Anais do X Encontro do CELSUL – Círculo de Estudos Linguísticos do Sul
UNIOESTE - Universidade Estadual do Oeste do Paraná
Cascavel-PR | 24 a 26 de outubro de 2012 | ISSN 2178-7751

SPEECH RATE EFFECTS ON L2 VOWEL PRODUCTION AND PERCEPTION

Hanna KIVISTÖ DE SOUZA[1]
Joan Carles MORA[2]

*RESUMO: Este estudo investigou o impacto da velocidade da fala (VF) na precisão da percepção e da produção de vogais em uma língua estrangeira (L2). Processos de comunicação são caracterizados pela variação da velocidade de fala e portanto, torna-se relevante saber como os aprendizes de uma língua estrangeira atuam em relação a esta diversidade. No Experimento I, 20 estudantes com um nível avançado de inglês (L1=catalão/espanhol) e um grupo de controle de falantes nativos de inglês (n=7) produziram as vogais inglesas contrastantes /iː/-/ɪ/ em 12 pares mínimos. Estes foram obtidos através de frases em uma tarefa de repetição retardada em três VF (claro, citação e rápido). Os resultados obtidos através de uma ANOVA mostraram um efeito principal significativo de VF (p< .001) nas dimensões temporais e espectrais das duas vogais, sendo que as vogais tornaram-se mais curtas e centralizadas com o aumento da velocidade articulatória. No Experimento II, 54 estudantes com nível avançado de inglês participaram de um teste de identificação contendo as vogais /iː/-/ɪ/ em palavras obtidas pelos falantes nativos do Experimento I nas mesmas três VF. ANOVA revelou um efeito principal significativo para a VF (F(2,52)=30,00; p<.001; η^2=.54) e qualidade de vogal (F(1,53)=33.94;p<.001; η^2=.39), assim como uma interação significativa VF*qualidade vogal (F(2,52)=97.39; p<.001; η^2=.79), sugerindo que a VF afetou a precisão da identificação das duas vogais de uma maneira diferente. No geral, os resultados indicam que as mudanças na VF afetam a percepção e a produção das vogais em uma L2, sugerindo que a VF deveria ser levada em conta na elaboração de ferramentas de treino fonético de uma L2.*
PALAVRAS-CHAVE: Velocidade de fala; Produção de vogais em L2; Percepção de vogais em L2.

*ABSTRACT: This study investigated whether changes in speech rate (SR) affect the accuracy of production and perception of non-native (L2) vowels. In daily interaction, speech is delivered at different rates. Knowing how L2 users behave when faced with such diversity might be beneficial for L2 speech training. In Experiment I, a group of Catalan/Spanish advanced EFL learners (NNSs, n=20) and a native-speaker control group (NSs, n=7) produced tokens of the contrasting English vowels /iː/ and /ɪ/ embedded in 12 minimal pairs. These were elicited through a delayed sentence repetition task at three SRs (labeled as clear, citation and fast). Mixed between-within ANOVAs showed a significant main effect of SR (p<.001) on the temporal and spectral dimensions of the two vowels, so that the vowels became increasingly shorter and centralized as articulatory speed increased. In Experiment II, a group of NNSs (n=54) performed a forced-choice word identification task for the /iː/-/ɪ/ vowel pair in the words elicited from the NSs in experiment I at the same three SRs. An ANOVA yielded significant main effects of SR (F(2,52)=30.00; p<.001;η^2=.54) and vowel quality (F(1,53)=33.94;p<.001; η^2=.39), as well as a significant SR*vowel quality interaction (F(2,52)=97.39;p<.001;η^2=.79) indicating that SR affected NNSs' identification accuracy for the two vowels differently. Overall, these findings indicate that changes in SR affect L2 vowel*

[1] PhD candidate, Universitat de Barcelona, hanna.kivistodesouza@gmail.com
[2] Lecturer, Universitat de Barcelona, mora@ub.edu

production and perception, suggesting that SR should be taken into account in the design of L2 phonetic training protocols.
KEYWORDS: Speech rate; L2 vowel production; L2 vowel perception.

1 Introduction

Daily communication is characterized by variability: we listen to males, females, children, native speakers, foreigners, different dialects, loud and soft voices, fast and slow talkers. Native speakers are naturally sensitive to this type of contextual variation and it does not hinder comprehension. Foreign language learners, on the other hand, require additional attentional resources in order to adapt to contextual variation. Every language learner knows how difficult it is to understand the foreign language in noisy environments or when spoken with an unfamiliar accent. Thus, in order to achieve a native-like competence in perception as well as production, the foreign language learner has to overcome the difficulties posed by contextual variation. However, most of the studies in the second language acquisition field have focused only on eliciting or presenting speech at one rate, usually at the normal reading rate (citation), thus studies testing L2 users at varying speech rates are needed. The aim of this study was to see how foreign language learners adapt to speech rate variation in production and perception.

