Concordia Working Papers in Applied Linguistics

Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014 © 2014 COPAL

Effects of Speech Rate on the Intelligibility of Non-Native Vowels

Hanna Kivistö-de Souza
Universitat de Barcelona

Joan Carles Mora
Universitat de Barcelona

Abstract

This paper examined the intelligibility of native and non-native vowel productions at three speech rates in native and L2 listeners. In Experiment I, native English speakers produced /iː/-/ɪ/ in a Delayed Sentence Repetition Task at three speech rates (clear/citation/fast). The vowels were presented to 54 Catalan listeners in a forced-choice word identification task. The results revealed a significant main effect of vowel quality and speech rate, and a significant vowel quality x speech rate interaction (all *ps*<.001), indicating that Catalan listeners' identification accuracy was affected by speech rate and that they distinguished the vowels by means of duration. In Experiment II, vowel productions of Catalan speakers obtained in the same conditions as in Experiment I were presented to 6 native English listeners. Overall intelligibility decreased as speech rate increased. The results suggest that pronunciation and perceptual training for /iː/-/ɪ/ at differing speech rates could be beneficial for Catalan EFL learners.

Second language (L2) learners face several challenges when communicating in the foreign language. Whereas at the initial stages of language learning, speech processing demands and lexical search hinder effective communication, at more advanced stages as the L2 context expands, L2 users are faced with coping with speech variability in the input. On the one hand, they have to adapt to speaker variation, such as

understanding native and non-native speakers, regional accents and differing voice qualities (e.g. male, female and infant voices). On the other hand, L2 learners need to be able to understand speech in different conditions such as speech varying in speed and in register and delivered through background noise. Whereas advanced L2 users living in immersion contexts have access to the necessary experience to face these daily communicative challenges with success, many advanced L2 learners who rarely use the language outside the classroom context face difficulties when being confronted with naturalistic language use. Our study sought to determine how advanced L2 learners in a foreign language setting cope with the processing and articulatory demands posed by speech rate variation.

Speech Rate Effects in Production

Previous research with native speakers has shown that speech rate variation affects segmental duration: as the speed increases, segments become shorter. This phenomenon has been observed in several domains: Voice Onset Time (Grandlund, Hazan & Baker, 2012; Rallo Fabra, 1997; Schmidt & Flege, 1995, 1996; Volaitis & Miller, 1992), syllable duration (Schmidt & Flege, 1995), vowel duration (Deterding, 1997; Flege, 1988; Moon & Lindblom, 1989; Schmidt & Flege, 1995) and phrase duration (Schmidt & Flege, 1995).

Like the temporal dimension, the spectral dimension of vowels is affected by speech rate variation in native speakers. 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, Walley, Flege & Sloane, 2000; Moon & Lindblom, 1989). Kivistö- de Souza (2011) examined the effect of speech rate on the production of L2 vowels in 20 Catalan/Spanish bilinguals producing the contrasting English vowels /i:/-/ɪ/ at three speech rates. Acoustic analyses revealed that the non-native speakers showed a similar behavior to that of the native speakers, that is, at the slowest speech rate the vowels occupied more peripheral positions, and at the fastest speech rate the vowels were centralized.

From the aforementioned studies, one could conclude that many speech rate effects are due to general articulatory characteristics that occur across languages and in native and non-native speakers alike. Lindblom's hyper-hypo theory (1990) states that vowels undergo inherent modifications due to speech rate variation. Vowels in fast speech are

subject to undershoot, i.e. they do not reach their targets due to the fast overlapping articulatory gestures. In slow speech, the articulatory targets are reached and tongue positions are more extreme because the speaker is trying to maximize vowel quality differences and aid comprehension. We thus have a fairly good understanding of what occurs in speech, and more specifically in vowel production, when speech rate is modified. However, studies examining the effect of speech rate on the perception of non-native vowels are scarcer.

Speech Rate Effects in Perception

Perceptual adjustment at category boundaries due to speech rate variation has been documented by several studies in the domain of VOT. Native speakers and advanced L2 learners have been shown to adapt to the category boundary shifts caused by changes in speech rate (Flege & Schmidt, 1995; Miller & Volaitis, 1989; Volaitis & Miller, 1992).

