Effects of Speech Style on the Use of Temporal and Spectral Cues in the Production and Perception of a Non-native Vowel Contrast

Hanna Kivistö de Souza

MA dissertation (July 2011)

Supervised by

Dr. Joan Carles Mora

Applied Linguistics and Language Acquisition in Multilingual Contexts

Department of English and German Philology

Universitat de Barcelona

TABLE OF CONTENTS

Table of contents……2
Index of figures and tables……..3
Abstract……...4
Introduction……5

Literature review
1. L1 Spanish learners' perception and production of the English /iː-ɪ/ vowel pair…….....6
2. Relation between second language speech perception and production……..8
3. Speech style effects on oral production……..9

Present Study
1. Research questions……10
2. Method……..12
2.1. Participants……...……..12
2.2. Materials……13
2.2.1. Perception……14
2.2.2 Production……...14
Delayed Sentence Repetition tasks……...14
Spanish/Catalan wordlists…….……16
3. Procedure……...16
4. Data Analysis……18
4.1. Perception……...18
4.2. Production……..18
5. Results……...19
6. Discussion and conclusions……...30

Acknowledgements……...32

References……33

Appendix……...37
1. Mann-Whitney U test for British and American speakers……i
2. Speed of delivery measures……...ii
3. Words used in the English DSR task and Catalan/Spanish wordlists……...iii
4. Paired samples t-tests and Wilcoxon Signed Rank tests for non-natives……..iv
5. ANOVA for speech style……...v
6. Paired samples t-tests for spectral dimensions……..vi
7. Spearman correlations for perception and production data……..…vii

INDEX OF FIGURES AND TABLES

Figures

1. Results of the perception task……...20

2. Realization of /iː/ and /ɪ /…….....22

3. Duration of /iː/ over speech styles…….....…….....25

4. Duration of /ɪ /over speech styles…….……25

5. Effects of speech style on vowel height and frontness…….27

6. Spectral distance over speech styles…….....28

Tables

1. Characteristics of the participants……12

2. Perception data……..20

3. English and Spanish/Catalan vowels……...….21

4. Speech style……...24

ABSTRACT

This paper examined Catalan/Spanish EFL learners' perception and production of the English /iː/- /ɪ / vowel pair. The main objective was to see whether changes in speech style affect the quality and quantity of the vowels and whether the non-native speakers' perception and production were related. Little research has been carried out about the effects of speech style in second language speech production and the research about the perception-production relation has presented inconclusive results. It was hypothesized that speech style would affect the spectral and temporal dimensions of the two vowels. Spanish/Catalan EFL learners have been shown to rely on temporal properties in the perception of this vowel contrast, whereas native English speakers perceive the contrast in terms of spectral and temporal cues. It was expected that the non-native speakers' reliance on duration would be reflected in the production. The results suggest that the Catalan/Spanish bilinguals had created categories for the English vowels, although perception of the vowels was not native-like. A correlation between perception and production was found for the non-native speakers. The non-native speakers were highly affected by changes in the speech style, supporting the hypothesis that in fast speech vowels became more centralized. The results are discussed in light of foreign language phonetic training.

Keywords: L2 vowel production, speech style effects, reliance on duration

INTRODUCTION

Babies have an extraordinary capacity for distinguishing human speech sounds. This ability, however, starts to decline very early on (e.g. Kuhl et al. 2003; Kuhl, 2004; Kuhl et al., 2007). From around six months of age onwards, experience with surrounding language(s) makes the child committed to his/her mother tongue (L1) through the creation of language-specific phonetic categories. In other words, the perception of speech sounds changes from universal to language-specific; whereas babies are able to discern minute differences between sounds, adults identify speech sounds relying on cues that contrast between different categories, ignoring subtle differences (Bongaerts et al. 1997). Adult second language learners' inability to discern minute differences results in a foreign accent.

Adult second language speech learning has a perceptual basis according to two influential models, Best's Perceptual Assimilation Model (PAM; Best, 1995; PAM-L2; Best and Tyler, 2007) and Flege's Speech Learning Model (SLM; Flege 1995). PAM states that non-native speech sounds are classified into native or new non-native representations depending on their perceived similarity to L1 sounds through the perception of articulatory gestures in the speech signal.

The SLM builds on the idea that the capacity to learn new speech sounds is maintained in adulthood and can be used for second language phonological learning. The development of new phonetic categories in the vowel space is believed to depend on the age of first exposure to the L2 (second language) and the robustness of the L1 categories at the time as well as on the perceptual distance between the L1 and the L2 sounds. Category formation for a sound that is perceived as different from an L1 sound is more likely than the development of a category for a sound that is perceived as similar to an existing L1 sound, in which case assimilation of the foreign sound to the existing native category is likely on the basis of equivalence classification and no new category is created. Whether a category is formed or not affects the accuracy in the perception and production of non-native speech sounds (MacKay et al., 2001).

LITERATURE REVIEW

The acquisition of the English /iː/- /ɪ / by L2 learners has been widely studied (e.g. Bohn & Flege, 1990; Flege et al. 1997; Escudero, 2000; Escudero & Boersma, 2004; Morrison, 2008 & 2009; Ylinen et al. 2009). This is partly due to the productive nature of the vowel pair (high frequency appearance and contrasting positions) and partly because EFL learners with varying L1s have shown to experience difficulties in perceiving and producing them in a native-like manner.

1. L1 Spanish learners' perception and production of the English /iː/ and /ɪ / vowel pair

Spanish and Catalan and English differ in the number of vowels as well as in the use of the phonetic properties that are used to distinguish among them. As a result, the Spanish and Catalan learners of English experience difficulties in implementing the /iː/ - /ɪ/ contrast. Whereas Spanish and Catalan have one high front vowel /i/, English has two, /iː/ and /ɪ/ occupying roughly the same portion of the vowel space (Flege, 2003; Mora, 2005). Spanish/Catalan /i/ is spectrally closer to the English tense than to the English lax vowel (Flege et al. 1997), although its duration is more similar to the English lax vowel. Spanish and Catalan also have fewer vowels (5 and 7 [Eastern Catalan] respectively) than English (GA 11, BR 12). It has been hypothesized that speakers of languages with less crowded vowel spaces, like Spanish, will have more difficulties in discerning small differences present in languages with larger vowel inventories, like English (Fox et al. 1995; Frieda & Nozawa, 2007).

Speakers of Standard Southern British English and General American English use a combination of spectral and temporal cues, with more reliance on quality than quantity, in

discerning between /iː/ and /ɪ/ (Escudero & Boersma, 2004; Bohn & Flege, 1990). Spanish and Catalan native speakers, on the other hand, have been shown to rely mainly on duration when distinguishing between these two non-native vowels (e.g. Mora & Fullana, 2007; Cebrian, 2006 & 2007; Cerviño & Mora, 2009; Escudero, 2006; Escudero & Boersma, 2004).

Previous research has encountered difficulties in explaining why Spanish learners of English attend to a secondary cue instead of the primary cue used by native speakers. Since Spanish or Catalan do not use duration contrastively, the reliance on duration cannot be attributed to transfer from the L1 (e.g. Bohn & Flege, 1990). Several hypotheses for the reliance on duration have been given.

An explanation coming from the English language classrooms in Spain is that Spanish learners of English are often taught that the distinction between /iː/ and /ɪ/ is based on duration (e.g. Morrison, 2008): long vs. short. Nonetheless, this would not explain why Spanish and Catalan speakers learning English in naturalistic contexts also use duration to distinguish between these two vowels. As suggested earlier, a more likely explanation can be found in the vowel inventories of the two languages. It is possible that the smaller vowel inventory of Spanish/Catalan speakers does not make the spectral differences between the English tense and lax salient enough to be discerned (Bohn & Flege, 1990). It has also been proposed that duration has some universal properties that make it available even for second language learners whose L1 does not use duration contrastively. According to Bohn (1995), if spectral cues are not salient enough and temporal cues are present, second language learners will attend to the latter. Another hypothesis following the same line of thought presented by Escudero and Boersma (2004) suggests that opting for the creation of a duration distinction is easier to implement and more productive than splitting the existing Spanish category, because by adding a duration distinction to the five existing Spanish vowels, the learner obtains 10 interlanguage English vowels.

According to Francis and Nusbaum (2000) learning a new phonetic contrast is a matter of adjusting the attentional weight given to individual acoustic dimensions. Following this, we could say that Spanish/Catalan learners of English have to learn to

redirect their attention to the spectral dimension in order to discern the English /iː/ - /ɪ/ vowel pair in a native-like manner.

2. Relation between second language speech perception and production

Research on second language speech perception and production has been conducted extensively, however, fewer studies try to relate them. As different as the two processes might be in nature, they are usually thought of as at least partially related as both of them are necessary for human communication.

Second language research addressing the relationship between perception and production has so far proposed four inconclusive accounts i) perception leads production, ii) production leads perception, iii) perception and production are related and iv) the development of perception and production are unrelated.

Koerich (2006) studied Brazilian EFL learners' perception and production of a paragogic /i/ with a sentence reading and a discrimination task. She discovered that the students who produced the paragogic /i/ were also the ones who were unable to discriminate between CVC and CVCi. Also Flege et al. (1999) found a clear relationship between perception and production in their study of L1 Italian speakers' English vowels in which perception was assessed with a categorical discrimination test and production with intelligibility ratings.