2 Speech rate effects on production and perception

According to Lindblom's (1990) hyper-hypo theory, speakers modify their articulation effort according to the listener's needs. Thus, in noisy conditions or when speaking to a person with a hearing impediment, for example, speakers tend to hyperarticulate in order to make the speech sounds as distinct from each other as possible. Among the strategies used by English speakers in adverse listening conditions are: decreasing the speech rate, increasing the duration and frequency of pauses, increasing pitch range and the mean fundamental frequency, releasing word final consonants, increasing the Voice Onset Time (VOT) distinction between voiced and voiceless plosives as well as the spectral distance between front and back vowels (HAZAN; BAKER, 2011; GRANDLUND; HAZAN; BAKER, a ser publicado). On the other hand, in normal communicative encounters, the speaker makes less articulatory effort and the speech sounds may not reach their target forms (resulting in undershoot) since comprehension is not threatened. A similar theory was put forward by Johnson (JOHNSON; FLEMMING; WRIGHT, 1993; JOHNSON, 2000), who claims that clear speech is hyperarticulated, i.e., the phonemes are dispersed in relation to the realizations in normal speech.

Previous studies have confirmed that speech rate affects the temporal dimension, so that as the speed increases, the segments become shorter. The phenomenon has been observed in VOT (GRANDLUND; HAZAN; BAKER, a ser publicado; RALLO FABRA, 1997; SCHMIDT; FLEGE, 1995; 1996; VOLAITIS; MILLER, 1992), syllable duration (SCHMIDT; FLEGE, 1995), vowel duration (SCHMIDT; FLEGE, 1995; DETERDING, 1997; MOON; LINDBLOM, 1989) and phrase duration (SCHMIDT; FLEGE, 1995). Native speakers have been shown to perform consistently, whereas non-native speakers have shown variation depending on their proficiency level (VOT in SCHMIDT; FLEGE, 1995; 1996).

Previous studies have also confirmed that the spectral dimension in vowels is affected by speech rate. Native speakers have been shown to produce more dispersed vowels in slow speech (expanded vowel space) and more centralized vowels in faster speech (JOHNSON;

FLEMMING; WRIGHT, 1993; JOHNSON, 2000; FRIEDA *et al*., 2000, MOON; LINDBLOM, 1989). In perception, several studies have established that native speakers and advanced foreign language learners adapt to the category boundary shifts caused by changes in the speech rate (VOT: FLEGE; SCHMIDT, 1995; MILLER; VOLAITIS, 1989; VOLAITIS; MILLER, 1992). However, to our knowledge, no previous studies have tried to determine how non-native speakers react to speech rate variation in the production and perception of non-native vowels.

3 The present study

The present study investigates how the perception and production of L2 vowels is affected by three speech rates, established by means of temporal properties. They were labeled as: *clear*, *citation* and *fast*. *Clear* speech was defined as slow-paced and hyperarticulated (speed of delivery [segments/seconds]: 6.15). *Citation* speech (speed of delivery: 9.06) was defined as the "normal" speech rate. *Fast* speech (speed of delivery: 16.41) was understood as a more casual speech style present in everyday conversations.

We chose to concentrate on the English /iː/-/ɪ/ vowel pair because implementing the contrast has been shown to be challenging for Spanish/Catalan learners of English. Spanish and Catalan have only one high front vowel /i/ occupying roughly the same portion of the vowel space as the English /iː/-/ɪ/ (FLEGE, 2003; MORA, 2005). In addition, Spanish and Catalan do not use duration contrastively, whereas English uses duration as a secondary cue to discern between this vowel pair. Previous research has suggested that speakers of languages with less crowded vowel spaces, such as Spanish (5) and Catalan (7), will have more difficulties in establishing the smaller spectral differences present in languages with more crowed vowel spaces (British English 12) (FOX; FLEGE; MUNRO, 1995; FRIEDA; NOZAWA, 2007). As a consequence, Spanish and Catalan learners have been shown to assimilate the English /iː/-/ɪ/ vowels to the native /i/ and to distinguish them based on duration cues (long-short) instead of spectral cues.

We carried out two experiments to examine Spanish/Catalan advanced EFL learners' production and perception of the English tense and lax high front vowels. In Experiment I, we tested the production of the vowel pair in order to determine whether the foreign language learners had pronunciation difficulties in one or several speech styles. Following Schmidt and Flege's (1995; 1996) rationale, we expected speech rate to be an indication of whether category formation for the English /iː/-/ɪ/ had taken place. So that, 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, whereas accurate production in extremely slow and fast speech styles would indicate that robust phonetic categories had been formed. In addition, we expected to replicate the findings of previous research carried on with native speakers (JOHNSON; FLEMMING; WRIGHT, 1993; JOHNSON, 2000; FRIEDA *et al.*, 2000), showing that in clear speech, the vowels would have more extreme positions, whereas in fast speech overlapping articulation gestures would result in more centralized vowels. In Experiment II, we tested the perception of the English /iː/-/ɪ/ at different speech rates with a participant population analogous to the Spanish/Catalan EFL learner group who participated in Experiment I. We hypothesized that if the speakers in Experiment I were having difficulties with the *fast* or *clear* speech, this would also be likely to be reflected in difficulties in the perception of the contrast by the participants in Experiment II.