Some studies (e.g., Frieda et al., 2000; Johnson, 2000; Johnson et al., 1993;) have established that the perceptual vowel space is expanded in relation to the acoustic vowel space. Johnson (Johnson et al., 1993) termed this the hyperspace effect. According to these studies, listeners prefer vowel prototypes that have the most extreme values, i.e. vowels that maximize perceptual contrasts and are produced at slower speech rates to those produced at faster speech rates.

Effects of speech rate on the perception of L2 speech have been scarce. Munro and Derwing (2001) tested how speech rate affected accentedness and comprehensibility of non-native sentence productions in English raters. They used computer software to compress and expand non-native sentences to 10% faster and 10% slower speech rates than those of the original sentences. Their results revealed that the fast stimuli were rated as the most comprehensible and less accented and that variations in speaking rate made a small but significant contribution to the ratings. The speaking rate for optimal intelligibility was somewhat faster than the rate actually used by the majority of non-native speakers. Shi and Farooq (2012) studied the intelligibility (as measured by key-word recognition) of L2 speech at five speech rates in quiet and noisy listening conditions in monolingual and bilingual listeners. Both listener groups' performance worsened when the speech rate increased and although the bilingual listeners benefitted from slower speech rates, they did not do so to the extent of monolingual listeners.

The literature reviewed here suggests that listening to speech at faster or slower than normal rates is different from listening to speech at normal speed. Whereas non-native listeners may benefit from slower speech rates, especially in the initial stages of acquisition when speech processing is less automatized, native listeners seem to prefer faster speech rates. To our knowledge, no previous studies have tried to determine the effect of speech rate variation in the perception of non-native vowels.

THE PRESENT STUDY

The present study consisted of two experiments that tested the intelligibility of native vowel productions in L2 listeners at varying speech rates (Experiment I) and the intelligibility of L2 vowel productions in native listeners at the same speech rates (Experiment II). The three speech rates studied were labeled as clear, citation and fast. The clear speech rate was understood as slow-paced and hyperarticulated and had a mean speed of delivery [segments per second] of 6.15. The citation speed [mean speed of delivery: 9.06] was understood as normal reading speed and the fast speech rate [mean speed of delivery: 16.41] was defined as casual, corresponding to the fast-paced speed of a conversation between two native speakers of a language.

We examined the perception of English /iː/-/ɪ/ vowel pair because implementing the contrast has been shown to be challenging for Catalan learners of English. The high front vowel region of Catalan has only one high front vowel /i/. Cebrian's (2006) acoustic analyses indicate that English /iː/ and Catalan /i/ are highly similar. This is further confirmed by high perceptual assimilation scores and goodness of fit ratings by Catalan and English speakers. Acoustic comparisons show that /ɪ/ is spectrally closest to Catalan /e/. Whereas English /iː/ tends to be assimilated to native Catalan /i/, /ɪ/ has more varied assimilation patterns and Cebrian found it to be assimilated to Catalan /e/, /i/ or /ɛ/. We could thus conclude that whereas English /iː/ can be easily assimilated to Catalan /i/, English /ɪ/ is a new vowel whose perception causes more problems. Additionally, Catalan speakers have been shown to discern the English /iː/-/ɪ/ contrast mainly based on temporal cues, making little use of spectral cues (Cebrian, 2006, 2007; Cerviño & Mora, 2009; Kivistö-de Souza, 2011; Mora & Fullana, 2007). This non-native like cue-weighting strategy might be problematic when the temporal dimension is affected by speech rate variation.

EXPERIMENT 1

Experiment I examined the intelligibility of native speaker /iː/-/ɪ/ vowel productions at three speech rates in Catalan EFL learners. We posed the following research question: Do changes in speech rate affect the perception of English /iː/-/ɪ/ in non-native listeners?

We hypothesized that the L2 listeners would have difficulties in perceiving the L2 vowels accurately due to incomplete category formation. Additionally, we expected that they would use a cue-weighting strategy different from that of native listeners, namely, that they would rely on temporal rather than spectral cues in vowel identification. If this were to be the case, the identification of the L2 vowels should be especially difficult at higher speech rates, since segmental duration is the dimension affected the most by speech rate variation.

METHOD

Participants. Participants were L1 English speakers (n=5) and L1 Catalan listeners (n=54). The native speakers of English were all female (mean age: 26.00). Three of them spoke the General American variety and two Standard Southern British English. Mann-Whitney U-tests found no significant differences in the formant values of the two dialect groups. The L1 Catalan listeners were Catalan-dominant Catalan/Spanish bilinguals (mean age: 21.6) who reported to use English 16.3% of the time.

