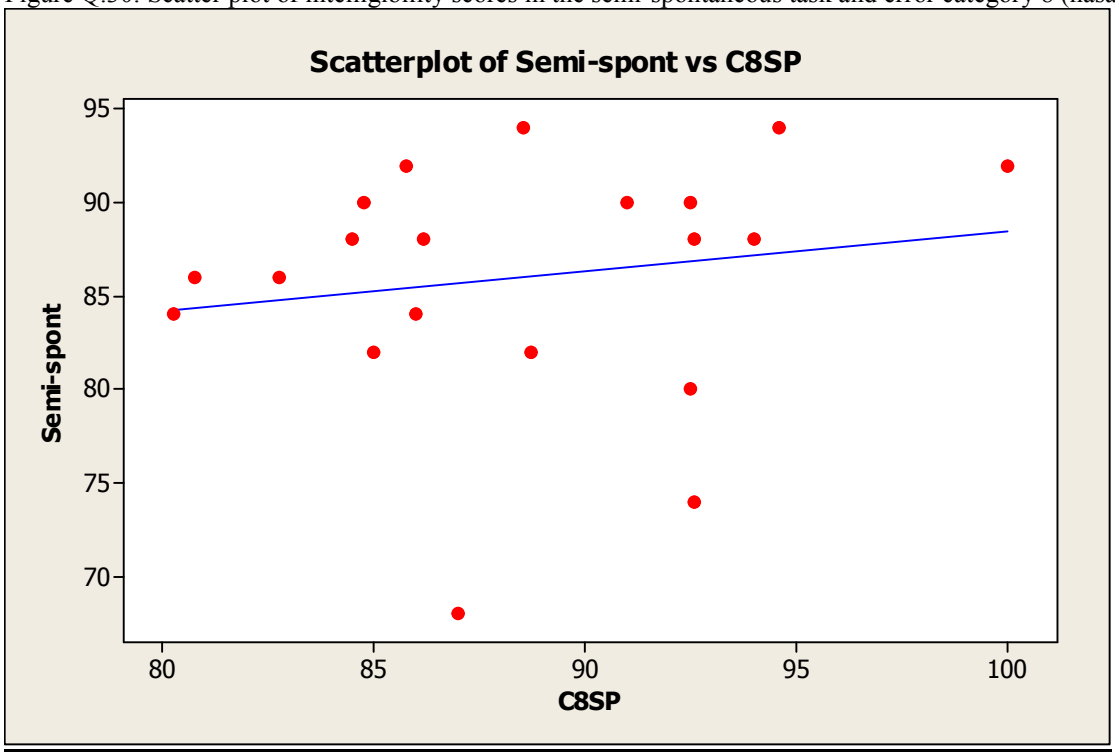


Figure Q.31: Scatter plot of intelligibility scores at sentence level and speech rate