Other studies, nonetheless, have failed to find a clear relationship. Flege et al. (1997) found an overall relation between the accuracy in perception and production of the English /iː/ - /ɪ/ contrast. However, the Spanish speakers in the study seemed to identify the English tense and lax vowels well, but failed to produce a significant spectral distance for them, suggesting that Spanish learners of English perceived vowels better than produced them. The opposite was found by Baker et al. (2008). In their study of L1 Korean adults and children, children outperformed adults in the production of English /ɪ/ and /ʊ/, but not in perception, leading the researchers to hypothesize that perception and production are more related in early learners than in late learners. Nevertheless, in Tsukada et al.'s (2005) experiment studying an analogous population, Korean children were shown to produce and

perceive English vowels more accurately than Korean adults, indicating a relation between perception and production. Yet, there are studies that have found no relation between second language speech perception and production. Peperkamp & Bouchon (in press), for example, failed to find correlations between perception and production of English /iː/ - /ɪ/ vowel contrast in French/English bilinguals.

The mixed results obtained in perception-production studies are not easy to explain. Differences in data elicitation methods, participants (L1, language experience and use, proficiency, age of onset of L2 learning and age at task) and measures could explain the heterogeneous results. Different results can be obtained even within the same participants depending on whether they are analyzed at individual or group level (Flege, 1993). As Peperkamp & Bouchon (in press) remark, comparisons between perception and production are difficult due to the different nature of the experimental methods used to measure performance in both. To recapitulate, it is clear that defining a perception-production relationship is not straightforward nor static. On the contrary, it seems that Perception-production relationship in L2 is affected by the degree of competence in L2, age of acquisition and degree of exposure and might differ according to the class of sounds under examination (e.g. consonants-vowels ; initial-final position and so forth) (Llisterri, 1995).

3. Speech style effects on oral production

Speech is characterized by variability: in our daily conversations, we listen to males, females, children, native speakers, foreigners, different dialects, loud and soft voices, fast and slow talkers. For a foreign language learner, adapting to this variation requires additional attentional resources. To the best of our knowledge, no studies have previously investigated foreign language learners' production of non-native vowels in slower and faster speaking rates.

The aim of the study was to determine what occurs to the /iː/ - /ɪ/ vowel pair in different speech styles. The three speech styles studied in this paper were established by means of temporal properties and labeled as careful, citation and fast. Citation form was defined as the "normal" speech rate. Fast speech was understood as a more casual speech

style present in everyday conversations. Careful speech, on the contrary, was defined as slow-paced, clear and hyperarticulated. This followed the careful-casual speech continuum proposed by Johnson et al. (1993). Based on previous research (Johnson et al., 1993; Johnson, 2000; Frieda et al., 2000), it was expected that in careful speech, the vowels would have more extreme positions in order to make the spectral differences more prominent and that in the fast speech, reduction processes would take place resulting in more centralized vowels (Deterding, 1997).

Speech rate was also seen as a way of determining whether category formation had taken place. Following Schmidt and Flege's (1996) rationale, accurate production of the target L2 vowel in normal speech rate could be the result of conscious attention and imitation instead of actual category formation. It was hypothesized that if the Spanish/ Catalan speakers had in fact created a category for the L2 sounds, those categories should be robust enough to persist speech rate changes.

PRESENT STUDY

This paper studies foreign language speech learning and category formation through Catalan/Spanish learners' perception and production of the English tense-lax vowel contrast. The aim of the study is to assess non-native learners' use of spectral and durational cues and to determine how the Spanish/Catalan learners' usage of these cues is affected by speech style changes in production and duration manipulation in perception.

1. RESEARCH QUESTIONS

The following research questions and hypotheses were posed:

1. **Do Catalan/Spanish bilinguals rely on duration in perception?**

   Reliance on duration in perception was expected based on previous research (Mora & Fullana, 2007; Escudero, 2006; Escudero & Boersma, 2004). It was hypothesized that this would be transferred to production, at least to a certain degree.

2. **Are the Catalan/Spanish bilinguals assimilating the English /iː/ and /ɪ/ to their native /i/ or have new L2 categories been formed?**

   It was expected that Catalan/Spanish bilinguals would not have fully developed English tense and lax vowel categories. It was also assumed that they would not be using their exact existing category either. It was hypothesized that their interphonology vowels would be somewhere in between the L1 and L2 categories.

3. **Do changes in speech style affect the accuracy of the production of English /iː/ and /ɪ/?**

   Based on previous research, it was hypothesized that speech style would affect the accuracy of the production of these vowels. Vowel duration was expected to be affected the most, since several studies (e.g. Miller & Volaitis 1989; Volaitis & Miller, 1992; Schmidt & Flege, 1996) have shown that as speech rate becomes slower, the syllable duration and VOT increase. For the spectral dimension, it was thought that the vowels would be the most distinct in the careful speech and less distinct (more centralized) in the fast speech. The non-native speaker's accuracy in production was expected to worsen in the fast speech in comparison to normal speech if categories for /iː/ and /ɪ/ had not been formed.

4. **Is there a relation between perception and production? If so, what happens to this relation when speech style is modified?**

   Based on previous research, it was assumed that a relation between perception and production would be found. However, the actual nature of this relation was not clear.

2. METHOD

2.1. Participants

In total, 22 students of English Studies taking a first year course in English Phonetics and Phonology volunteered to take part in the study and received course credit for their participation. Of the 22 students, two were left outside analysis due to their L1 (Chinese and Urdu), since the study concentrated on the perceptual cue weighting of L1 Spanish/Catalan EFL students.

Non-native experimental group (n=20)						
L1	**Age (m)**	**Sex**	**Daily use of English (m)**	**Outside university exposure to English**	**Considers him/herself fluent in English**	**Familiarity with English dialects**
Spanish/ Catalan 100%	22.85 (4.90)	Female 60% Male 40%	17.72% (14.27)	**high**: at work/with native friends 70% **low:** read and watch TV 30%	Yes 80% No 20%	British 70% American 20% Both 10%
Native English control group (n=7)						
Native dialect	**Age(m)**	**Sex**		**Daily use of English (m)**	**Daily use of Sp/Cat (m)**	
British 40% American 60%	27.7 (4.6)	Female 100% Male 0%		67.8 % (13.81)	29.29% (13.36)	

Table 1. **Characteristics of the participants.**

90 % of the non-natives were bilingual in Catalan and Spanish (35% dominant in Spanish, 30 % dominant in Catalan and 25 % considered themselves balanced) and the remaining 10 % (two participants) were monolingual in Spanish. All participants reported normal hearing.

In order to compare the Spanish/Catalan bilinguals' perception and production to native standards, a control group of native English speakers was recruited. Seven L1

English natives volunteered for the experiments. Three of the natives were from Southern England and four from the United States. British and American participants' vowel height, frontness and duration measures as well as perception scores were submitted to a set of Mann-Whitney U tests (see appendix 1). None of the tests yielded significant differences, indicating that the British and American participants did not behave significantly differently.

2.2. Materials

In order to relate perception to production, tasks for both dimensions were used. A vowel identification task was chosen to study perception. Vowel discrimination tasks require ability to hear differences between sounds, meaning that the categories under study do not necessarily have to be established in order to perform successfully (Flege, 2003). In vowel identification, on the other hand, the subject is obliged to use his/her existing vowel category representations in order to identify the vowel.

Production was studied by means of a Delayed Sentence Repetition (DSR) task (Flege et al., 1995). DSR task was chosen in order to minimize direct imitation from sensory memory, list effects and speech dysfluencies. DSR tasks make the subject rely on his/her mental representation of the sound by making direct imitation impossible. At the same time, DSR tasks provide highly controlled speech since every participant hears the same model and produces the same sentences. Since the target words were only presented aurally, reliance on orthographic cues was minimized.

2.2.1. Perception

The vowel identification task used in this study was previously used in Moya-Gale (2010), Mora & Cerviño-Povedano (2010) and Gilabert et al. (2010). The task was an adaptation of Ylinen et al.'s (2009) vowel identification task in which phonetic cue weighting of Finnish EFL learners was studied by means of normal and duration manipulated stimuli.

The participants heard a word from the headphones and were shown two written options constituting a minimal pair (e.g. *beat –bit*) on the computer screen. Their task was

to identify the word they heard by choosing the correct option. The task began with a set of practice trials and the totality of the task took approximately 10 minutes.

The stimuli used in the task were spoken by six Southern British natives, three males and three females. There were 12 target words (six minimal pairs with the /iː/ - /ɪ/ contrast). The participants heard in total 72 natural tokens of the target words. In order to better study the suspected reliance on duration, the naturally produced tokens were manipulated for duration so that the tense vowel received the duration of the lax vowel and vice versa. The quality of the vowel was left intact. This produced in total 72 manipulated tokens. Altogether, the participants heard 144 target words, half natural, half manipulated, in a randomized order.

2.2.2. Production

Production was elicited by means of Delayed Sentence Repetition Task for the English vowels and wordlist reading for the Spanish/Catalan vowels.