Instrument. A forced-choice word identification task was created with stimuli obtained from native English speakers. Six CVC /iː/-/ɪ/ minimal pairs (*bead-bid, beat-bit, heed-hid, heat-hit, seed-Sid, seat- sit*) were elicited in carrier phrases at the three (clear/citation/fast) speech rates in Delayed Sentence Repetition tasks. The target /iː/-/ɪ/ word stimuli were segmented and excised from the carrier phrases and normalized for amplitude. The mean duration values from the five native English speakers were submitted to a two-way repeated measures ANOVA with Speech Rate (clear/citation/fast) and Vowel Type (tense/lax) as within subjects factors. The main effects of Speech Rate ($F(2,3)=396.98$; $p<.001$, $\eta2=.996$) and Vowel Type ($F(1,4)=69.60$; $p=.001$, $\eta2=.946$) were significant, as well as the Speech Rate x Vowel Type interaction ($F(2,3)=105.03$; $p=.002$, $\eta2=.986$). These results indicate that the duration of the tense vowel differed significantly across speech rates (the faster the speech rate, the shorter the vowel), but not the duration of the lax vowel, whose duration was similar in clear and

citation speech (Figure 1). Spectral characteristics of the stimuli can be seen in figure 2.

The task consisted of 180 items (5 speakers x 6 minimal pairs x 3 speech rates), which were presented in randomized order through computer software. The listeners heard the auditory stimuli through headphones and had to identify the corresponding word by selecting one of two pictures appearing on the screen. Prior to the perception task, the listeners completed a short familiarization task in order to learn the word-picture correspondences. Orthographic effects were expected to be minimized by using pictures.

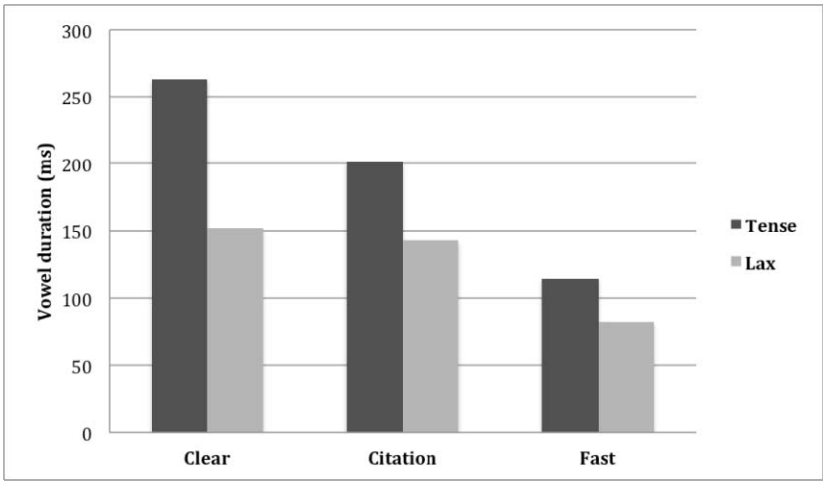

Figure 1. Mean duration values of the stimuli.

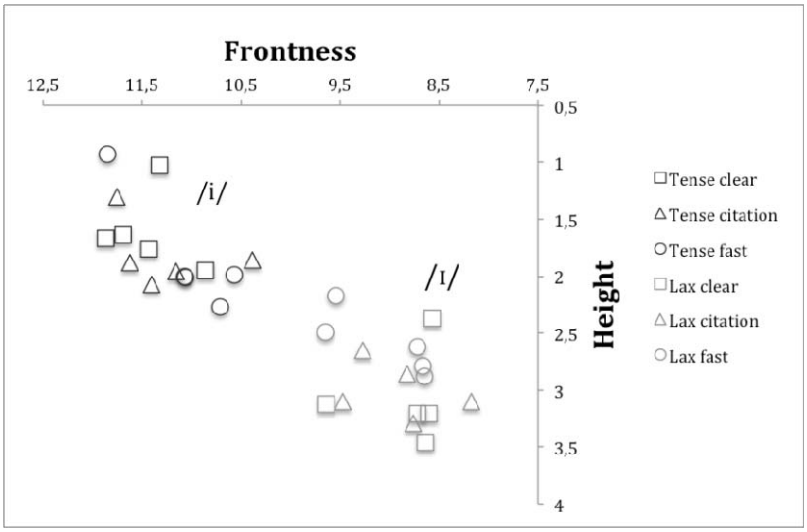

Figure 2. Mean quality values/speaker in a normalized Bark scale.

