

TEMPORAL CHARACTERISTICS OF ENGLISH SPOKEN BY NATIVE SPEAKERS OF CROATIAN

Damir Horga and Vesna Mildner

Department of Phonetics, University of Zagreb, Croatia

ABSTRACT

A group of 20 Croatian native speakers with various levels of English proficiency and a group of 7 English native speakers read a short passage of English prose, which was recorded and digitized, and the PCquirer computer program was used to obtain waveforms and spectrograms to measure selected variables. 10 English language experts assessed the level of subjects' English pronunciation proficiency. Croatian speakers were compared with the group of English native speakers by means of discriminative analysis in the following sets of variables: segmental durations (35 selected sounds), sentence and clause pauses durations, feet duration and several rate-of-speech variables. Regression analysis was used to show which of the temporal characteristics participated in predicting the pronunciation proficiency. It was shown that not all of the temporal characteristics of Croatian speakers speaking English that are significantly different from native English speakers had the corresponding influence on foreign accent assessment.

1. INTRODUCTION

Learning foreign language after the critical period increases the strength of foreign accent due to the loss of neurolinguistic plasticity and flexibility [5, 3]. Temporal organization of segments, syllables, stress groups and breath groups is language specific. When learning a foreign language the student's first language articulatory and perceptual skills will influence her/his foreign language temporal organization [8, 6]. The second-language Speech Learning Model [4, 9] attempts to explain that the inaccurate production of second language sounds is due to the learner's attitude that some differences between L1 and L2 sounds are phonetically irrelevant. This impedes the formation of new L2 phonetic categories. The model claims that L2 sound categories can be established if the dissimilarity between L2 sound and the nearest L1 sound is great enough and if the age of learning is early enough. The model also postulates that the sounds of L1 and L2 share the same phonological space and that is why in highly proficient bilingual speakers it is not only the phonetic space of the L2 that is influenced by the phonological space of L1, but vice versa as well.

After the L1 phonological system is mastered, it causes the filtering out of those acoustic and perceptive characteristics of L2 that are not important in L1 from the phonological point of view. The sensitivity level of distinguishing the L2 sounds depends upon the phonological system of L1.

That point of view is in agreement with the "gestalt theory" (theory of form) according to which the firmly structured perceptual forms (and the phonemes of the L1 are such forms)

are very difficult to destroy and do not permit a new form (the form of a L2 phoneme) to be perceived and articulated correctly.

From the point of view of information theory the L2 student's code is not tuned to the code of the native speaker. That is manifested as inadequate functioning of the decoding processor and results in the occurrence of semantic noise.

In general, this gives rise to a system of errors that has gained the name of interlanguage. Interlanguage is the result of L1 and L2 contact and of different operators influencing the learning process.

These processes are influenced by the learners' age of learning, gender, length of residence in the second language speaking country and especially by the relative frequency of the second and native language use [12].

The aim of the present study was to investigate the temporal characteristics of English spoken by Croatian speakers and to show which of the temporal characteristics are the best predictors of their pronunciation proficiency. This was tested according to two criteria: the temporal measurement of the pronunciation of native English speakers and the assessment of foreign accent by experts. It was supposed that the same variables which significantly differentiate Croatian English and native English speakers will also contribute to the assessment of foreign accent in Croatian English speakers.

2. MATERIALS AND METHOD

A group of 20 male native speakers of Croatian (CE), aged 20 to 24, with different levels of English speaking proficiency (10 students of English language and literature and 10 students majoring in other subjects at the Zagreb University) and a group of 7 male native adult speakers of English (NE) read a short passage of English prose (119 syllables). The students started to learn English between the ages of 6 to 14, thus having been learning it for 2 to 13 years. Every subject had two minutes of practice in reading the passage and after that their reading was recorded in a sound-proof room on MD.

The acoustic signals were digitized at 11 kHz and analyzed by means of PCquirer computer program. The temporal parameters were measured on the basis of spectrographic, waveform and auditory criteria.