Delayed Sentence Repetition tasks

In the DSR task that was created, the participant was presented with the target word in a carrier sentence. The task was to listen to the dialogue and repeat the target sentence after the second repetition of the distracting sentence. There were three DSR tasks for the three speech styles (careful, citation and fast) studied. Each of the tasks had the same 12 minimal pairs, in order to make comparisons possible, but the order of the items varied from task to task.

The words used in the DSR task were chosen taking into account syllable length, postvocalic voicing context and word familiarity as well as reliability in taking vowel format measurements (see appendix 2). All the tokens were monosyllabic CVC words. Half of the target vowels were followed by a voiced consonant, half of them were followed by a voiceless consonant, since vowel duration varies according to the voicing of the following consonant (e.g. Morrison, 2008). All the words to be used were intended to be familiar for the non-native participants in order to avoid lexical knowledge effects (Mora, 2005).

The target words were embedded into the carrier sentences: "*beat* is the next word" (citation form), "I didn't say *bit*, I said *beat*" (careful speech) and "I would say *beat* is the next word" (fast speech). The careful speech was operationalized through an exaggerated contrast as in Frieda et al. (2000) and only the second word was acoustically analyzed in further analyses.

In order to create the DSR tasks, a native Southern British male speaker was recorded speaking the 24 target sentences in the three speech styles. Since the participants were using the male speaker as their model, special care was taken that the model's productions were good exemplars of the token words for each speech style in question.

The tokens to be used in the DSR task were chosen from the repetitions by computing the means for quality and duration and by choosing the word whose quality and duration measures were closest to the mean value. If the token was unclear or creaky, the next closest match was chosen. Once the tokens to be used were selected, they were normalized for peak and mean amplitude and filtered at 60Hz to eliminate any low-frequency noise with GSUPraat Tools.

The vowel durations of the three speech styles were compared in order to make sure that careful speech was slower than citation form and that fast speech was faster than citation form. In addition, a measure of speed of delivery was established by counting the number of segments in the target sentence and dividing it by the duration in seconds of the same sentence. Speed of delivery was submitted to a repeated measures ANOVA with speech style as dependent variable to make sure that the three speech styles differed (see appendix 3). The results confirmed that the three speech styles differed significantly in their speed.

The three DSR tasks were created with Goldwave, a digital audio editor. Mini-dialogues of distracter sentence, target sentence, distracter sentence and silence were created for each target word and the 24 mini-dialogues obtained like this were merged into one task. In order to create additional time pressure, the silences in the fast speech task were shorter than in the citation form task (2.5sec. vs. 3sec., respectively).

The order in which the target words were presented was the same for all the participants, but the order changed from task to task so that the participants would not be using the first presented task as their model for the following tasks. Each of the 24 target

words was presented only once, although using repetition was discussed, but abandoned due to task length. In addition, the use of minimal pairs already offered repetition of the target vowels in different contexts. The DSR tasks took between 3:14min (fast) and 4:50min (careful) to complete.

Spanish/Catalan wordlists

A list of Catalan and Spanish words (appendix 2) containing the Catalan/Spanish vowel /i/ was created in order to make comparisons between L2 and L1 vowels possible. The participant was asked to read the list in his/her strongest language.

Ten target words were created for each L1. The items on the lists were aimed to be the best possible phonological matches for the English target words in order to make comparisons reliable. Because of this in cases where exact matches were not possible (e.g. *hit* > *jida*), non-words were deemed to be the best option.

Due to differences between the languages in question, all the Catalan and Spanish words were disyllabic, with the exception of *ida*, with stress on the first syllable where the target vowel /i/ appeared. There were three repetitions of each word in random order and the syllable stress was indicated by underlining. The participants were instructed to read the words as if they were in isolation, in their normal speaking speed.

3. PROCEDURE

The participants were tested individually in the University of Barcelona phonetics laboratory. The production tasks were done inside a soundproof booth in order to guarantee good sound quality. The perception task was carried out in a quiet corner of the room.

Each participant carried out the tasks in the same order: questionnaire, practice, DSR, native wordlists and perception task. The duration of the testing session was around 30 minutes for native participants and 45 minutes for non-native participants.

After filling in a language background questionnaire, the non-native participants had the opportunity to familiarize themselves with the target words through a Microsoft Office Power Point presentation. The participant saw all the target words written on a computer screen in a random order and had the possibility to click the words to listen to them. After this, the participant was asked if there were some words that were not familiar and translations to L1 were provided if needed. The main goal behind the familiarization task was to make sure that the participants knew the test words, since lexical knowledge has a facilitating effect on the perception of non-native phonemic contrasts (Mora, 2005). Another goal was to make the participants feel at ease, before starting with the production experiments.

DSR tasks were carried out before the perception task in order to avoid influence from the perception task (Flege, 1993; Frieda et al., 2000). The order of the production tasks was the same for all the participants: citation, careful speech and fast speech. The chosen order made the differences between the speech styles pronounced by presenting the most dissimilar styles sequentially.

The participants were instructed to pay attention to the male speaker's pronunciation and speed. In the careful speech, the participants were instructed to exaggerate the two contrasting words and in the fast speech they were asked to imitate the model's speed as much as they could. If the participant forgot the instructions about the speech style or was unable to produce more than two consecutive words, the tape recorder was paused, instructions were repeated and the task was resumed. After completing each task, the participants were encouraged to have a small break. Following the DSR tasks, the non-native participants were asked to read the L1 word/non-word list in their normal speaking speed.

The test session finished with the vowel identification task. The participants were explained that they would hear similar words to those they had produced before and that they would have to identify them by pressing the correct key on the keyboard. They were told that some of the words might sound strange, but that they should not worry about that and try to identify the word. Instructions to answer as fast as possible and guess if unsure were given.

4. DATA ANALYSIS

4.1. Perception

For the vowel identification task, percentages of correct identification were obtained, first separately for each subject and then for the native and non-native groups as a whole. Percentages of correct identification were calculated for the tense and lax vowels in natural and manipulated trials.

4.2. Production

The participants' productions were measured for duration and vowel quality (F0, F1, F2) with speech analysis software Praat (Boersma, P. & Weenink, D., 2011). The duration measures for the vowels were obtained by measuring the vowel from the onset of voicing to a visible decrease in amplitude.

Next the vowel quality measures were converted from frequencies (Hz) into bark scale by using the following formula: $B= 26.81 / (1+(1960/F)) – 0.53$ (Traunmüller, 1997). Bark scale offers a more reliable psychoacoustic measure, making comparisons between perception and production possible. The obtained bark measures were then normalized for speaker characteristics such as gender and vocal tract size enabling the contrasting of the participants' productions. The procedure involved subtracting the B1 value from the B2 value (for vowel frontness) and subtracting the B0 value from the B1 value (for vowel height) (Syrdal and Gopal, 1986). The normalized bark values were then employed to calculate means for the vowel duration, height and frontness for the two vowels in the three speech styles for every participant. Vowel duration, height and frontness means were also calculated for the Spanish and Catalan words separately as well as under a common L1 category.

A measure of speed of delivery (segments/sec) was calculated for each speech style to determine that the non-native participants had been able to follow the instructions and produce speech at three different speeds. Each participants' speed of delivery measures in

the three speech styles were submitted to a repeated measures ANOVA with Bonferroni adjustment (see appendix 3). The results confirmed that all the participants realized three significantly different speech styles.

In order to better study the effects of speech style on the realization of the two vowels, the Euclidean distance between the tense and the lax vowel in all the speech styles was computed. Euclidean distances were also obtained for the Spanish/Catalan vowel and the English vowels for the non-native speakers for the sake of determining the distance between the native and non-native categories.

In the production tasks some data was missing due to the participants' inability to produce the target word as a result of memory constraints and/or time pressure. The overall percentage of missed data for the DSR tasks was 3.4% and 1.4% for the native wordlist[1]. There was no missing data in the perception task. The missing values were replaced by a mean value obtained from the same person's same vowel and same voicing context productions. It was judged that this procedure would have the least effect on the mean value of the existing tokens.

5. RESULTS

Normal distribution of the perception and production data was examined prior to any statistical analyses. The production data was normally distributed, whereas the perception data was skewed due to ceiling and floor effects (native and non-native speakers respectively). Non-parametric tests were used for analyses involving perception data. The results will be discussed by research questions.

1. Do Catalan- Spanish bilinguals rely on duration in perception?

In order to determine whether perception of the manipulated tokens differed significantly from the perception of natural tokens, the ID scores were submitted to a Wilcoxon Signed Rank Test.

[1] One non-native participant only had five wordlist items recorded as a result of a problem with the recording equipment.