RESULTS

Percentage of correct identification (=intelligibility score) was calculated for each vowel and speech rate (Table 1). These were subjected to a one-way ANOVA with two within-subjects factors: Speech Rate (clear/citation/fast) & Vowel Type (tense/lax).

Table 1. Mean intelligibility scores by L2 listeners. Standard deviations in brackets

Speech rate	Clear			Citation			Fast		
Vowel type	tense	lax	total	tense	lax	total	tense	lax	total
ID accuracy	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 Speech Rate ($F(2,52)=30.00$; $p<.001$; $\eta 2=.54$) and Vowel Type ($F(1,53)=33.94$; $p<.001$; $\eta 2=.39$), as well as a significant Speech Rate xVowel Type interaction ($F(2,52)=97.39$; $p<.001$; $\eta 2=.79$), indicating that speech rate affected the intelligibility of the two vowels differently (Figure 3). Post-hoc paired samples t-tests confirmed that the intelligibility of the two vowels varied by means of speech rate (Table 2). The tense vowel was most intelligible at the slowest (clear) speech rate and its identification accuracy dropped as speech rate increased, reaching below chance level (30.74%) at the fastest speech rate. The lax vowel, on the other hand, was most intelligible at the fastest speech rate and its identification dropped as speech rate became slower, reaching near-chance level (55.25%) at the slowest speech rate. In other words, the two vowels behaved in opposite ways.

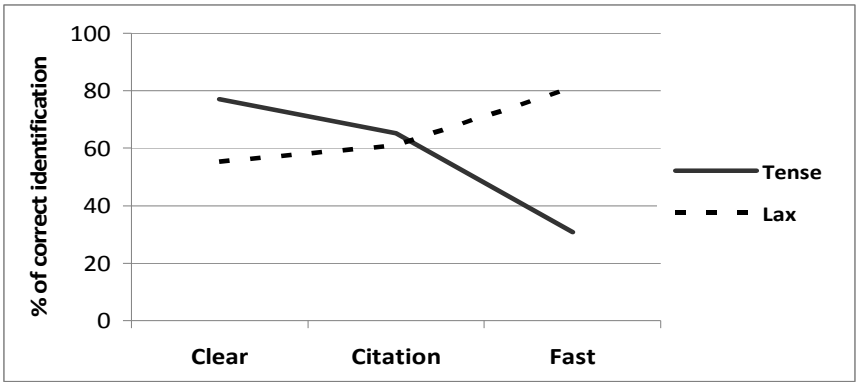

Figure 3. Intelligibility of /iː/ and /ɪ/ words by L2 listeners across speech rates.

Table 2. Results of paired samples t-tests on the intelligibility scores.

Contrast		Paired comparisons
Tense-lax over speech rates	clear	$t(53)=8.19$, $p<.001$*
	citation	$t(53)=2.20$, $p<.032$*
	fast	$t(53)=-14.89$, $p<.001$*
Tense over speech rates	clear - citation	$t(53)=7.04$, $p<.001$*
	clear - fast	$t(53)=16.14$, $p<.001$*
	citation - fast	$t(53) =14.10$, $p<.001$*
Lax over speech rates	clear - citation	$t(53)=-3.47$, $p=.001$*
	clear - fast	$t(53)=-8.61$, $p<.001$*
	citation - fast	$t(53)=-8.11$, $p<.001$*

Our tentative explanation as to why the two vowels behaved differently is motivated by the cue-weighting strategies of the participants. As stated above, previous research has established that Catalan speakers distinguish English /iː/-/ɪ/ in terms of duration differences, perceiving and producing /ɪ/ as relatively short and /iː/ as relatively long, making little or no use of spectral differences. Native English speakers, on the contrary, have been shown to rely mainly on spectral cues in discerning the vowel pair (e.g. Bohn & Flege, 1990; Escudero & Boersma, 2004). If the L1 Catalan listeners in our study used duration to distinguish the two vowels, it makes sense for the tense vowel to be most accurately perceived in clear speech where its duration was the longest, whereas the lax vowel was most intelligible in fast speech, where its duration was the shortest. Likewise, the intelligibility of the tense vowel dropped significantly in fast speech because its duration approximated that of the lax vowel. The same drop in intelligibility can be observed for the lax vowel in clear speech.