Four groups of temporal variables were chosen: segmental duration, foot duration, speech rate, sentence pauses and clause pauses durations.

Ten university teachers of English assessed each CE speaker twice on a ten-point scale for three pronunciation characteristics (general pronunciation, sounds and prosody) by listening to the recordings of CE speakers in two randomized orders. The experts were instructed to establish a reference

Table 1. Discriminative analyses of CE and NE speakers. (Abbreviations: St - stops, Af - affricates, F- fricatives, So - sonorants, SV - short vowels, LV - long vowels, D- diphthongs, SP - sentence pauses, CP - clause pauses, FI - feet isochrony, R- rate)

	Can R	Chi- Sqr	p	Centroids		Most differe- ntiating var.
				CE	NE	
St	.86	29.79	.00	0.97	-2.76	VOT [t]
Af	.22	1.16	.56	0.13	-0.36	Closure [tʃ]
F.	.71	15.24	.03	0.58	-1.65	[s]
So	.63	11.23	.02	-0.45	1.29	[ŋ]
SV	.66	13.02	.02	0.50	-1.44	[ʌ]
LV	.59	9.77	.08	-0.42	1.20	[ə:]
D	.48	6.05	.20	0.31	-0.81	[əʊ]
SP	.45	4.89	.56	-0.28	0.81	Pause #6
CP	.71	15.70	.02	-0.58	1.16	Pause #4
FI	.60	9.45	.31	0.43	-1.23	Foot #8
R	.69	14.18	.05	0.55	-1.57	Rate of art.

grade 10 for native speakers of English.

The difference between CE and NE speakers in temporal variables was determined by discriminative analysis and the predictability of pronunciation quality on the basis of temporal characteristics of CE speakers by regression analysis.

3. RESULTS

The results will be presented in two parts to answer the two principal questions that were mentioned before: the question about temporal differences between English spoken by CE and NE speakers, and the question about the structure of temporal characteristics which contribute to the foreign accent of CE speakers when they speak English.

3.1. Temporal differences of English spoken by CE and NE speakers

The 119 syllables of the passage were read on the average in 31.7 sec by CE and in 26.6 sec by NE speakers. There were 35 variables of segmental durations ([p], [t], [k], [tʃ], [f], [v], [θ], [s], [ʃ], [x], [w], [r], [ŋ], [l], [i:], [ε], [æ], [ə], [ə:], [ʌ], [ɑ:], [ai], [ei], [əʊ], [iə]). For stops and affricates the durations of closure, VOT and friction were measured. The 35 segmental variables were grouped and analyzed in seven sound categories: stops, affricates, fricatives, sonorants, short (lax) vowels, long (tense) vowels and diphthongs.

There were six between-sentence points and six between-clause points and duration of the pauses at these places constituted two sets of variables: the duration of sentence pauses and the duration of clause pauses.

The sentence "| I was | coming | back | from Devonshire | to London | on my | old | motor-bike |", divided into 8 feet represented 8 variables to test the foot isochrony.

Finally, seven measures constituted the set of rate

Table 2. Regression analyses of temporal characteristics on condensed assessments of pronunciation quality. (Abbreviations as in Table 1)

	R	R ²	p	Most influencing var.	BETA
St	.76	.58	.04	Closure of [k]	.58
Af	.42	.17	.20	Closure of [t]	-.37
F.	.79	.63	.05	[ʃ], [x]	-.59, .45
So	.34	.12	.73	[ŋ]	.36
SV	.45	.20	.63	/ɪ/	.33
LV	.37	.14	.81	/ɑ:/, /æ/	.27, -.27
D	.35	.12	.72	[əʊ]	.26
SP	.76	.58	.05	Pause #6	.86
CP	.75	.56	.06	Pauses #3 and #4	-.74, 1.51
FI	.79	.63	.09	Foot #6	.63
R	.84	.71	.02	Number of syll. in hesit.	.63

variables. They were: rate of articulation (duration of speaking excluding pauses), rate of speaking (duration of speaking including pauses), average duration of interpause articulatory blocks, average duration of pauses, number of pauses, number of syllables in full hesitation pauses or repairs and the total duration of pauses.