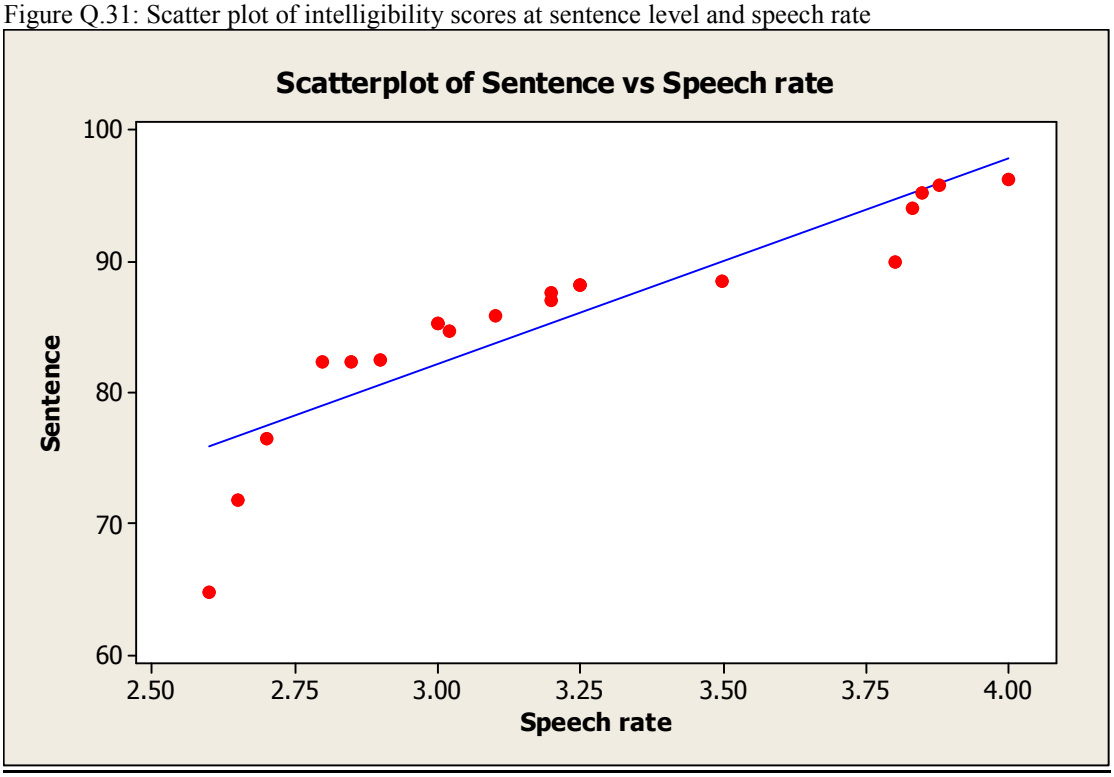


Figure Q.32: Scatter plot of intelligibility scores at sentence level and pause frequency

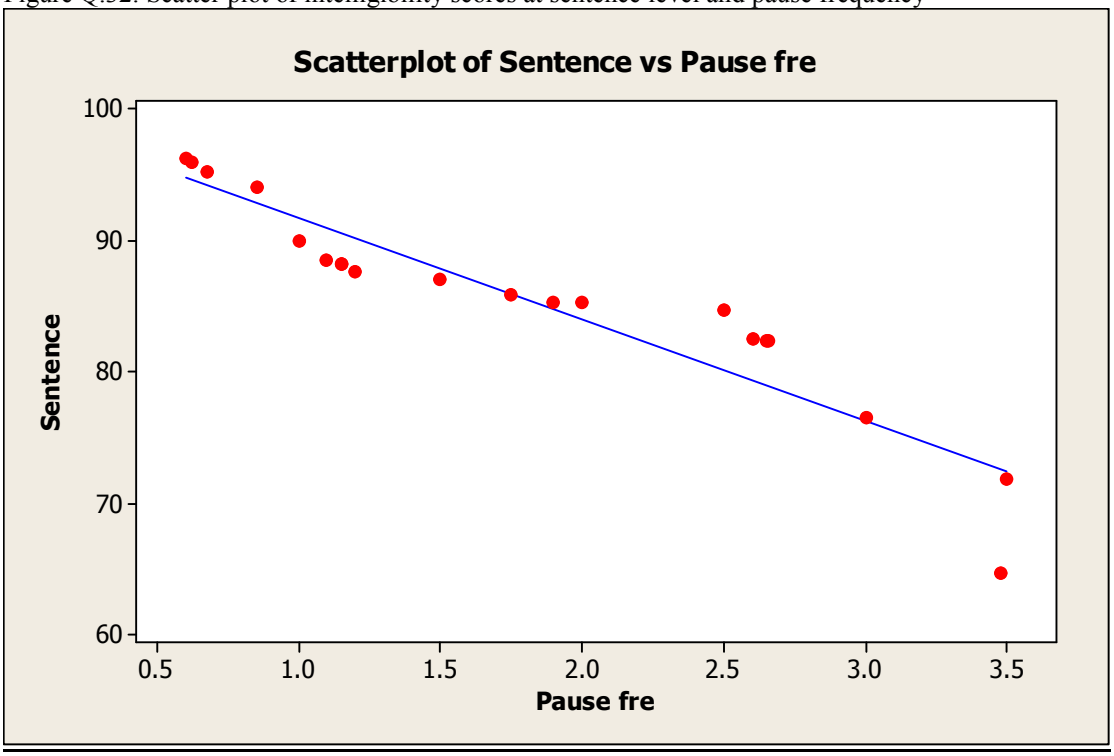


Figure Q.33: Scatter plot of intelligibility scores at sentence level and pause duration