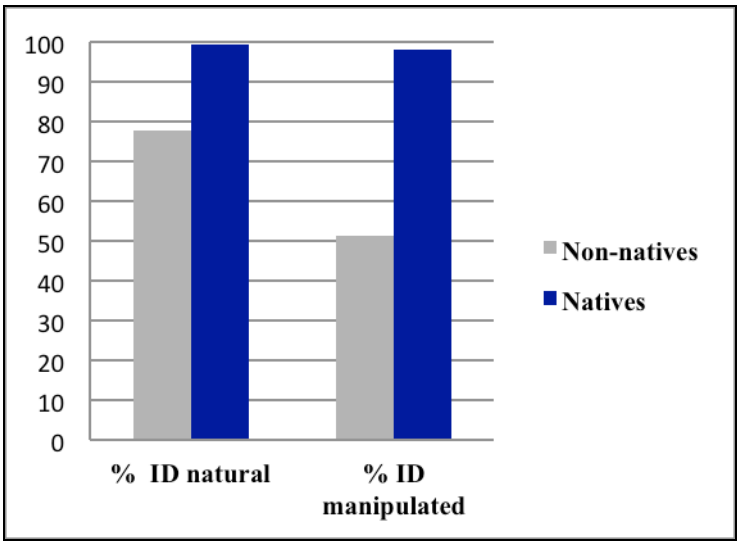

Figure 1.
Results of the perception task. (% of correct identification).

No significant difference was found for the native speakers (see table 2), which was expected since the manipulation of duration cues should not have an effect on native English speakers relying mostly on spectral cues. A significant difference was found for the non-natives (Z= -3.92 ; p<.001), indicating that manipulation of duration cues had a significant effect on word identification. This suggests that the Spanish/Catalan bilinguals relied mostly on duration cues in identifying English /iː/ and /ɪ/ confirming the initial hypothesis and previous research.

	% ID natural	% ID manipulated	Wilcoxon
Non-natives	77.56 (12.16)	51.31 (20.54)	z= -3.92 **p<.001***
Natives	99.20 (1.57)	98.01 (2.09)	z= -1.65 p=.098

Table 2.
Perception data. Mean percent identification of natural and manipulated tokens (*SD* in parenthesis) and the results of the Wilcoxon Signed Rank Test. Asterisk indicates significant difference.

2. Are the Catalan/Spanish bilinguals assimilating the English /iː/ and /ɪ/ to their native /i/ or have new L2 categories been formed?

In order to determine whether the Spanish/Catalan participants assimilated the English /iː/ and /ɪ/ to their native /i/ a set of paired samples t-tests for the vowel duration, height and frontness for the English vowels and the Spanish/Catalan vowel was carried out (see appendix 4). Spectral and durational differences between the two English vowels as well as the English vowels and the Spanish/Catalan vowels were significant in all the measures. In terms of duration, native vowels had a shorter duration than both of the English vowels (see table 3 for descriptives), indicating that the Spanish/Catalan speakers realized both of the English vowels with longer duration than their L1 vowel. Measures for vowel height showed that the L1 vowel was realized higher than both of the English vowels. Vowel frontness measures indicated that the native vowel was more fronted than both of the non-native vowels.

		Non-natives (n=20)	Natives (n=7)
Duration (ms)	**Tense**	202.24 (26.25)	198.36 (22.32)
	Lax	154.89 (29.39)	141.52 (21.66)
	L1	118.87 (13.73)	-
B1-B0	**Tense**	2.28 (0.27)	1.86 (0.26)
	Lax	2.61 (0.40)	3.00 (0.31)
	L1	1.64 (0.29)	-
B2-B1	**Tense**	10.02 (0.78)	11.16 (0.49)
	Lax	9.29 (0.73)	8.88 (0.48)
	L1	10.86 (0.71)	-
Spectral distance	**Tense-lax**	0.86 (0.63)	2.56 (0.85)
	Tense-L1	1.08 (0.76)	-
	Lax-L1	1.88 (0.85)	-

Table 3.
English and Spanish/Catalan vowels. Mean values (SD in parenthesis) for vowel duration, height, frontness and spectral distance for the English vowels (in citation form) and the native Catalan/Spanish vowel.

After determining that the non-native speakers were using different vowels for the L1 and the L2, another set of paired samples t-tests was performed in order to see whether the Spanish/Catalan participants were making a significant difference between the English /iː/ and /ɪ/ (appendix 4). The English vowels differed significantly in temporal and spectral dimensions. The tense vowel was longer, higher and more fronted than the lax vowel (see figure 2). These results indicate that the Spanish/Catalan bilinguals, were not assimilating the target language vowels to their L1 category and that the two target vowels were realized differently.

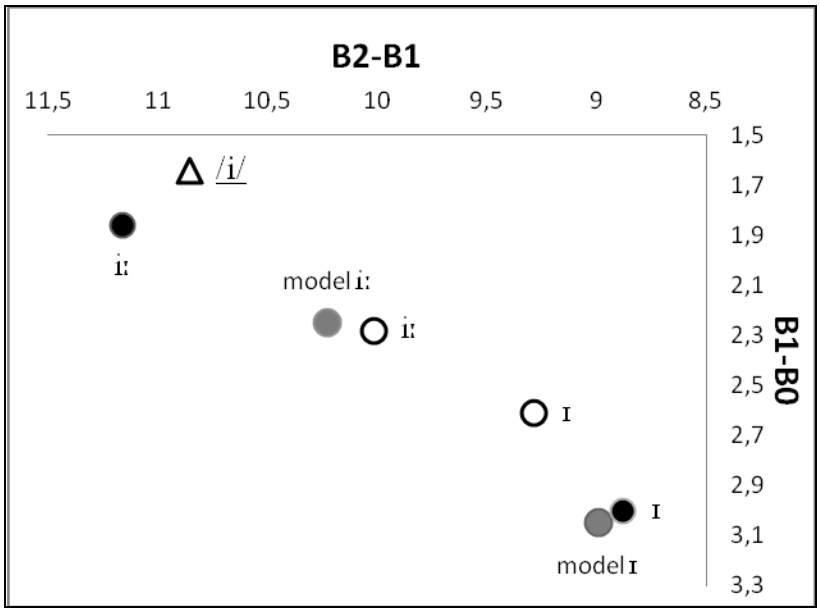

Figure 2.
Realization of /iː/ and /ɪ/. Tense and lax vowel (in citation form) by non-native speakers (empty dots), native speakers (black dots) and the DSR model (grey dots). The native Cat/Sp /i/ is indicated by underlining.

In order to determine if also the participants with lower competence were in fact using different L2 vowel(s), the non-natives were assigned to high and low proficiency groups (n=10 each) through a median split based on their score in the identification of the natural stimuli. It was judged that the identification of natural tokens in the perception task would also measure general L2 phonological proficiency. Wilcoxon Signed Ranks test revealed that all of the dimensions were significantly different for L1 and L2 (see appendix 4), indicating that even the participants with a poorer phonological competence, were able to make a distinction between the L1 and the L2 vowels. Also all of the dimensions were significantly different for the English tense and lax vowels.

Although the non-native speakers seem to be realizing two distinct target vowels, comparison to native control group indicates that the tense and the lax vowel of the native speakers differ significantly in vowel height, frontness, duration and spectral distance (for comparisons, see table 3 in the previous page). The native controls realize /iː/ and /ɪ/ with a significantly larger spectral distance and in fact the Spanish/Catalan speakers' target vowels are located in between the native speaker realizations. However, the non-native speakers' /iː/ vowel is almost native-like if the DSR model is taken as a reference instead of the native control group.

3. Do changes in speech style affect the accuracy of the production of English /iː/ and /ɪ/?

To explore changes in vowel duration, height, frontness and spectral distance due to changes in speech style, a mixed between-within subjects analysis of variance with Bonferroni adjustment was conducted with speech style (careful/citation/fast) as within subjects factor and L1 (native/non-native) as between subjects factor (appendix 5). In addition, a series of paired samples t-tests were conducted to further explore the differences in the spectral dimension (appendix 6)

		Non-natives			Natives		
		Careful	Citation	Fast	Careful	Citation	Fast
Duration (ms)	Tense	228.38 (43.67)	202.24 (26.25)	101.49 (22.40)	252.45 (21.91)	198.36 (22.32)	116.90 (16.92)
	Lax	128.02 (26.42)	154.89 (29.39)	87.37 (17.25)	150.76 (17.94)	141.52 (21.66)	83.21 (17.52)
B1-B0	Tense	2.18 (0.30)	2.28 (0.27)	2.24 (0.42)	1.71 (0.33)	1.86 (0.26)	1.83 (0.43)
	Lax	2.71 (0.53)	2.61 (0.40)	2.45 (0.53)	3.08 (0.59)	3.00 (0.31)	2.54 (0.26)
B2-B1	Tense	10.54 (0.68)	10.02 (0.78)	9.64 (0.90)	11.40 (0.40)	11.16 (0.49)	11.01 (0.45)
	Lax	9.28 (0.78)	9.29 (0.73)	9.16 (0.92)	8.85 (0.48)	8.88 (0.49)	9.08 (0.42)
Spectral Distance	Tense - Lax	1.43 (1.07)	0.86 (0.63)	0.64 (0.69)	2.90 (0.83)	2.56 (0.85)	2.07 (0.77)

Table 4.
Speech style. Means (and standard deviations in parentheses) for duration, height, frontness and spectral distance over speech styles for non-native and native speakers.

Duration

Significant differences (appendix 4) were found for the tense vowel over all speech styles and for the lax vowel in citation-fast and careful- fast. Partial eta squared showed large effect size and no interaction was found for speech style*native language. Both groups produced consistently shorter vowels when speed increased, but the difference was more pronounced in the native speakers (see figures 3 and 4).