The results of discriminative analyses of CE and NE speakers for sets of variables (Table 1) show that within segmental variables the majority of them differentiate those two groups of speakers at acceptable level of significance: stops (.00), fricatives (.03), sonorants (.02), short vowels (.02), and long vowels (.08). Only durations of affricates and diphthongs are not significantly different for two groups. Within suprasegmental variables the durations of clause pauses and temporal characteristics of rate of speech are significantly different ($p=.02$ and $p=.05$, respectively), while durations of sentence pauses and values of foot isochrony are not significantly different for two groups.

3.2. The influence of temporal characteristics of CE speakers on their foreign language accent

As the experts showed very high coefficients of reliability (between .93 and .94) it was reasonable to keep all 10 of them in further calculations. The first and the second assessments and the assessments of the three characteristics of pronunciation also showed very high correlations (between .95 and .98) so it was possible to condense all the grades of each expert for a particular speaker to one grade. The average CE speakers grades were from 4.88 to 9.22 with the total average grade 6.49.

The condensed assessments of the quality of pronunciation were used as criterion variable and the temporal characteristics of CE speakers as predictors in regression analysis.

The results in Table 2 reveal that multiple correlation is significantly for only two sets of segmental variables, namely by durations of stops ($p=.04$, $R=.79$) and durations of fricatives

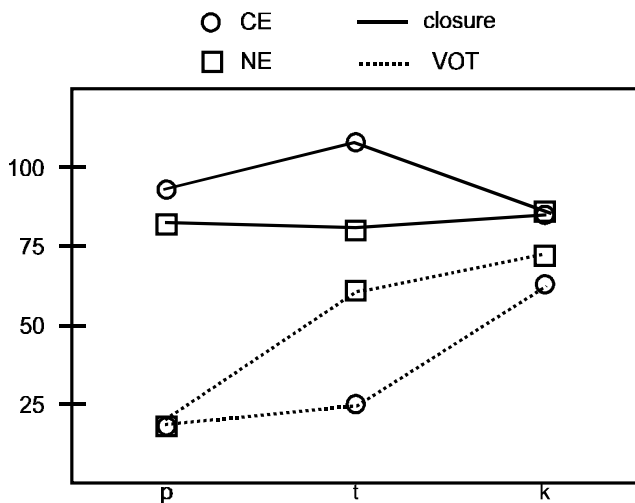


Figure 1. Mean stop (closure and VOT) durations (ms) for CE and NE speakers.

($p=.05$, $R=.79$). All of suprasegmental sets of variables are significantly predictors of foreign accent assessment: sentence pauses ($p=.05$, $R=.76$), clause pauses ($p=.06$, $R=.75$), foot isochrony ($p=.09$, $R=.79$), and rate of speech ($p=.02$, $R=.84$).

4. DISCUSSION AND CONCLUSIONS

The differences at the segmental level between CE and NE speakers were expected in stops which are aspirated in English (VOT: [p]-58 ms, [t]-70 ms, [k]-80 ms [10]) and nonaspirated in Croatian (VOT: [p]-15.3 ms, [t]-19.1 ms, and [k]-25.2 ms [1]), in the difference between short and long vowels and in durations of diphthongs which are not characteristic sounds in Croatian. The discriminative analysis showed that all sound categories, except affricates and diphthongs, differ in CE pronunciation from NE pronunciation but not all the sounds participate in that difference. Within stops (Figure 1) it is the sound [t] (closure: 108 ms for CE vs. 80 ms, for NE $p=.03$ and VOT: 25 ms vs. 61 ms, $p=.00$, respectively). In general, there is a strong tendency for CE speakers to pronounce English stop consonants with the longer closure duration and shorter VOT than NE speakers, but the difference is significant only for the sound [t].