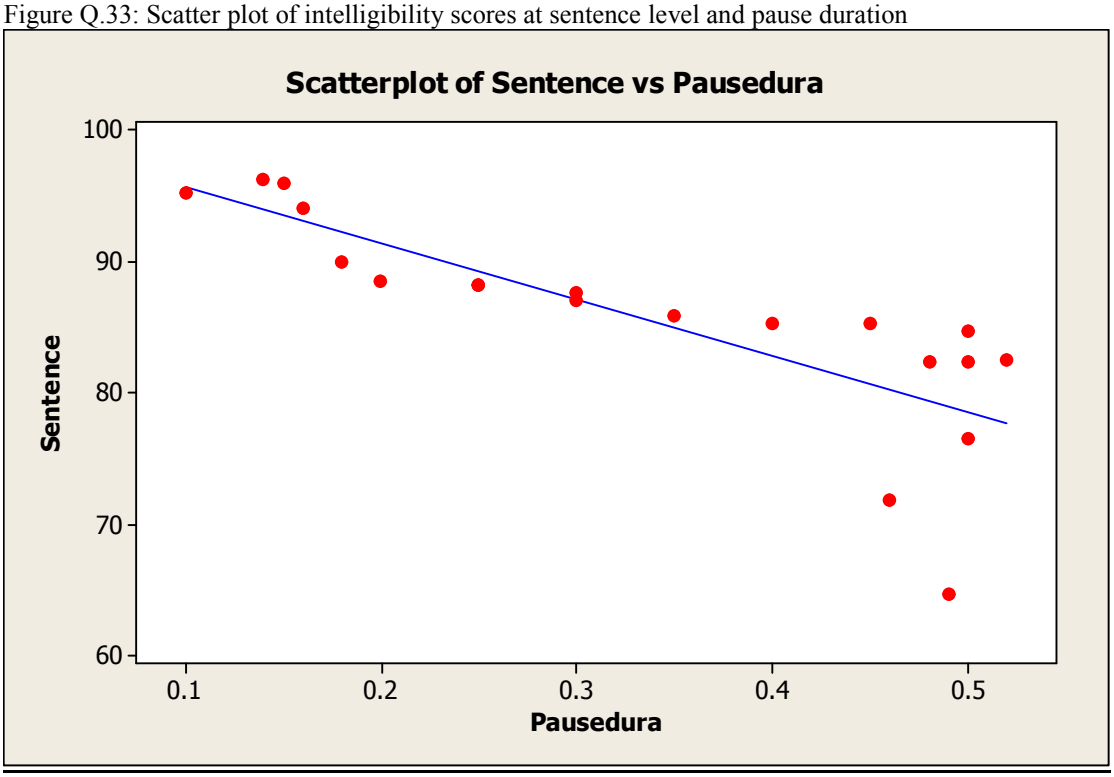


Figure Q.34: Scatter plot of intelligibility scores at passage level and speech rate