DISCUSSION

In this experiment we presented L1 Catalan listeners with instances of native English /iː/-/ɪ/ productions at three speech rates. The results indicate that the preferred cue-weighting strategy of L1 Catalan EFL learners is ineffective and even detrimental when speech is presented at different rates. Since the segmental duration of the vowels is highly affected by speech rate variation, listeners are forced to rely on the more constant spectral cues for identification, a task that seems very difficult for L1 Catalan EFL learners due to their lack of experience with small-scale spectral differences (Bohn, 1995).

As Experiment I indicated that L1 Catalan listeners used duration cues to identify English /iː/ and /ɪ/, even when duration distinctions were minimized due to speech rate variation, we were interested to see if this would also occur in production. That is, whether Catalan speakers would fail to produce the two English vowels with perceivable spectral differences so that native English listeners would have trouble perceiving the vowels accurately. We devised Experiment II for this purpose.

EXPERIMENT II

Experiment II examined the intelligibility of L1 Catalan speakers' English /iː/-/ɪ/ vowel productions at three speech rates by native English listeners. We posed the following research question: Do changes in speech rate affect the perception of L2 English /iː/-/ɪ/ by native listeners?

We hypothesized that difficulties in perception would occur since native English speakers have been shown to rely on spectral cues in discerning the /iː/-/ɪ/ vowel pair, whereas Catalan speakers have been shown to rely on temporal cues. Thus, if the Catalan speakers realized the two vowels with small spectral differences, and the temporal differences were neutralized due to speech rate, native English listeners should have problems in correctly identifying these L2 English productions of /iː/ and /ɪ/.

METHOD

Participants. Participants were L1 Catalan speakers (n=9) and L1 English listeners (n=6). In addition, two native English speakers (both female) were included in order to offer base-line data for native vowel perception at varying speech rates. The L1 Catalan speakers (5 male, 4 female) were Catalan-dominant Catalan/Spanish bilinguals (mean age: 22.6) who were completing the first year of an English degree at university. The L1 English listeners (mean age: 30.00) were 5 female and 1 male. They were from the United States (n= 4) and Southern England (n=2). They were living in Barcelona at the time of data collection and reported to use English 52% of the time and their L2 (either Spanish or Catalan) 44% of the time.

Instrument. A design similar to that used in Experiment I was used in Experiment II. A forced-choice word identification task was created containing the Catalan speakers' L2 English word stimuli and the base-

line native English speaker word stimuli. The same six minimal pairs were elicited at the three speech rates in Delayed Sentence repetition tasks. The procedure in preprocessing the word stimuli for auditory presentation was the same as that used in Experiment I. The L2 speakers mean vowel duration values were submitted to a two-way repeated measures ANOVA with Speech Rate and Vowel Type as within subjects factors. Significant main effect was found for both Speech Rate ($F(2,7)=72.96$; $p<.001$, $\eta 2=.954$) and Vowel Type ($F(1,8)=40.45$; $p<.001$, $\eta 2=.835$). The Speech Rate x Vowel Type interaction was also significant ($F(2,7)=21.41$; $p=.001$, $\eta 2=.859$) (Figure 4). The results indicate that the three speech rates differed significantly in their duration and that whereas the effect was constant for the tense vowel (the faster the speech rate, the shorter the vowel), the same did not occur for the lax vowel: the L2 speakers produced shorter lax vowels in the clear speech rate than in the citation speech rate, probably in order to maximize the temporal contrast between the tense and the lax vowel in this hyperarticulated speech rate. Spectral characteristics of the stimuli can be seen in figure 5.

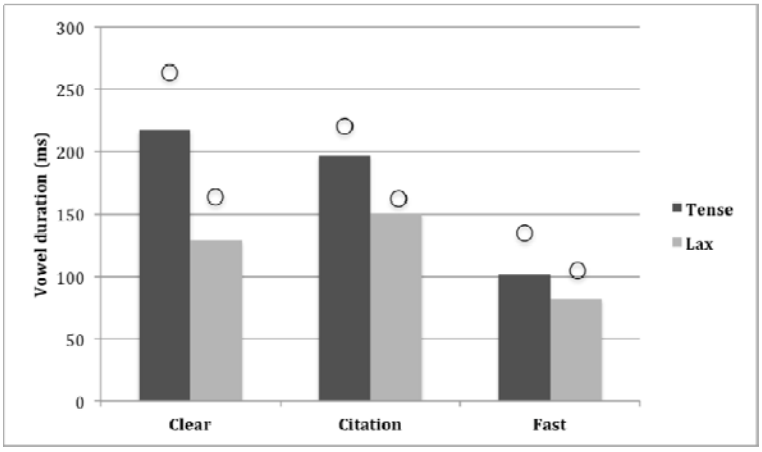

Figure 4. Mean duration values of the stimuli. L1 Catalan speakers in bars, native English speakers' means are represented by the empty dots above the bars.