In fricatives the sounds [v] and [s] are significantly shorter (CE 52 ms vs. NE 63 ms, $p=.05$ and CE 50 ms vs. NE 77 ms, $p=.01$) and [ʃ] and [x] are significantly longer (CE 143 ms vs. NE 123 ms, $p=.09$ and CE 98 ms vs. 68 ms, $p=.03$). The differences between sonorants are mostly influenced by the duration of [w] which is significantly ($p=.01$) shorter in CE pronunciation (67 ms) than in NE speakers (96 ms).

From the duration point of view vowels in Croatian generally can be classified into three categories: long stressed, short stressed, and unstressed vowels with the following approximate average durations in spontaneous speech: 100 ms, 80 ms, and 60 ms, respectively [1]. The durations of English long and short vowels in spontaneous speech are: 99 ms vs. 143 [11], or for stressed vowels about 130 ms and unstressed vowels about 70 ms [7]. As shown in Figure 2 in CE speakers the

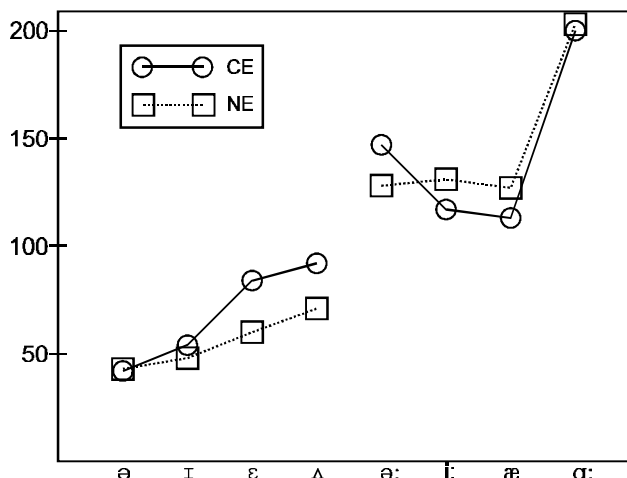


Figure 2. Mean vowel durations for CE and NE speakers.

difference between short and long vowels is smaller than in NE speakers so they only approximate the difference characteristic for the native speakers though in Croatian the difference between long and short vowels also exists and the positive transfer is possible. It also proves the second-language Speech Learning Model because though the two categories of vowels are established their authentic (native) values are not reached in CE speakers. This result shows that second-language phonological and phonetic systems in second-language speakers can be mastered at different levels of proficiency: with phonological level carrying mainly communicative function and phonetic level carrying mainly function of speakers' identification with the language. Generally, CE short vowels are longer (63 ms) than NE ones (45 ms), while the long vowels are shorter (119 ms vs. 144 ms, respectively). But the discriminative analysis showed the major influence only of the short vowels [ε] and [Δ] and the long vowel [ə:] with $p=.00$ and $p=.01$, respectively.

Due to the fact that there are no diphthongs in Croatian and that English diphthongs would be pronounced by Croatian speakers as a combinations of two vowels, it was supposed that diphthongs in CE speakers' pronunciation would be longer than in NE speakers. But that hypothesis was not proved and the durations of the diphthongs in both groups are practically the same (126 ms for CE vs. 124 ms for NE).

Though there were no significant differences between sentence pauses, CE speakers make generally shorter pauses than NE speakers (582 ms vs. 632 ms). The clause pauses on the other hand, though practically of the same average duration (CE 302 ms vs. NE 306 ms), differ significantly due to the pause in "I heard a scream, | a woman's scream." This result differs from Flege's investigations of pause influence on foreign accent [2]. That study provided an attempt to determine whether the speech fluency affects global foreign accent judgments. It was shown that removing pauses from the sentences spoken by Chinese subjects who speak English did not change foreign

accent scores. This suggested that foreign accent judgments depend mainly on segmental and suprasegmental articulation.