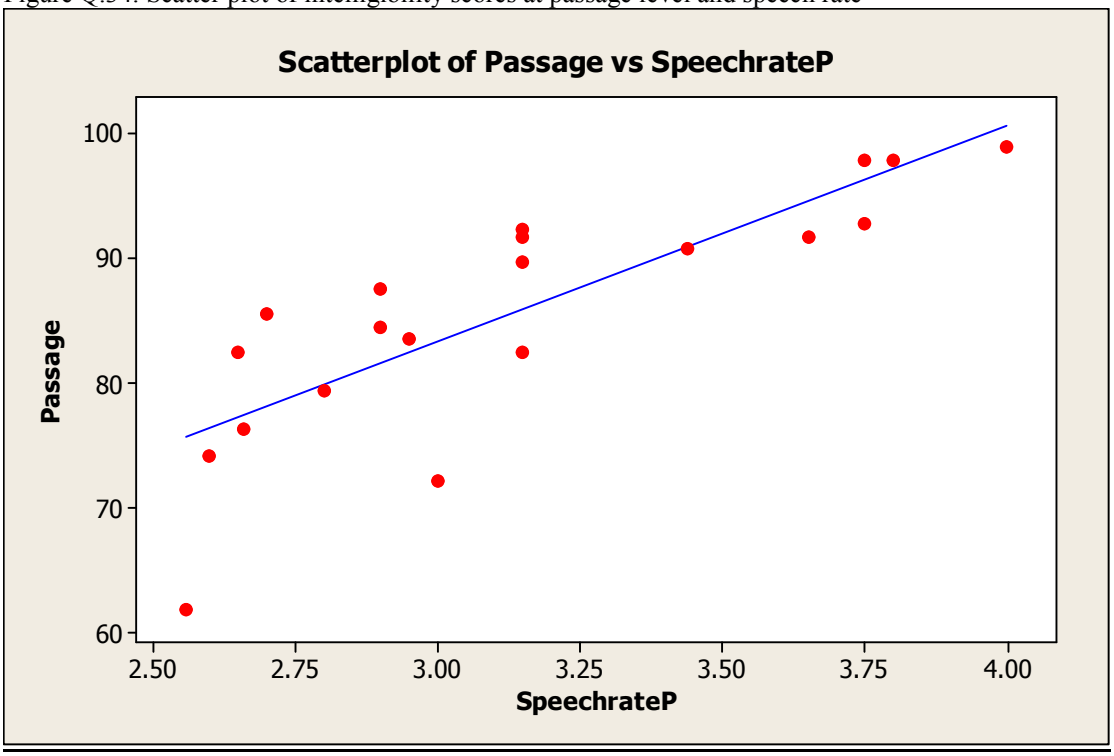


Figure Q.35: Scatter plot of intelligibility scores at passage level and pause frequency

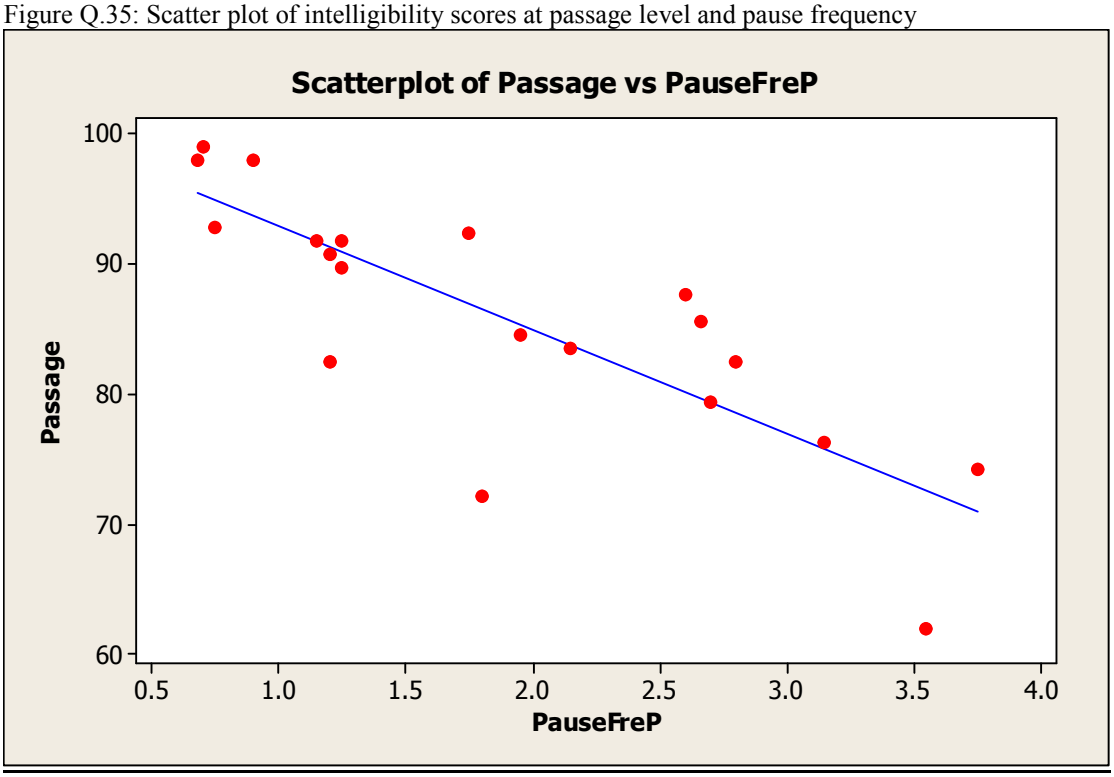


Figure Q.36: Scatter plot of intelligibility scores at passage level and pause duration