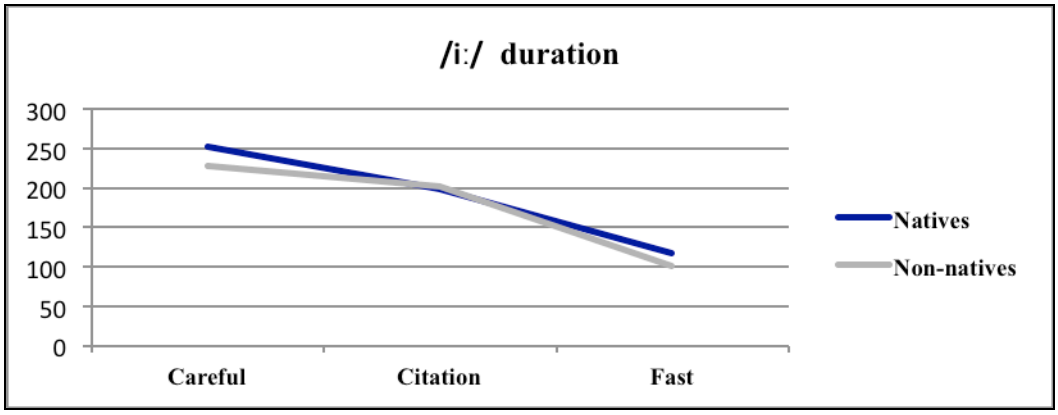

Figure 3. Duration of /iː/ over speech styles.

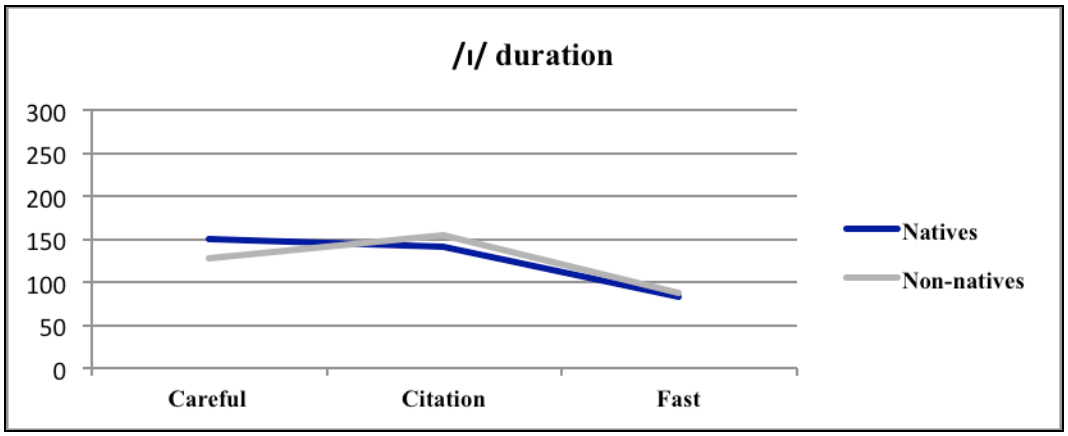

Figure 4. Duration of /ɪ/ over speech styles.

The non-native speakers approximated to the native speakers in citation form, making duration differences between the two groups non-existent in this speech style. The non-native speakers as a group produced the lax vowel with a longer duration in citation form than in careful form. A possible explanation for this is that the non-native speakers were establishing the contrast in the careful speech mostly by means of duration (short-long) and exaggerated the shortness of the lax vowel. The duration dimension was significantly affected by the speech style, confirming the hypothesis of vowel shortening in increased speed and lengthening in slow speed.

Height

Significant differences were obtained for the lax vowel (citation-fast and careful-fast), but not for the tense vowel. Partial eta squared indicated a large effect size and there was no interaction between speech style and native language (see appendix 4). Further t-tests showed that the height difference between careful and fast speech styles was significant for non-natives and natives, but that the height difference between citation and fast was only significant for the native speakers (see appendix 5). These results are consistent with the hypothesis that when speech becomes faster, vowels became more centralized.

Frontness

Significant differences were obtained for the tense vowel (careful-fast), but not for the lax vowel. Partial eta squared showed a large effect size and no interaction was found between speech style and native language (see appendix 4). Further t-tests indicated significant differences over all speech styles for the tense vowel for the non-native speakers, but not for the native speakers (appendix 5). The results of the non-natives are consistent with the hypothesis of centralization.

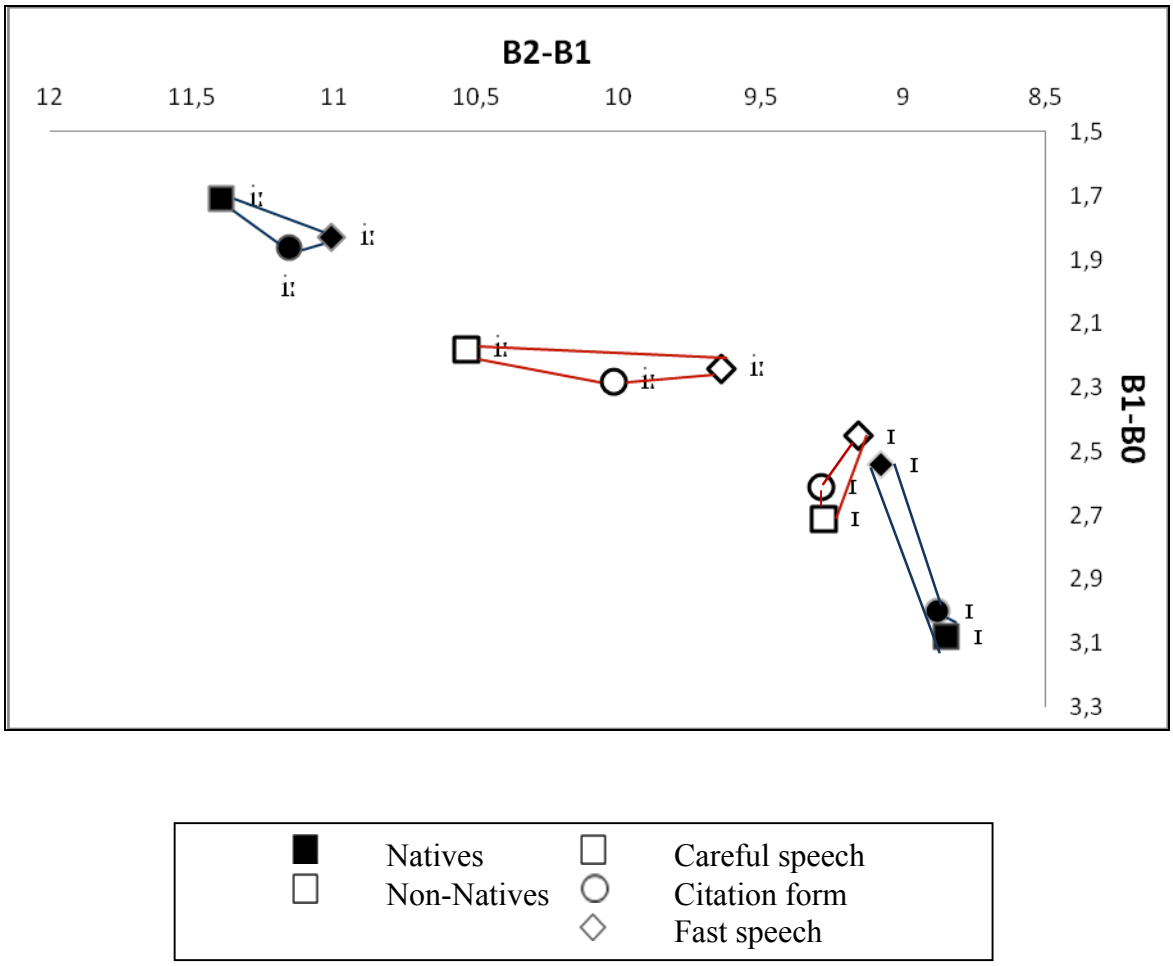

Figure 5. Effects of speech style on vowel height and frontness. The shape of the tense and lax vowel categories is indicated by connected lines.

Spectral distance

Significant differences were found for careful-citation and careful-fast. Partial eta squared indicated a large effect size and there was no speech style*native language interaction. Further t-tests showed that the difference between careful-fast speech was significant for natives and non-natives, but the difference between citation and fast speech was only significant for non-natives. This indicated that both L1 groups, produced /iː/ and /ɪ/ as more distinct in careful speech and that in fast speech, the distinction between the vowels became smaller (see figure 6)

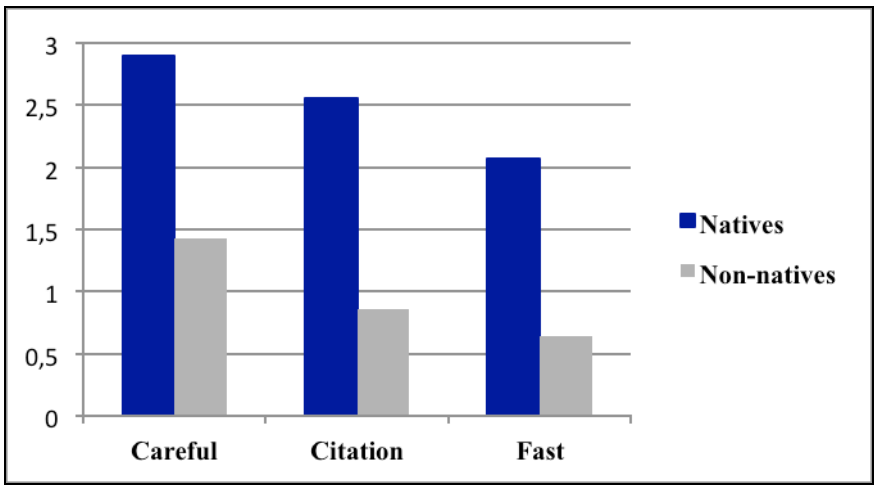

Figure 6. Spectral Distance over speech styles.