4 Experiment I

This experiment investigated whether changes in speech rate would have an effect on the accuracy of the production of non-native vowels. Following previous research conducted with native speakers, it was expected that *fast* speech would result in shorter and more centralized vowels due to the overlap of articulatory gestures. However, it was also hypothesized that the differences in speech rate would affect accuracy of production in the two non-native vowels to a greater extent in non-native speakers than in native speakers due to the varying degrees to which the former would have accomplished target-like category formation for these vowels in their interlanguage phonology.

4.1 Method

4.1.1 Participants

20 Catalan-Spanish advanced learners of English and a control group of 7 native English speakers participated in Experiment I.

Table 1. Characteristics of the participants. SD between parentheses

	Age	**Sex**	**L2 use**
Catalan/Spanish EFL leaners	22.85 (4.9)	12female 8male	17.72% (14.27)
Native English speakers	27.70 (4.6)	7female	29.29% (13.36)

4.1.2 Materials

A Delayed Sentence Repetition (DSR) task (FLEGE; MUNRO; MACKAY, 1995) was used to elicit the vowel productions. A DSR task minimizes direct imitation from sensory memory, list effects and speech dysfluencies. Orthographic effects on production were avoided by presenting the stimuli only aurally. Three DSR tasks were created, one for each speech style (*clear, citation* and *fast*). Each of the tasks had the same 12 /iː/-/ɪ/ minimal pairs (Table 2), but the order of the lexical items varied from task to task. 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).

Table 2. Minimal pairs used in Experiment I

/iː/	bead	beat	deed	heed	heat	keys	peak	Pete	seat	seed	seen	team
/ɪ/	bid	bit	did	hid	hit	kiss	pick	pit	sit	Sid	sin	Tim

A Southern British male speaker provided the model for the DSR tasks. The 12 minimal pairs x3 speech styles were normalized for peak and mean amplitude, filtered at 60Hz to eliminate any low-frequency noise and embedded into carrier sentences ("*beat* is the next word" -*citation*, "I didn't say *bit*, I said <u>*beat*</u>" - *clear*, and "I would say *beat* is the next word"- *fast*). *Clear* speech was established by means of an exaggerated contrast as in Frieda *et al*. (2000) and only the second token was analyzed. An articulation rate was established by counting the number of segments in the target sentence and by dividing it by the duration in seconds of the same sentence in order to confirm that the three speech styles differed in speed. A repeated measures ANOVA with articulation rate as independent and speech style as dependent variable confirmed that the three speech styles differed significantly ($F(2,22)=1168.62$; $p<.001$) in their speed.

4.1.3 Procedure

After filling in a language background questionnaire, the participants were recorded one by one in a soundproof booth in the phonetics laboratory in the University of Barcelona. The participant heard the mini-dialogues from the headphones and repeated the target sentence after the second repetition of the distractor sentence. The DSR tasks took between 3:14min (*fast*) and 4:50min (*clear*) to complete. The order of the production tasks was the same for all the participants: *citation, clear* and *fast*. The chosen order made the differences between the speech styles more pronounced by presenting the most dissimilar styles sequentially.

4.2 Results

The participants' productions were measured for duration and vowel quality (F0, F1, F2) with speech analysis software Praat (BOERSMA; WEENIK, 2011). The vowel quality measures were converted from frequencies (Hz) into a bark scale by using the following formula: B= 26.81 / (1+(1960/F)) – 0.53 (TRAUNMÜLLER 1997). The obtained Bark measures were then normalized for speaker characteristics such as gender and vocal tract size in order to make the production data comparable across speakers. The normalization 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; GOPAL, 1986). The normalized Bark values were then employed to calculate means for the vowel height and frontness for the two vowels in the three speech styles for every participant. Also, the mean Euclidean distance between the tense and the lax vowel in all the speech styles was computed.

Finally, 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 participant's speed of delivery measures in the three speech styles were submitted to a repeated measures ANOVA with Bonferroni adjustment. The results confirmed that all the participants realized three significantly different speech styles ($p<.001$).

In order to determine whether speech rate had an effect on the production of non-native vowels, mixed between-within ANOVAs with Bonferroni adjustment were conducted with *speech rate* (*clear/citation/fast*) as the within subjects factor and *L1* (*native/non-native*) as the between subjects factor. Descriptive statistics are presented in Table 3 and the results of the statistical analyses are presented in Table 4.