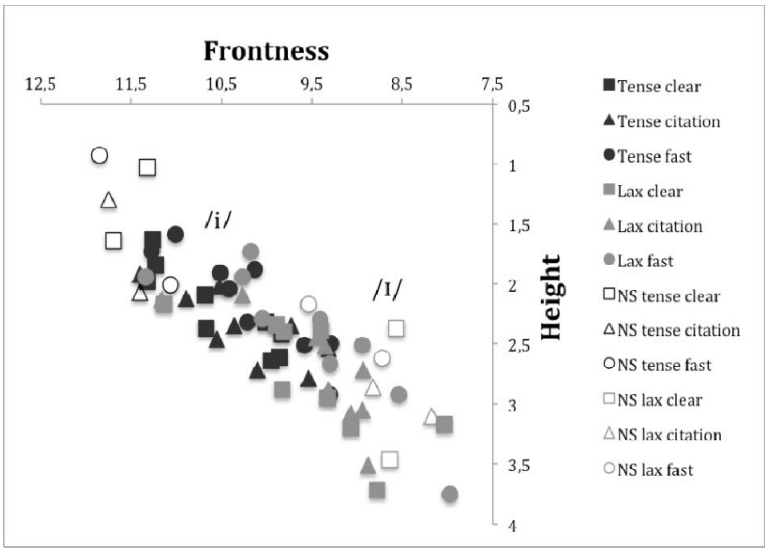

Figure 5. Mean quality values/speaker in a normalized Bark scale.

The task consisted of 346[1] items (11 speakers x 6 minimal pairs x 3 speech rates), which were presented in a randomized order through computer software. As in Experiment I, the listeners completed a short familiarization task prior to the perception task in order to learn the word-picture correspondences.

RESULTS

As in Experiment I, the mean percentage of correct identification (=intelligibility score) was calculated for each vowel and speech rate (Table 3). Mean intelligibility ratings for each word were subjected to an ANOVA separately for the two speaker groups (native/L2) with Speech Rate (clear/citation/fast) as the within-subjects factor and Vowel type (tense/lax) as the between-subjects factor.

[1] Some of the non-native data had to be excluded because either the word was not produced at one or several speech rate(s) or the produced word was clearly not the target-item (e.g. "Pete" instead of "seat"). In these cases, all the instances of the word were excluded. Consequently, the number of words included in the task by each speaker varied from 24 (four words missing) to 36 (all words included).

Table 3. Mean intelligibility scores by native listeners. Standard deviations in brackets.

Vowel Type	Speech rate	Native speakers (n=2)	L2 speakers (n=9)
Tense	clear	97.22 (6.48)	72.22 (31.98)
	citation	97.22 (6.48)	60.74 (36.43)
	fast	97.22 (6.48)	52.22 (37.53)
Lax	clear	94.44 (8.20)	70.65 (31.26)
	citation	98.61 (4.81)	64.49 (36.61)
	fast	94.44 (8.20)	69.20 (28.53)
Combined	clear	95.83 (7.37)	71.42 (31.45)
	citation	97.91 (5.63)	62.63 (36.37)
	fast	95.83 (7.37)	60.80 (34.19)