It was not proved that syllable-timed Croatian influences the pronunciation of English as far as the foot timing is concerned. Though we would expect that the variability of the 8 measured feet would be greater in CE speakers under the influence of syllable-timed Croatian (L1) the variability of the feet compared to the average duration of the foot was 62% for both groups of speakers.

Finally, the rate of articulation measured as the average duration of the syllable was shown to be the most influential of the rate variables (CE 199 ms vs. 180 ms for NE, $p=.05$).

The multiple correlations (Table 2) show that the set of suprasegmental variables (especially the rate of speaking) are better predictors of the foreign accent in CE speakers than the segmental variables, and within segmental variables the consonants are better than vowels and diphthongs.

Some of the variables which were expected on the basis of discriminative analysis to predict the quality of the English pronunciation did not show that (Compare the most differentiating and the most influencing variables in Tables 1 and 2). Within the segmental duration good predictors are: closure of [k] and duration of fricatives [f], [ʃ] and [x]. Durations of sentence and clause pauses showed to be good predictors at stylistically marked places. Final foot in the sentence also showed to be the good predictor probably because of the difference in final lengthening. The number of hesitations is the best predictor within the rate variables.

In conclusion, it can be said that speakers showed the following deviations from standard English pronunciation determined by the results of discriminative as well as regression analysis: longer closure and shorter VOT of stops, longer fricative noise in [s], [ʃ] and [x], shorter sentence pauses, slower rate of articulation and greater number of hesitations.

The results also reveal that, in general, the temporal variables which are in the process of speech production [9] mainly under the influence of the higher levels of the production (conceptualizer and formulator) than the lower levels (articulator) in foreign accent realization and judgment are of greater significance.

ACKNOWLEDGEMENTS

This research was supported by the grant 130721 of Ministry of Sciences and technology of Republic of Croatia.

The help of N. Lazi in preparing the manuscript is gratefully acknowledged.

REFERENCES

- [1] Bakran, J. Bakran, J. 1996. *Zvu na slika hrvatskoga govora*. IBIS grafika, Zagreb.
- [2] Flege, J. E. 1988. Factors affecting degree of foreign accent in English sentences. *Journal of Acoustical Society of America*, 84, 1, 70-79.
- [3] Flege, J. E. 1991. Age learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of Acoustical Society of America*, 89, 1, 395-411.
- [4] Flege, J. E., Munoro, M., Mackay, I. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of Acoustical Society of America*, 97, 5, 3125-3134.
- [5] Flege, J. E., Schmidt A.M. and G. Wharton 1996. Age of Learning Affects Rate-Dependent Processing of Stops in a Second Language. *Phonetica*, 53, 143-161.
- [6] Horga, D. 1994. Horvatsko-russkaja interferencija (na material'e glasnyh). *Vestnik Sank-Peterburgskogo universiteta*, 2, 1 (2), 70-77.
- [7] Klatt, D.H. 1976. Linguistic use of segmental duration in English: Acoustic and perceptual evidence. *Journal of Acoustical Society of America*, 59, 1208-1221.
- [8] Laeufer, C. 1996. The Acquisition of a Complex Phonological Contrast: Voice Timing Patterns of English Final Stops by Native French Speakers. *Phonetica*, 53, 117-142.
- [9] Levelt, W.J.M. 1995. *Speaking. From Intention to Articulation*. A Bradford book, MIT Press.
- [10] Lisker, L. Abramson, A. S. 1967. Some Effects of Context on Voice Onset Time in English. *Language and Speech*, 10, 1-28.
- [11] Lisker, L. Abramson, A. S. 1995. A Long-Short Vowel Dichotomy in Fluent English? In Elenius, K and Branderud, P. (eds.) *Proceedings of The XIIIth International Congress of Phonetic Sciences*. Vol. 3, 226-229.
- [12] Strange, W. 1995. Phonetics of Second-Language Acquisition: Past, Present, Future. In Elenius, K and Branderud, P. (eds.) *Proceedings of The XIIIth International Congress of Phonetic Sciences*. Vol. 4, 76-83.