Table 3. Means (SD in parentheses) for vowel duration, height, frontness and spectral distance over speech styles for non-native and native speakers

		Non-natives			Natives		
		Clear (cl)	**Citation (ct)**	**Fast (fa)**	**cl**	**ct**	**fa**
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. Results of Mixed between-within ANOVA for speech style and native language (significance at p<.05 level is marked with an asterisk)

		ANOVAs		
		Multivariate tests	**Pairwise comparisons for speech style**	
Duration	**Tense**	$F(2,24)=145.61$; $p<.001$* ; $\eta^2=.924$	**cl-ct**	$p<.001$*
			ct-fa	$p<.001$*
			cl-fa	$p<.001$*
	Lax	$F(2,24)=82.30$; $p<.001$*****; $\eta^2=.873$ L1*SR interaction p<.010	**cl-ct**	$p=.377$
			ct-fa	$p<.001$*
			cl-fa	$p<.001$*
B1-B0	**Tense**	$F(2,24)=1.16$; $p=.329$; $\eta^2=.089$	**cl-ct**	$p=.395$
			ct-fa	$p=1.00$
			cl-fa	$p=1.00$
	Lax	$F(2,24)=8.96$; $p=.001$*; $\eta^2=.428$	**cl-ct**	$p=1.00$
			ct-fa	$p=.002$*
			cl-fa	$p=.003$*
B2-B1	**Tense**	$F(2,24)=6.87$; $p=.004$* ;$\eta^2=.364$	**cl-ct**	$p=.079$
			ct-fa	$p=.149$
			cl-fa	$p=.003$*
	Lax	$F(2,24)=.123$; $p=.885$; $\eta^2=.010$	**cl-ct**	$p=1.00$
			ct-fa	$p=1.00$

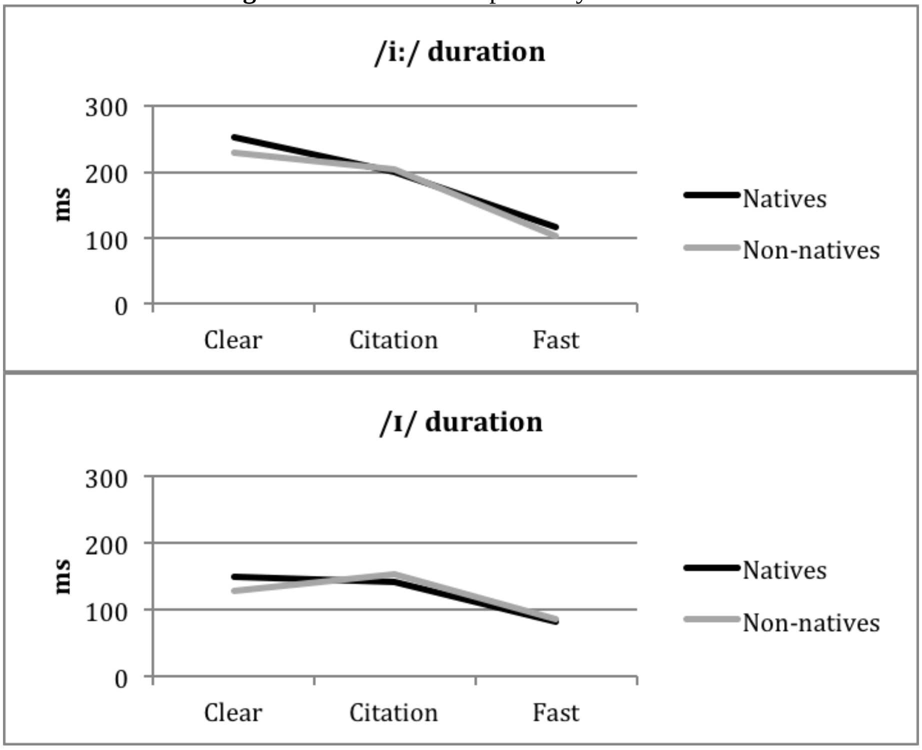
		cl-fa	*p*=1.00
Spectral distance tense-lax	*F*(2,24)=14.03; *p*<.001* ; η^2 =.539	**cl-ct**	*p*=.050*
		ct-fa	*p*=.086
		cl-fa	*p*<.001*

The results for the temporal and spectral dimensions will be discussed separately.

4.2.1 Temporal dimension

As expected, vowel duration was largely affected by speech rate. Both vowels became increasingly shorter as the speed increased and the differences were significant over all the speech rates. The duration difference was more pronounced in native speakers. There was a *L1*speech rate* interaction in the lax vowel, since the Spanish/Catalan participants produced longer lax vowels in the *citation* than in the *clear* speech, whereas the native English speakers' vowels were consistently shorter.