Overall the speech style effect hypothesis was supported. As expected, the dimension that was most affected by the speech style was duration. In the spectral dimension (see figures 5 and 6), the two vowels behaved differently. The tense vowel was more affected in vowel frontness, whereas the lax vowel was more affected in vowel height. This is not unexpected if we take into account their positions in the vowel space; because the tense vowel is already high, it was more affected for the back-front dimension and because the lax vowel is already centralized, it was more affected in height.

Although not reaching significance in all occasions, a trend consistent with the initial hypotheses could be seen for speech style: when speed increased, the vowels became more centralized and when speed became slower, the vowels became more dispersed and realized with bigger spectral differences. However, the non-natives and the natives were found to behave differently. Overall, the non-natives supported the initial hypothesis more strongly, whereas for the native speakers, the effects of the speech style on the two vowels were mostly mild. This supports the idea that whereas the native speakers' vowels were stable and fully developed and consequently are not strongly affected by speech style changes, the non-native speakers' developing L2 categories were not robust enough to endure changes in speech style.

4. Is there a relation between perception and production? If so, what happens to this relation when speech style is modified?

The relation between perception and production was explored with Spearman's rank order correlation (see appendix 7).

The native speakers overall did not show correlations between perception scores and production measures, suggesting that perception and production were not correlated in native speakers. This finding is not surprising if we take into account the small variation the native English speakers showed in the perception and production tasks.

The non-native speakers, on the other hand, showed correlations for the spectral dimension and perception scores. Spectral distance in careful and fast speech showed large positive correlations with both perception measures. This indicated that the better the Spanish/Catalan speakers were able to identify the tense and the lax vowel in the perception, the bigger the spectral distances they created for them in careful and fast speech. This finding provides moderate support to the idea that second language speech perception and production are related.

The relation between perception and production was also affected by the speech style, indicated by the fact that citation productions did not correlate with perception, but careful and fast did. Following the initial idea about L2 phonetic categories being robust enough to stand speech style changes if they had in fact been created, it could be hypothesized that the participants' productions in citation form did not indicate category formation but conscious attention (imitation) to the model. When the tasks were made more challenging by fast and careful speech, the participants who had in fact created L2 categories were still able to make a distinction between /iː/ and /ɪ/, whereas the participants who had not formed L2 categories were obliged to realize the two vowels closer to each other. The correlation of perception and fast and careful speech would indicate that the participants who were more native-like (relying in spectral differences) in perception, were also the ones who had the strongest L2 categories in production, persisting changes in speech style.

6. DISCUSSION AND CONCLUSIONS

This paper investigated the perception and production of English /iː/-/ɪ/ by Spanish/Catalan EFL learners. The perception task confirmed reliance on temporal cues over spectral cues, indicating that the non-native participants had not reached a native-like perception (Escudero, 2000; Morrison, 2008 & 2009).

In the production tasks, the non-native participants' performance indicated that category assimilation to the native /i/ had not taken place and that two distinct vowel sounds were used for English. These findings are quite surprising in light of previous research showing that the /iː/ - /ɪ/ contrast is difficult for Spanish/Catalan learners. An explanation for the obtained results could be that the elicitation method used in this study was designed to elicit the best possible target vowels from the participants by giving a native model. Had a spontaneous sentence reading task been used, for example, the realization of the target vowels might be much poorer. It should also be taken into account that the non-native vowels produced by the Spanish/Catalan speakers were significantly different to the /iː/ - /ɪ/ produced by the native English speakers, which is why caution should be exercised in interpreting the results. Although a significant difference was found for the two target vowels, it does not necessarily mean that this difference is actually discernable in speech. A more complete picture could be obtained by using both objective and subjective vowel measurements. For example, a forced choice vowel identification and goodness rating task could be created with the non-native speakers' productions and presented to a panel of native raters. This would show if the significant difference obtained through acoustic analysis is in fact audible to native speakers.

The spectral values obtained for the L1 vowel placed the Catalan/Spanish vowel in a position that is not in line with previous research which states that the Spanish /i/ is less fronted and lower than the English /iː/. There are some possible explanations for these results. The data elicitation method differed for the L1 vowel and the English vowels. It is possible, that reading words in isolation had an impact on the quality of the L1 vowel. In addition, the participants in this study were Catalan/Spanish bilinguals and not Spanish monolinguals as in previous studies. Since the main interest was to determine whether the

non-native participants realized their L2 vowels differently to their L1 vowel, the Catalan and Spanish dominant bilinguals as well as Spanish monolinguals were treated as one group. Although the Catalan and the Spanish /i/ have been reported to have similar spectral values, it is possible that treating them as one is not appropriate. Finally, the study only used 20 Spanish/Catalan bilinguals and 7 English native speakers. It is possible that with a higher sample different results would have been obtained.

Support was found for the existence of a relationship between second language speech perception and production. The fact that the correlations were significant for some production and perception measures, but not for others, would indicate that the relationship between perception and production in second language speech learning is complex and requires further research.

The production tasks also revealed differences in the vowels over speech styles. In general, the native speakers' vowels were not spectrally highly affected by speech style, but the Spanish/Catalan bilinguals showed significant spectral differences. This could be interpreted as supporting the idea that category formation had not actually taken place since speed changes had a strong influence on the vowels, especially on /iː/. However, if compared to native speakers (see figures 5 and 6), it can be seen that both groups follow the same trend: vowels produced in careful style being in the most extreme positions and vowels produced in the fast speech being in the most centralized positions.

As results of this and previous studies suggest, following Flege's (1995) prediction that accuracy in second language speech production is dependent on accuracy of second language speech perception, L2 speech perception and production should be taken into account when designing applications for L2 phonetic training. Speech style changes in perception were not studied in this paper. However, future research should consider doing this in order to see how foreign language learners' accuracy in perception is affected by speech style. The preliminary results obtained in this study suggest that non-native speakers' interlanguage speech sounds are malleable. It is premature to predict whether teaching careful, citation or fast speech would provide more native-like results in foreign language learners, but it seems that increasing variability to L2 pronunciation teaching by means of different voices, accents and speech styles can only be favorable by reflecting the reality of the world outside the classroom.

Acknowledgements

I want to start by thanking my supervisor Dr. Joan Carles Mora for sharing his expertise with me, for spending so many hours solving my doubts and for offering invaluable support and ideas at every phase of the project. My wholehearted thank you.

Secondly, I want to express my gratitude for the people who volunteered for the experiments. Without their participation this Master's thesis would not exist. I especially wish to thank the native English speakers and the fellow MA students who took time off of their busy schedules to help and encourage me.

Finally, I want to thank my family and friends for all the support I have received during these months. Susanna, kiitos lukemattomista tsemppaussessioista. Mumski, kiitos kannustuksesta. Diogo, obrigada por acreditar em mim. Sem a sua paciência e o seu apoio isto não tinha sido possível.

Thank you.

REFERENCES

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech, 51(4),* 317-342.

Best, C. (1995). A direct realist view of cross-language speech perception. In Strange, W. (ed), *Speech Perception and Linguistic Experience: Issues in Cross-language research.* Timonium, MD: York Press.

Best, C., Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementaries. In Bohn, O. and Munro, M. (Eds.), *Language Experience in Second Language Speech Learning In honor of James Emil Flege*. John Benjamins Publishing Company.

Boersma, P., Weenink, D. (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.26, retrieved 24 May 2011 from http://www.praat.org/

Bohn, O. (1995). Cross language speech perception in adults: First language transfer doesn't tell it all. In Strange, W. (ed)*, Speech Perception and Linguistic Experience: Issues in Cross-language research.* Timonium, MD: York Press, pp. 279-304

Bohn, O., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(3), 303-28.

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

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics 34 (2006*), 372-387.

Cebrian, J. (2007). Old sounds in new contrasts: L2 production of the English tense-lax vowel distinction. In Trouvain, J. & Barry, W. (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, vol. 3 (pp. 1637-1640), 6-10 August 2007, Saarbrücken: Universität des Saarlandes*.

Cerviño, E., Mora, J. C. (2009) When over-reliance on duration does not mean perceiving duration differences. In Ashby, M. and Maidment, J. (Eds.), *PTLC2009 Proceedings: Phonetics Teaching and Learning Conference*. (pp. 27-30) London: University College London.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. Journal *of the International Phonetic Association, 27*: 47-55.

Escudero, P. (2000). Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers. Unpublished master's thesis, University of Edinburgh, Scotland, UK.