The ANOVA yielded a significant main effect of Speech Rate ($F(2,88)=3.57$; $p=.032$; $\eta2=.075$) for the L2 speakers. The main effect of Vowel Type ($F(1,89)=1.56$, $p=.215$ $\eta2=.017$) and the Speech Rate xVowel Type interaction ($F(2,88)=2,37$; $p=.099$ $\eta2=.051$) did not reach significance. For native speakers, all the contrasts were non-significant (Speech Rate: $F(2,21)=.96$, $p=.398$ $\eta2=.084$; Vowel Type: $F(1,22)=.54$, $p=.470$ $\eta2=.024$; Speech rate x Vowel type ($F(2,21)=.96$, $p=.398$ $\eta2=.084$). These overall results show that whereas the L1 English vowels were highly intelligible at all speech rates, this was not the case for the L2 English vowel productions, whose intelligibility was affected by speech rate. The L2 English vowels produced at the clear speech rate were significantly more intelligible than the L2 English vowels produced at the fast speech rate ($p=.046$). Independent samples t-tests were carried out to further examine the effect of speech rate on the native and L2 speaker vowels (Table 4). The paired comparisons revealed that the previously found effect of speech rate for the L2 speakers was due to the unequal identification of the tense and the lax vowel at the fast speech rate. Whereas the lax vowel had the mean intelligibility rating of 69.20%, the tense vowel was correctly identified only in 52.22% of the cases. In other words, the intelligibility of the tense vowel dropped to chance level when spoken at the fastest speech rate (Figure 6). We suggest that this is due to the non-native-like spectral values of the tense vowel in the fast speech rate (Figure 5). Only two L2 speakers produced tense vowels that were comparable to those produced by the native English speakers, whereas the rest produced vowels that were spectrally closer to English /ɪ/. When spectral values are ambiguous, the listeners might try to attend to secondary temporal cues (Bohn, 1995). In this case, the duration of the L2 English tense vowels was closer to the

duration of the L1 English lax vowel. This may have further confused the native L1 English listeners, which would explain the poor intelligibility score the tense vowels received at the fastest speech rate.

In general, the intelligibility of the L2 English vowels was rather low at all speech rates (71.42 - 60.80%), which would seem to indicate that the interlanguage English vowel categories of the L1 Catalan speakers were still in development. When looking at the spectral values of the L2 English stimuli (Figure 5), this indeed seems to be case: the spectral values of the L2 English tense and lax vowels overlap.

Table 4. Results of the independent samples t-tests on the intelligibility scores.

Contrast		Speaker group	
Vowel	Speech rate	L2	Native
Tense	clear	*t*(89)=.237; p=.813	*t*(22)=.920; p=.368
	citation	*t*(89)=-.490; p=.625	*t*(22)=-.596; p=.557
Lax	fast	*t*(89)=-2.433; p=.017*	*t*(22)=.920; p=.368

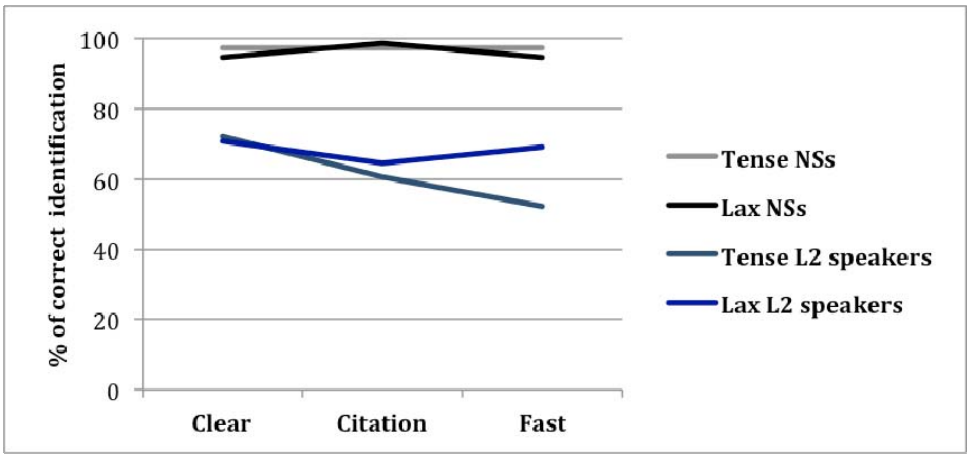

Figure 6. Intelligibility of /iː//ɪ/ in native listeners over speech rates.

DISCUSSION

In this experiment we presented L2 English vowels produced by L1 Catalan speakers at three speech rates to native English listeners. The results indicate that L2 English vowels spoken at the fastest speech rate were the least intelligible, whereas the L2 English vowels spoken in slower speech rates were more intelligible. Intelligibility of the L1 English

vowels in the control condition produced by native English speakers did not decrease as a function of speech rate. It would thus seem that native speakers are able to adapt their articulatory movements to different speech rates without it affecting intelligibility negatively. This does not seem to be the case for the L2 English speakers whose interlanguage categories are still developing. Following Schmidt and Flege's (1995, 1996) rationale, we believe that whereas accurate vowel production at the normal (citation) speech rate could be an indication of conscious attention to form, accurate production at atypical speech rates (slow and fast) could be an indication of robust L2 category formation. Our results indicate that the vowel categories of the L2 speakers were not robust enough to stand speech rate variation, which negatively affected intelligibility.