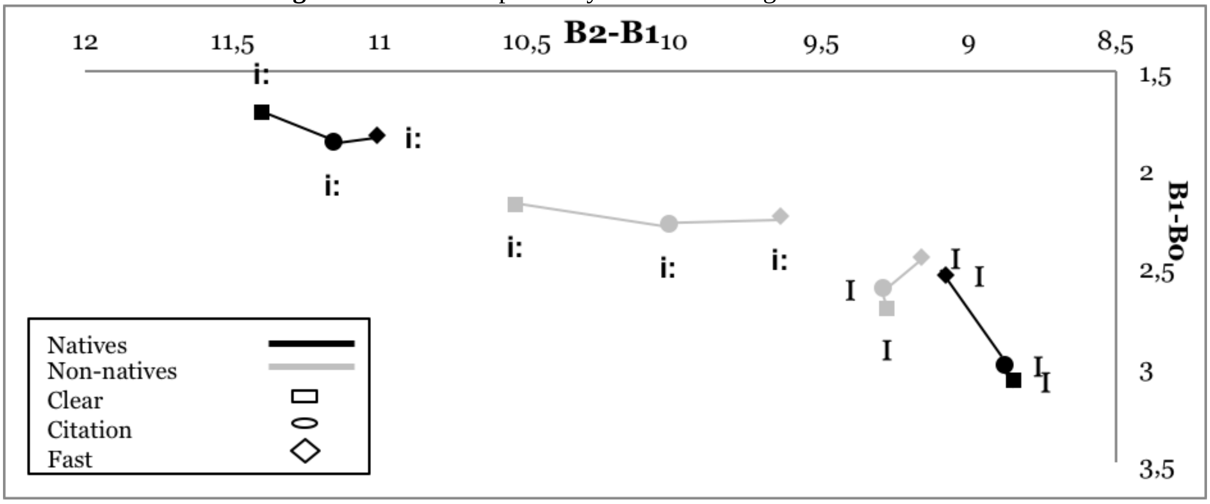
Figure 1. Effects of the speech style on vowel duration



4.2.2 Spectral dimension

As can be seen in Figure 2, the two vowels behaved differently. The lax vowel was significantly affected in height whereas the tense vowel was significantly affected in frontness. This is not surprising if we bear in mind their position in the vowel space; since the lax vowel is already more centralized, it seems logical that its height is more affected. Similarly, the tense vowel is already high, so it is not surprising that it was more affected in frontness. The spectral distance between the two vowels became significantly smaller when the speech became faster in both groups. However, it is important to notice that the native speakers consistently produced larger spectral distances between the vowels. These results support the initial hypothesis that in fast speech, the vowels became more centralized.

Figure 2. Effects of speech style on vowel height and frontness



4.3 Discussion

The results of Experiment I confirm that speech rate has an effect on the production of non-native as well as native vowels. Overall, the non-native vowels were more affected by the speech style, i.e. they became more dispersed than the native speaker vowels, suggesting that the vowel categories were not fully developed. It is also interesting to notice that the two vowels were differently affected by the speech rate. Both vowels became shorter when the speech style increased, although non-native speakers produced longer lax vowels in *citation* speech than they did in *clear* speech. In the spectral dimension, native speakers produced larger spectral distances between the vowels over all speech rates, showing the existence of two distinct vowel categories, whereas non-native speakers produced the vowels closer to each other. Both groups showed the same tendency in the spectral dimension, namely that the contrastive vowels occupied the most extreme positions, whereas the vowels spoken in the *fast* condition were the most centralized. This would suggest that fast speech inherently requires articulatory movements of a smaller magnitude and as a result the vowels become more centralized than in *clear* speech.

5 Experiment II

After establishing in Experiment I that speech rate affects non-native vowel production, we sought to determine whether similar effects are also found in non-native vowel perception. If production reflects what has been perceived (FLEGE, 1995), we would expect to find a similar effect of speech rate in vowel perception. Thus, we hypothesized that non-native speakers would be more accurate at identifying vowels in *clear* speech than in *fast* speech because in *clear* speech the vowels were more target-like whereas in *fast* speech they were more centralized and as a consequence, more difficult to identify for a speaker who does not have fully developed L2 vowel categories.

5.1 Method

5.1.1Participants

Fifty-four Catalan-Spanish advanced learners of English (45 female, 9 male) participated in this experiment. Their mean age was 21.6 (SD= 5.2) years and they used English daily on average 16.3% (SD= 9.1) of the time.