Escudero, P. (2006). The phonological and phonetic development of new vowel contrasts in Spanish learners of English. In Baptista, B. and Watkins, M. (Eds.), *English with a Latin Beat. Studies in Portuguese/Spanish- English Interphonology*. John Benjamins Publishing Company.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26(4)*, 551-585.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America, 93*(3), 1589-1608.

Flege, J. E. (1995) Second-language Speech Learning: Theory, Findings, and Problems. In Strange, W. (ed), *Speech Perception and Linguistic Experience: Issues in Cross-language research.* Timonium, MD: York Press, Pp. 229-273

Flege, J.E. (2003). Methods for assessing the perception of vowels in a second language. In Fava, E. and Mioni, A. (Eds.), *Issues in Clinical Linguistics*. Padova: UniPress, Pp. 19-44.

Flege, J. E., Bohn, O., Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J. E., MacKay, I. R. A., Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106*(5), 2973-2987.

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

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America, 97*(4), 2540-2551.

Francis, A., Nusbaum, H. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 2002, Vol. 28. No. 2, 349-366.

Frieda, E. , Nozawa, T. (2007). You are what you eat phonetically. The effect of linguistic experience on the perception of foreign vowels. In Bohn, O. and Munro, M. (Eds.), *Language Experience in Second Language Speech Learning In honor of James Emil Flege*. John Benjamins Publishing Company.

Frieda, E., Walley, A., Flege, J. E., Sloane, M. (2000). Adult's perception and production of the vowel /i/. *Journal of Speech, Language and Hearing Research*. Vol. 43. February 2000. 129-143.

Gilabert, R., Mora, J. C., Muñoz, C. (2010). Phonological short-term memory and working memory capacity in foreign language acquisition. Paper presented at SLRF 2010 (Second Language Research Forum. Invited colloquium on Aptitude and Individual Differences in SLA), University of Maryland, Washington, USA.

Johnson, K. (2000). Adaptive dispersion in vowel perception. *Phonetica 57 (2-4)*,181-188.

Johnson, K., Flemming, E., Wright, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language, 69*(3), 505-28.

Koerich, R. (2006). Perception and production of a vowel paragoge by Brazilian EFL students. In Baptista, B. and Watkins, M. (Eds.), *English with a Latin Beat. Studies in Portuguese/Spanish-English Interphonology*. John Benjamins Publishing Company.

Kuhl, P., (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843

Kuhl, P., Conboy, B., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., Nelson, T. (2007). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society. B (2008) 363*, 979-1000.

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

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. *Proceedings of the 13th International Congress of Phonetic Sciences. Vol* 4.(pp. 92-99). Stockholm, Sweden, August 13-19, 1995.

MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America, 110*(1), 516-528.

Miller, J., Volaitis, L. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*. 1989, 46 (6). 505-512.

Mora, J. C. (2005). Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and non-words by Spanish/Catalan bilingual learners of English. *Proceedings of the ISCA Workshop on Plasticity in Speech Perception (PSP 2005); London, UK; 15-17th June 2005.*

Mora, J.C., Cerviño-Povedano, E. (2010) Phonological short-term memory and L2 vowel perception. Paper presented at the 20th Annual Conference of the European Second Language Association (EUROSLA-20), Reggio Emilia, Italy. September 1-4 2010.

Mora, J. C., Fullana, N. (2007). Production and perception of English /i:/-/ ɪ / and /æ/- /ʌ/ in a formal setting: Investigating the effects of experience and starting age. In Trouvain, J. & Barry, W. (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, vol. 3 (pp. 1613-1616), 6-10 August 2007, Saarbrücken: Universität des Saarlandes*.

Morrison, G. S. (2008). L1-Spanish speakers' acquisition of the English /i /-/I/ contrast: Duration-based perception is not the initial developmental stage. *Language and Speech, 51*(4), 285-315.

Morrison, G. S. (2009). L1-Spanish speakers' acquisition of the English /i/-/I/ contrast II: Perception of vowel inherent spectral change. *Language and Speech, 52*(4), 437-462.

Moya-Galé, G. (2010). Overcoming non-native over-reliance: a study on English vowel duration manipulation and neutralization. Unpublished MA Thesis. Universitat de Barcelona.

Peperkamp, S., Bouchon, C. (in press). The relation between perception and production in L2 phonological processing. In: *Proceedings of Interspeech 2011.*

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

Syrdal, A., Gopal, H. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America. 79*, pp. 1086-1100.

Traunmüller, H. (1997). Auditory scales of frequency representation. [online resource]. Retrieved 30th May 2011 from http://www.ling.su.se/staff/hartmut/bark.htm

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., Flege, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics 33 (2005*), pp. 263-290.

Volaitis, L., Miller, J. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *The Journal of the Acoustical Society of America*. 92 (2)

Ylinen, S., Uther, M., Latvala, A., Vepsälainen, S., Iverson, P., Akahane-Yamada, R., Näätänen, R. (2009). Training the brain to weight speech cues differently: A study of

Finnish second-language users of English. *Journal of Cognitive Neuroscience*. 22:6, pp. 1319-1332.

APPENDIX

Appendix 1.

Means and standard deviations (in parenthesis) for the British and American native speakers' perception and production (in citation) measures and the results of the Mann-Whitney U test.

		British	American	Mann-Whitney U test
Perception (% of ID)	**Natural**	100 (0)	98.61 (1.96)	*z*=-1.323 *p*= .186
	Manipulated	98.14 (1.60)	97.91 (2.65)	*z*= .000 *p*= 1,00
Duration (ms)	**Tense**	199.84 (19.28)	197.25 (27.17)	*z*= -.354 *p*= .724
	Lax	127.17 (19.28)	152.28 (18.17)	*z*=-1.768 *p*= .077
B1-B0	**Tense**	1.99 (0.10)	1.77 (0.31)	*z*=-1.411 *p*= .157
	Lax	2.93 (0.29)	3.05 (0.36)	*z*=-.707 *p*= .629
B2-B1	**Tense**	11.22 (0.50)	11.12 (0.56)	*z*=-.354 *p*= .724
	Lax	8.98 (0.71)	8.80 (0.36)	*z*= -.707 *p*= .480
Spectral distance	**Tense-lax**	2.43 (0.97)	2.65 (0.89)	*z*=.000 *p*= 1.00

Appendix 2.

Words used in the English DSR tasks and in the Catalan/Spanish wordlist reading.

English minimal pairs

Bead	bid
Beat	bit
Deed	did
He'd	hid
Heat	hit
Keys	kiss
Peak	pick
Pete	pit
Seat	sit
Seed	Sid
Seen	sin
Team	Tim

Spanish words	Catalan words
Dida	Dida
Ida	Ida
Jida	Shida
Pica	Pica
Pita	Pita
Quiso	Kissa
Sida	Sida
Sina	Sina
Tima	Tima
Vida	Vida

Appendix 3.

Speed of delivery measures for the English native speaker who produced the words in the DSR tasks and for the Spanish/Catalan participants.

	Speed of delivery (segments/seconds)			ANOVA
	Careful (m)	Citation (m)	Fast (m)	
DSR model	**6.15**	**9.06**	**16.34**	$F(2.22)=1168.62$; $p<.001$; $\eta^2=.991$
Non-native participants	6.82	10.62	16.41	$F(2.18)=279.18$; $p<.001$; $\eta^2= .969$
p01	6.74	11.65	19.37	$F(2.22)=443.15$; $p<.001$; $\eta^2 =.976$
p03	7.14	9.23	14.95	$F(2.22)=322.72$; $p<.001$ $\eta^2= .967$
p04	6.80	10.33	15.10	$F(2.22)=859.52$; $p<.001$; $\eta^2=.987$
p06	6.42	9.68	15.61	$F(2.22)=1478.25$; $p<.001$; $\eta^2=.993$
p07	6.62	9.85	13.57	$F(2.22)=794.15$; $p<.001$; $\eta^2=.986$
p08	6.91	10.24	18.10	$F(2.22)=1191.20$; $p<.001$; $\eta^2=.991$
p11	7.01	10.46	17.08	$F(2.22)=862.19$; $p<.001$; $\eta^2=.987$
p12	6.55	11.36	17.49	$F(2.22)=458.05$; $p<.001$; $\eta^2=.977$
p13	6.86	10.91	13.81	$F(2.22)=254.53$; $p<.001$; $\eta^2=.959$
p15	7.76	13.31	18.78	$F(2.22)=2082.30$; $p<.001$; $\eta^2=.995$
p16	7.27	11.29	16.20	$F(2.22)=985.95$; $p<.001$; $\eta^2=.989$
p17	6.48	11.68	17.64	$F(2.22)=2624.16$; $p<.001$; $\eta^2=.996$
p21	6.35	11.09	16.71	$F(2.22)=392.38$; $p<.001$; $\eta^2=.973$
p22	6.64	10.72	17.80	$F(2.22)=565.95$; $p<.001$; $\eta^2=.981$
p23	7.74	13.15	18.69	$F(2.22)=616.56$; $p<.001$; $\eta^2=.982$
p24	6.23	8.94	15.16	$F(2.22)=943.79$; $p<.001$; $\eta^2=.988$
p25	6.37	9.80	18.05	$F(2.22)=2118.51$; $p<.001$; $\eta^2=.995$
p26	7.32	9.90	13.76	$F(2.22)=504.50$; $p<.001$; $\eta^2=.979$
p28	6.83	10.09	14.99	$F(2.22)=540.80$; $p<.001$; $\eta^2=.980$
p29	6.41	8.76	15.46	$F(2.22)=1640.17$; $p<.001$; $\eta^2=.993$

Appendix 4.