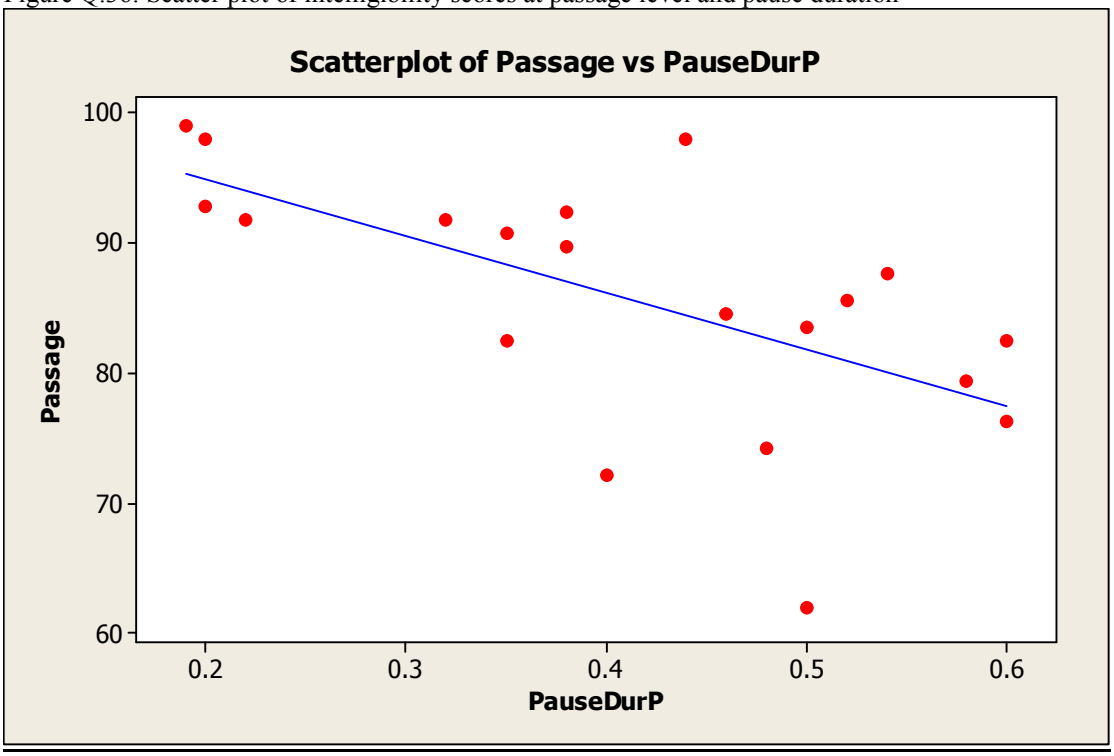


Figure Q.37: Scatter plot of intelligibility scores in the semi-spontaneous task and speech rate