5.1.2 Materials

A forced choice identification task was used to measure non-native identification accuracy of the same two vowels (/iː-ɪ/) as in Experiment I. Six minimal pairs from the original 12 (*beat-bit, bead-bid, seed-Sid, seat-sit, heat-hit, heed-hid*) were chosen from the native speaker productions from Experiment I. After their extraction and isolation they were preprocessed for presentation. The test thus consisted of 12 words x 3 speech styles x 5 speakers = 180 items. The order of the items was randomized. Speaker variation, in addition to the studied speech rate variation, was included in order to reflect speech processing in daily communication more closely.

5.1.3 Procedure

The perception experiment was conducted in a computer room in the University of Barcelona. After filling in a language background questionnaire, the participants completed a short familiarization task in which they were able to listen to the test items and see the pictures that were associated with them. By using pictures, instead of words, the impact of orthographic cues was minimized. In the perception task, the participants were presented aurally with a word and they had to identify it by clicking on the corresponding picture. The task took approximately 10 minutes to complete.

5.2 Results

In order to determine the identification accuracy, a percentage of correct responses was calculated for each participant for the three speech styles and the two vowels. Next, one-way ANOVA with two within subject factors (*speech rate* (*SR*): *clear/citation/fast* and *vowel quality*: tense/lax) was conducted. The descriptives are shown in Table 5.

Table 5. Mean percentage of correct identification (SD in parentheses) for the two vowels over speech rates

	ID tense cl	ID lax cl	ID total cl	ID tense ct	ID lax ct	ID total ct	ID tense fa	ID lax fa	ID total fa
n=54	77.08 (14.84)	55.25 (19.09)	65.98 (14.06)	65.20 (15.26)	60.99 (17.71)	63.06 (14.99)	30.74 (17.02)	81.11(14.21)	55.93 (9.58)

The ANOVA yielded significant main effects of *SR* (F(2,52)=30.00; p<.001;η^2=.54) and *vowel quality* (F(1,53)=33.94;p<.001; η^2=.39), as well as a significant *SR*vowel quality* interaction (F(2,52)=97.39;p<.001;η^2=.79). This shows that the speech rate affected the two vowels differently as can be seen in Figure 3. Paired samples T-tests were ran in order to explore more fully the effects of the speech rate on the vowels. The results are shown in Table 6.

Figure 3. Effects of speech style on the accuracy of identification of the two vowels

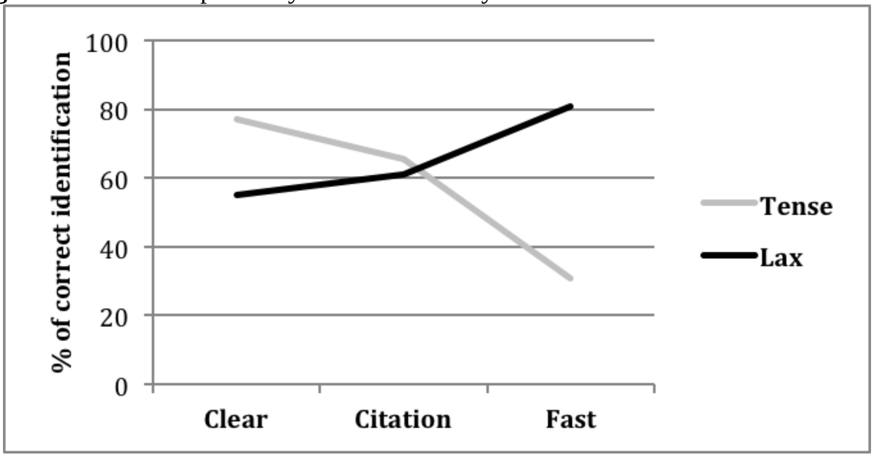


Table 6. The results of paired samples T-test over speech styles for the two vowels. Significance at p<.05 level is marked with an asterisk

	ID tense- ID lax over speech styles
ID tense cl – ID lax cl	*t*(53)=8.19, *p*<.001*
ID tense ct- ID lax ct	*t*(53)=2.20, *p*<.032*
ID tense fa – ID lax fa	*t*(53)=-14.89, *p*<.001*
	ID tense over speech styles
ID tense cl- ID tense ct	*t*(53)=7.04, *p*<.001*
ID tense cl- ID tense fa	*t*(53)=16.14, *p*<.001*
ID tense ct – ID tense fa	*t*(53) =14.10, *p*<.001*
	ID lax over speech styles
ID lax cl- ID lax ct	*t*(53)=-3.47, *p*=.001*
ID lax cl- ID lax fa	*t*(53)=-8.61, *p*<.001*
ID lax ct – ID lax fa	*t*(53)=-8.11, *p*<.001*