Results of paired samples t-tests and Wilcoxon Signed Rank tests for the non-native speakers' L1-English vowels and English tense-lax vowels. English vowels in citation form. (significance at $p<.05$ level is marked with asterisk).

		Non-natives	
		All (n=20)	**Low (n=10)**
Duration	**L1 x tense**	$t(19)=-15.13,\ p<.001$*	$z= -2.80,\ p=.005$*
	L1 x lax	$t(19)=-5.99,\ p<.001$*	$z= -2.70,\ p=.007$*
	Tense-lax	$t(19)=10.03,\ p<.001$*	$z= -2.280,\ p=.005$*
B1-B0	**L1 x tense**	$t(19)=-7.51,\ p<.001$*	$z= -2.80,\ p= .005$*
	L1 x lax	$t(19)=-12.66,\ p<.001$*	$z= -2.80,\ p= .005$*
	Tense x lax	$t(19)=-4.42,\ p<.001$*	$z= -2.59, p= .009$*
B2-B1	**L1 x tense**	$t(19)=5.23,\ p< .001$*	$z= -2.80,\ p=.005$*
	L1 x lax	$t(19)=8.04,\ p<.001$*	$z= -2.70,\ p=.007$*
	Tense x lax	$t(19)=5.23,\ p< .001$*	$z= -2.39 , p= .017$*

Appendix 5.

Results of Mixed between-within Anova for speech style and native language. (significance at $p<.05$ level is marked with asterisk).

		Pairwise comparisons		Multivariate tests
Duration	**Tense**	careful-citation	***p*<.001***	$F(2,24)=145.61$; ***p*<.001*** ; $\eta^2=.924$
		citation-fast	***p*<.001***	
		careful-fast	***p*<.001***	
	Lax	careful-citation	*p*=.377	$F(2,24)=82.30$; ***p*<.001***; $\eta^2=.873$
		citation-fast	***p*<.001***	
		careful-fast	***p*<.001***	
B1-B0	**Tense**	careful-citation	*p*=.395	$F(2,24)=1.16$; $p=.329$; $\eta^2=.089$
		citation-fast	*p*=1.00	
		careful-fast	*p*=1.00	
	Lax	careful-citation	*p*=1.00	$F(2,24)=8.96$; ***p*=.001***; $\eta^2=.428$
		citation-fast	***p*=.002***	
		careful-fast	***p*=.003***	
B2-B1	**Tense**	careful-citation	*p*=.079	$F(2,24)=6.87$; ***p*= .004*** ;$\eta^2=.364$
		citation-fast	*p*=.149	
		careful-fast	***p*=.003***	
	Lax	careful-citation	*p*=1.00	$F(2,24)=.123$; $p=.885$; $\eta^2=.010$
		citation-fast	*p*=1.00	
		careful-fast	*p*=1.00	
Spectral distance tense-lax		careful-citation	***p*=.050***	$F(2,24)=14.03$; ***p*<.001*** ; $\eta^2=.539$
		citation-fast	*p*=.086	
		careful-fast	***p*<.001***	

Appendix 6.

Results of paired samples t-tests for spectral dimensions. Significance at $p<.05$ level is marked with asterisk.

			Non-natives	Natives
B1-B0	**Tense**	careful-citation	*t*(19)= 1.272 ; *p*=.219	*t*(6)= .956 ; *p*=.376
		citation-fast	*t*(19)= .499 ; *p*=.624	*t*(6)= .338 ; *p*=.747
		careful-fast	*t*(19)=-.686 ; *p*=.501	*t*(6)=-.602 ; *p*=.569
	Lax	careful-citation	*t*(19)=-1.109 ; *p*=.281	*t*(6)=-.431 ; *p*=.681
		citation-fast	*t*(19)= 1.920 ; *p*=.070	*t*(6)= 3.350 ; ***p*=.015***
		careful-fast	*t*(19)= 2.310 ; ***p*=.032***	*t*(6)= 3.259 ; ***p*=.017***
B2-B1	**Tense**	careful-citation	*t*(19)= -3.057 ; ***p*=.006***	*t*(6)=-1.001 ; *p*=.355
		citation-fast	*t*(19)= 2.621 ; ***p*=.017***	*t*(6)= 1.346 ; *p*=.227
		careful-fast	*t*(19)= 4.710 ; ***p*<.001***	*t*(6)= 2.267 ; *p*=.064
	Lax	careful-citation	*t*(19)= .085 ; *p*=.933	*t*(6)=.277 ; *p*=.791
		citation-fast	*t*(19)= .949 ; *p*=.355	*t*(6)=-.833 ; *p*= .437
		careful-fast	*t*(19)=-.908 ; *p*=.375	*t*(6)= 1.489 ; *p*=.187
Spectral distance tense-lax		careful-citation	*t*(19)=-3.185; ***p*=.005***	*t*(6)=-1.070 ; *p*=.326
		citation-fast	*t*(19)=1.548; *p*=.138	*t*(6)= 1.525 ; *p*=.178
		careful-fast	*t*(19)=5.006; ***p*<.001***	*t*(6)= 3.453 ; ***p*=.014***

Appendix 7.

Spearman correlations for perception and production data. Significance at $p<.05$ level is marked with asterisk.

			Non-natives (n=20)		Natives (n=7)	
			% ID natural	**% ID manipulated**	**% ID natural**	**% ID manipulated**
Duration	**Tense**	Careful	r_s= -.044 ; p=.853	r_s = -.047 ; p=.845	r_s=.535 ; p=.216	r_s= -.056 ; p=.905
		Citation	r_s= -.226 ; p=.338	r_s= -.206 ; p=.384	r_s=.178 ; p=.702	r_s=.131 ; p=.780
		Fast	r_s=.038 ; p=.872	r_s= -.008 ; p=.972	r_s =.089 ; p=.849	r_s= -.037 ; p=.937
	Lax	Careful	r_s= -.198 ; p=.403	r_s= -.235 ; p=.318	r_s=-.356 ; p=.433	r_s= -.075 ; p=.873
		Citation	r_s= -.431 ; p=.058	r_s= -.309 ; p=.185	r_s= -.223 ; p=.631	r_s= -.168 ; p=.718
		Fast	r_s= -.087 ; p=.714	r_s= -.230 ; p=.330	r_s= -.089 ; p=.849	r_s=.075 ; p=.873
B1-B0	**Tense**	Careful	**r_s= -.464 ; p=.039***	**r_s= -.536 ; p=.015***	**r_s= -.802 ; p=.030***	r_s= -.187 ; p=.688
		Citation	r_s= -.050 ; p=.835	r_s= -.021 ; p=.930	r_s= -.134 ; p=.775	r_s=.056 ; p=.905
		Fast	r_s= -.154 ; p=.518	r_s=.004 ; p=.987	r_s= -.045 ; p=.924	r_s=.037 ; p=.937
	Lax	Careful	r_s=.216 , p=.360	r_s=.107 ; p=.653	r_s= -.668 ; p=.101	r_s= -.225 ; p=.628
		Citation	r_s=.002 ; p=.992	r_s=.236 ; p=.317	**r_s= -.802 ; p=.030***	r_s= -.468 ; p=.290
		Fast	r_s=.048 ; p=.840	r_s=.204 ; p=.388	r_s= -.579 ; p=.173	r_s=.056 ; p=.905
B2-B1	**Tense**	Careful	r_s=.607 ; p=.005*	**r_s=.510 ; p=.022***	r_s=.045 ; p=.924	r_s= -.561 ; p=.190
		Citation	r_s=.049 ; p=.838	r_s= -.076 ; p=.750	r_s=.134 ; p=.775	r_s= -.112 ; p=.811
		Fast	r_s=.315 ; p=.176	r_s=.214 ; p=.365	r_s= -.223; p=.631	r_s= -.243 ; p=.599
	Lax	Careful	r_s= -.339 ; p=.143	r_s= -.375 ; p=.103	r_s=.668 ; p=.101	r_s=.206 ; p=.658
		Citation	r_s= -.006 ; p=.980	r_s= -.305 ; p=.191	r_s=.490 ; p=.264	r_s=.112 ; p=.811
		Fast	r_s= -.079 ; p=.740	r_s= -.155 ; p=.513	r_s =.401 ; p=.373	r_s= -.449 ; p=.312
Spectral distance tense-lax		Careful	**r_s=.631 ; p=.003***	**r_s=.554 ; p=0.11***	r_s= -.401 ; p=.373	r_s= -.356 ; p=.434
		Citation	r_s=.194 ; p=.412	r_s=.413 ; p=.070	r_s= -.178 ; p=.702	r_s= -.037 ; p=.937
		Fast	**r_s=.599 ; p=.005***	**r_s=.566 ; p=.009***	r_s = -490 ; p=.264	r_s=.019 ; p=.968