GENERAL DISCUSSION AND CONCLUSIONS

This paper presented the results of two intelligibility experiments with non-native vowels. In Experiment I, L2 listeners identified /iː/-/ɪ/ word stimuli as produced by native English speakers at three speech rates. The results indicate that the L2 English listeners were weighting duration cues, which hindered the intelligibility of the vowels when the duration of the stimuli was not constant due to speech rate variation. In Experiment II, native English listeners identified /iː/-/ɪ/ as spoken by L2 speakers at the same three speech rates. The results show that the overall intelligibility of the vowels was rather low and it improved when speech rate decreased.

Native speakers process speech rapidly and have learned to attend to the natural phonetic variability present in everyday speech. They are able to distinguish lexically relevant phonetic information from information that is indexical in nature and provides cues about the talker (e.g. emotional state, origin, social class), but does not change the meaning. As our results show, native English listeners are able to select the relevant phonetic cues in order to identify vowels at varying speech rates and ignore the irrelevant cues. L2 listeners on the other hand face a more difficult task. Their speech processing is slower and they do not share the amount of language experience native speakers have. The L2 English listeners in our study struggled when speech was presented at different rates. They were unable to attend to the spectral cues and to adapt to the temporal changes in vowel duration when identifying the non-native vowels. In producing the same vowels at the varying speech rates, they struggled with keeping the two vowels spectrally distinct, which affected their intelligibility. The effect speech rate had on the perception and

production of L2 English vowels by L1 English listeners indicates that the L1 Catalan speakers' L2 English vowel categories were not fully developed. Thus, speech rate could be used in future studies as a means of studying category formation.

As previous studies (e.g. Ylinen et al., 2009) have shown that it is possible to train L2 learners to re-weight acoustic cues, we believe that this type of perceptual training would be beneficial for Catalan EFL learners, as our results indicate that relying on temporal cues is detrimental when speech is presented at different rates. Training pronunciation and perception in noisy conditions and at differing speech rates could be more beneficial for L2 learners than focusing solely on the citation mode in optimal listening conditions, since this reflects real life communicative situations more closely.

ACKNOWLEDGEMENTS

We thank Eva Cerviño, Lidia Montero, Youssef Rochdi and Elena Safronova for their help with the data collection for Experiment I.

REFERENCES

Bohn, O.-S. (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* (pp. 279-304). Timonium, MD: York Press.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*, 303-328.

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

Cebrian, J. (2007). Old sounds in new contrasts: L2 production of the English tense-lax vowel distinction. In J. Trouvain & W. Barry (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. & 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., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551-585.

Flege, J. E., (1988). Effects of speaking rate on tongue position and velocity movement in vowel production. Journal of the Acoustical Society of America, *84*, 901-916.

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*, *43*, 129-143.

Grandlund, S., Hazan, V., & Baker, R. (2012). An acoustic-phonetic comparison of the clear speaking styles of late Finnish-English bilinguals. *Journal of Phonetics, 40*, 509-520.

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

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

Kivistö-de Souza, H. (2011). *Effects of speech style on the use of temporal and spectral cues in the production and perception of a non-native vowel contrast*. (Unpublished master's thesis). Universitat de Barcelona, Barcelona.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W.J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403-439). Kluwer academic: Amsterdam.

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

Moon, S. J., & Lindblom, B. (1989). Format undershoot in clear and citation speech: A second progress report. *STL-QPSR, 30*, 121-123.

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 J. Trouvain & W. Barry, (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, vol. 3* (pp. 1613-1616), 6-10 August 2007, Saarbrücken: Universität des Saarlandes.

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

Rallo Fabra, L. (1997). Speaking rate effects in Catalan and English: A cross language study. *Estudios de fonética experimental*, *9*, 243-262.

Schmidt, A. M., & Flege, J. E. (1995). Effects of speaking rate changes on native and nonnative speech production. *Phonetica*, *52*, 41-54.

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

Shi, L-F., & Farooq, N. (2012). Bilingual listeners' perception of temporally manipulated English passages. *Journal of Speech, Language and Hearing Research*, 55, 125-138.

Volaitis, L., & Miller, J. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America*, *92*, 723-735.

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*, 1319-1332.