The paired samples T-tests further confirmed that the speech rate affected the identification accuracy of the two vowels differently. The identification accuracy for the tense vowel became significantly worse when the speech rate increased, whereas the identification of the lax vowel became significantly more accurate when the speech rate increased. Why was an increase in speech rate beneficial in the identification of /ɪ/ but detrimental in the identification of /iː/? Our tentative explanation is motivated by the cue-weighting strategies of the participants. Previous research (CEBRIAN, 2006; 2007; CERVIÑO; MORA, 2009; ESCUDERO, 2006; ESCUDERO; BOERSMA, 2004; MORA; FULLANA, 2007; KIVISTÖ DE SOUZA, 2011) has confirmed that Spanish and Catalan speakers distinguish these contrasting English vowels in terms of duration differences, namely perceiving and producing /ɪ/ as relatively short and /iː/ as relatively long, making little use of the spectral differences. Native speakers, on the contrary, rely mainly on the spectral cues in discerning the vowel pair (e.g. ESCUDERO; BOERSMA, 2004; BOHN; FLEGE, 1990). If our participants used duration to distinguish between the two vowels, it would make sense that the tense vowel was most accurately identified in the *clear* speech in which its duration was the longest, whereas the lax vowel was most accurately identified in the *fast* speech, in which its duration was the shortest. Likewise, the identification accuracy of the tense vowel dropped significantly in the *fast* speech because its duration approximated to that of the lax vowel. The same drop in identification accuracy can be observed for the lax vowel in the *clear* speech, which could indicate that reliance on duration cues was an ineffective strategy to discern between the two vowels. In the *citation* speech, the tense vowel was better identified than the lax vowel, but the difference was not as striking as in the other two speech styles

5.3 Discussion

This experiment studied the effects of speech rate on the perception of non-native vowels. The Spanish/Catalan bilinguals were shown to identify the two vowels differently depending on the speech rate. The lax vowel was best identified at the fastest speech rate, whereas the tense vowel was most accurately identified at the slowest speech rate. We hypothesize that this is due to the cue-weighting strategies Spanish/Catalan speakers use. Reliance on duration cues to discern between the English /iː-ɪ/ vowel pair caused unequal identification of the two vowels at different speech rates.

6 General discussion and conclusions

This paper studied the effects of speech rate on the production and perception of non-native vowels. Speech rate significantly affected both the production and perception accuracy of the English /iː/ and /ɪ/. In production, the two groups produced both vowels shorter and more centralized as the speaking rate increased. This confirms previous research with native speakers and would suggest that speaking fast has a certain impact on vowel production regardless of language proficiency. Our production results also indicate that the non-native speakers struggled to produce target-like vowels in the *fast* speech, suggesting that the two vowel categories were not fully developed.

An important finding of the study was that the non-native speakers were found to treat the two vowels differently. In the production tasks, the non-native speakers produced longer lax vowels in the *citation* speech than in the *clear* speech, whereas the native speakers produced consistently longer vowels as the speech rate became slower. A possible explanation is that the non-native speakers were relying on duration differences in the production of the vowel pair and thus in the *clear* speech, established the contrast by exaggerating the shortness of the lax vowel and the length of the tense vowel. In the perception task, the lax vowel was identified most accurately in the fastest speech rate, whereas the tense vowel was best identified in the slowest speech rate. We suggest that this is due to the ineffective cue-weighting strategies the participants were using.

The implications of the study can be used to develop more effective pronunciation training programs. First, it is clear that Spanish/Catalan speakers of English benefit from perceptual training to create more effective cue-weighting strategies. Reliance on duration is clearly not an effective strategy to discern the English /iː-ɪ/ vowel contrast and directing the learners' attention on the spectral differences is necessary. Further research could seek to determine whether teaching pronunciation at clear, normal or fast speech rates has different results in the acquisition of more target-like pronunciation. We believe that adding variability to L2 pronunciation training through using speech at different speeds can only be beneficial for foreign language learners by reflecting the world outside the classroom.

Acknowledgements

We wish to thank Angelica Carlet for her help in the translation of the abstract into Portuguese. We would also like to thank Eva Cerviño, Lidia Montero, Youssef Rochdi and Elena Safronova for their help in the Experiment II data collection.

References

BOERSMA, P., WEENINK, D. *Praat:* doing phonetics by computer. Versão 5.2.26. Disponível em: <http://www.praat.org>. Accesso em: maio 2011. [software].

CEBRIAN, J. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, v. 34, p. 372-387, 2006.

CEBRIAN, J. Old sounds in new contrasts: L2 production of the English tense-lax vowel distinction. In: TROUVAIN, J.; BARRY, W. (Ed.). INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 16., 2007, Saarbrücken. *Proceedings*... Saarbrücken: Universität des Saarlandes, 2007. v. 3, p. 1637-1640.

CERVIÑO, E.; MORA, J.C. When over-reliance on duration does not mean perceiving duration differences. In: ASBY, M.; MAIDMENT, J. (Ed.). PHONETICS TEACHING AND LEARNING CONFERENCE, 2009, London. *Proceedings*... London: University College London, 2009. p. 27-30.

DETERDING, D. The formants of monopthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, v. 27, p. 47-55, 1997.