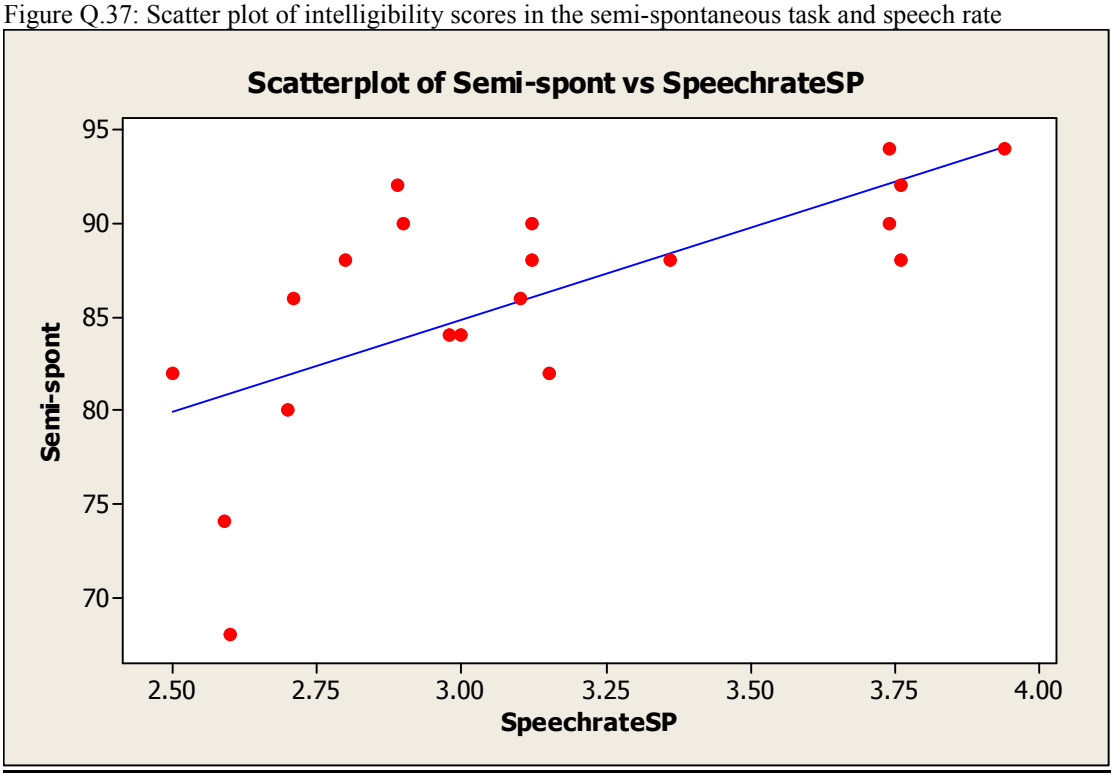


Figure Q.38: Scatter plot of intelligibility scores in the semi-spontaneous task and pause frequency

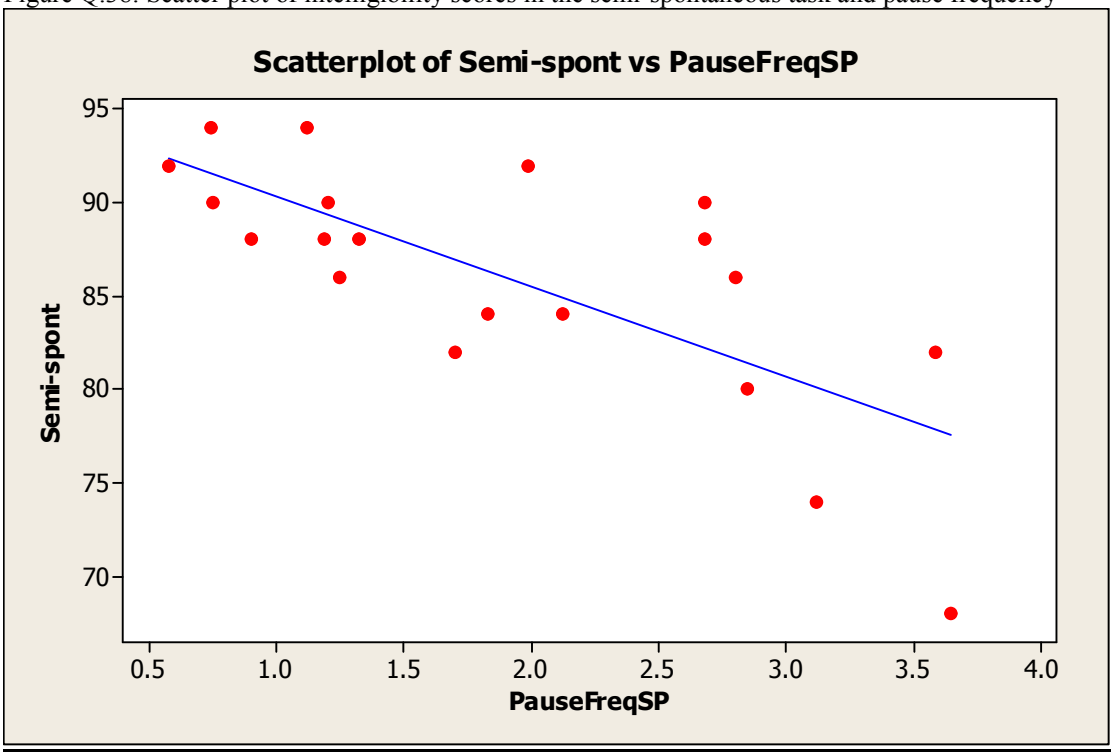


Figure Q.39: Scatter plot of intelligibility scores in the semi-spontaneous task and pause duration