ESCUDERO, P. The phonological and phonetic development of new vowel contrasts in Spanish learners of English. In: BAPTISTA, B.; WATKINS, M. (Ed.). *English with a Latin Beat:* studies in Portuguese/Spanish-English interphonology. John Benjamins Publishing Company, 2006.

ESCUDERO, P.; BOERSMA, P. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, v. 26, n. 4, p. 551-585, 2004.

FLEGE, J. E. Methods for assessing the perception of vowels in a second language. In: FAVA, E.; MIONI, A. (Ed.). *Issues in Clinical Linguistics*. Padova: UniPress, 2003. p. 19-44.

FLEGE, J. E. 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, 1995. p. 229-273.

FLEGE, J. E.; MUNRO, M.; MACKAY, I. R. A. Factors affecting strength of perceived foreign accent in second language. *The Journal of the Acoustical Society of America*, v. 97, n. 5, 1995.

FLEGE, J. E.; SCHMIDT, A. M. Native speakers of Spanish show rate-dependent processing of English stop consonants. *Phonetica*, v. 52, p. 90-111, 1995.

FOX, R. A.; FLEGE, J. E.; MUNRO, M. J. 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,* v. 97, n. 4, p. 2540-2551, 1995.

FRIEDA, E.; NOZAWA, T. You are what you eat phonetically: The effect of linguistic experience on the perception of foreign vowels. In: BOHN, O.; MUNRO, M. (Ed.). *Language experience in second language speech learning:* In honor of James Emil Flege. John Benjamins Publishing Company, 2007.

FRIEDA, E.; WALLEY, A.; FLEGE, J. E.; SLOANE, M. Adult's perception and production of the vowel /i/. *Journal of Speech, Language and Hearing Research*, v. 43, p. 129-143, 2000.

GRANDLUND, S.; HAZAN, V.; BAKER, R. An acoustic-phonetic comparison of the clear speaking styles of Finnish-English late bilinguals. *Journal of Phonetics*. A ser publicado.

HAZAN, V.; BAKER, R. Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, v. 130, n. 4, p. 2139-2152, 2011.

JOHNSON, K. Adaptive dispersion in vowel perception. *Phonetica*, v. 57, n. 2-4, p. 181-188, 2000.

JOHNSON, K.; FLEMMING, E.; WRIGHT, R. The hyperspace effect: phonetic targets are hyperarticulated. *Language*, v. 69, n. 3, p. 505-528, 1993.

KIVISTÖ DE SOUZA, H. Effects of speech style on the use of temporal and spectral cues in the production and perception of a non-native vowel contrast. Dissertação de Mestrado (não publicada), Universitat de Barcelona, Barcelona, 2011.

LINDBLOM, B. Explaining phonetic variation: A sketch of the H&H theory. In: HARDCASTLE, W. J.; MARCHAL, A. (Ed.), *Speech production and speech modeling*. Amsterdam: Kluwer academic, 1990.

MILLER, J.; VOLAITIS, L. Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*, v. 46, n. 6, p. 505-512, 1989.

MOON, S. J.; LINDBLOM, B. Format undershoot in clear and citation speech: A second progress report. *STL-QPSR*, v. 30, n. 1, p. 121-123, 1989.

MORA, J. C. Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and non-words by Spanish/Catalan bilingual learners of English. In: ISCA WORKSHOP ON PLASTICITY IN SPEECH PERCEPTION, 2005, London. *Proceedings...*, London, 2005. p. 43-46.

MORA, J. C.; FULLANA, N. 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. (Ed.). INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 16., 2007, Saarbrücken. *Proceedings...* Saarbrücken: Universität des Saarlandes, 2007. v. 3, p. 1613-1616.

MORRISON, G. S. L1-Spanish speakers' acquisition of the English /i:/-/ɪ/ contrast: duration-based perception is not the initial developmental stage. *Language and Speech*, v. 51, n. 4, p. 285-315, 2008.

RALLO FABRA, L. Speaking rate effects in Catalan and English: a cross language study. *Estudios de Fonética Experimental*, v. 9, p. 243-262, 1997.

SCHMIDT, A. M.; FLEGE, J. E. Speaking rate effects on stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica*, v. 53, p. 162-179, 1996.

SCMIDT, A. M.; FLEGE, J. E. Effects of speaking rate changes on native and nonnative speech production. *Phonetica*, v. 52, p. 41-54, 1995.

SYRDAL, A.; GOPAL, H. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America*, v. 79, p. 1086-1100, 1986.

TRAUNMÜLLER, H. Auditory scales of frequency representation. Disponível em: <http://www.ling.su.se/staff/hartmut/bark.htm>. Accesso em: 30 maio 2011.

VOLAITIS, L.; MILLER, J. 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*, v. 92, n. 2, 1992.