

# SPEAKING-RATE EFFECTS IN VOICELESS STOPS PRODUCED BY CATALAN SPEAKERS OF ENGLISH

Lucrecia Rallo Fabra

*Laboratori de Fonètica, Universitat de Barcelona, Spain*

## ABSTRACT

Seven native speakers of Catalan, who were relatively proficient in English, produced English /t/ at five speaking rates: very fast, fast, normal, slow and very slow. Duration of VOT and closure interval were measured among other temporal parameters. Analysis of the results revealed three different patterns of production: a group of subjects showed a Catalan-like rate effect: they produced /t/ with short-lag VOT values and they adjusted to changes in rate varying the duration of the closure interval; a second group showed an English-like rate effect adjusting the duration of VOT to rate changes. Finally, a third pattern of production mixed both the Catalan and English models: VOT and closure duration were equally varied across the rates. These results suggest that a group of proficient Catalan speakers of English with limited exposure to the L2, are able to modify VOT across speaking rates in the same way English monolinguals had done in earlier studies.

## 1. INTRODUCTION

It is widely acknowledged that a person speaking a second language tends to “transfer” the sound system of his native language to the target language. This phenomenon is most common with the sounds of the L1 that have a direct equivalent in the L2 (similar sounds, [1, 2]). In English /t/ is realized as a long-lag voiceless aspirated stop having a longer VOT duration (35-40 msec) as compared to the short-lag /t/ of Romance languages like Spanish or Catalan (10-15 msec). Previous studies have examined Spanish production of English stops. The findings show that Spanish speakers tend to produce English /t/ with VOT values more appropriate for Spanish /t/ or, in some cases, halfway between English and Spanish.

Research in speech perception conducted by Miller and cols. [3] showed that the temporal parameter known as VOT was highly influenced by speaking rate, having short durations at a fast rate speaking style and longer duration at a slow rate. The effects of rate were not only found in production but also in perception, causing variability in the location of the boundaries between two phonetic categories like /bi/ and /pi/.

Schmidt and Flege [5, 6] examined speaking rate effects in a group of native American speakers and two groups of Spanish native speakers differing in L2 experience. VOT was measured for /p/ as produced in the sentence “a green pea” at a normal and fast rate. The results of the acoustic analysis revealed that native American and experienced Spanish bilinguals adjusted VOT duration to speaking rate. As expected, VOT had a longer duration in the normal rate than in the fast rate. The values obtained for the experienced bilinguals closely approximated those obtained for the native English speakers. However, the

less-experienced Spanish group produced /p/ with values more appropriate for Spanish /p/.

In a previous study [4] we compared the effects of speaking rate in English and Catalan. With this aim, we recorded seven Catalan and seven American English monolinguals reading the word *tapo* in the context of a carrier sentence *No diguis tapo mai més* (Never say *tapo* again) and *I do miss tapo my friend*. The English monolinguals’ adjustment for rate was reflected in VOT duration, as we expected (fig. 1). However, the Catalan monolinguals kept VOT duration constant across the rates. Instead they increased the duration of the closure interval (fig. 2).

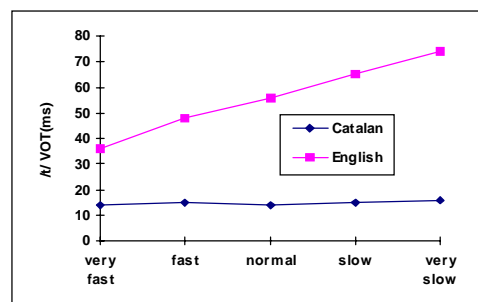


Fig. 1. The mean /t/ VOT values of the non-word *tapo* spoken at five rates by Catalan and American English monolinguals.

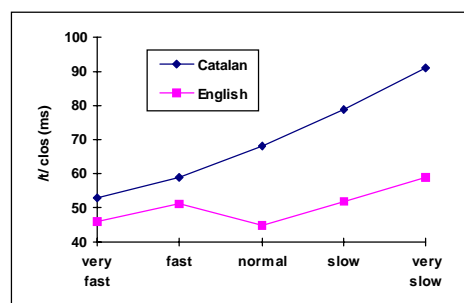


Fig 2. The mean closure interval durations of /t/ in the non-word *tapo* spoken at five rates by Catalan and American English monolinguals.

## 2. AIM

The aim of this study is to examine the speaking rate effects in Catalan speakers of English as a second language. We expect three possible patterns of results.

(1) Subjects may perform like the English monolinguals and adjust VOT duration to speaking rate. In that case VOT will have short durations at the fast rate and longer durations at the slow rate. In our previous study [4] the range of VOT durations was 35-75 msec. This would be evidence that these speakers are using pronunciation rules for a sound of the L2 that differs phonetically from the closest equivalent of the L1, such as /t/.

(2) The second logical possibility is that subjects transfer the pattern found in the Catalan monolinguals. This would mean that VOT duration would range between 10 /15 msec and it would not vary as a function of speaking rate. Instead, closure-interval duration would be adjusted to speaking-rate changes.

(3) Finally, it may be the case that Catalan speakers of English mix both the English and Catalan patterns. In that case, both VOT and closure-interval duration would vary to a lesser degree. After all, putting it in Grosjean's terms "a bilingual is not two monolinguals in one person".

### 3. METHOD

#### 3.1. Subjects

A total of seven native speakers of Catalan between the ages of 21 and 28 participated in our study. They all had started to learn English between 10 and 12 years old and were undergraduate (4 subjects) and graduated students (3 subjects) in English philology at the time of the investigation. All subjects had received formal instruction in English for an average of 10-15 hours a week. The percentage of L2 use was quite low (around 20%). Furthermore they had lived in an English-speaking country for at least one year and they did not speak any other foreign language apart from English.

#### 3.2. Speech Materials

Each subject produced a series of nonwords (*papo*, *dapo*, *tapo*) in the carrier sentence *I do miss \_ my friend* at five speaking rates following this sequence: very fast, fast, normal, slow, very slow. The test items were arranged in a pseudorandomized list three times each.

Subjects were told that they were to produce the nine-item list at five different rates. *Very fast* was defined as "as fast as you can speak without dropping sounds or distorting your speech" and very slow was defined as "as slow as you can speak without introducing pauses". These rates were given the values of 1.5, 1 and 0.5. The other two rates *fast* (defined as "halfway between very fast and normal") and *slow* (defined as halfway between very slow and normal) were assigned the values of 1.25 and 0.75, respectively.

#### 3.3. Procedure

Subjects were recorded in a sound booth, using a Marantz recorder, model CP-4300 and a *Shure* microphone. Before the actual elicitation task started, each subject was interviewed in English by the experimenter. The interview lasted about ten minutes. They were also given some training until they were comfortable with the task, i. e. until they could self-control and vary their speaking rate as requested.

Two different methods were used in the magnitude production task: the *block method* and the *sentence method*. As

subjects tend to stabilize at one particular rate, the former was preferred for the start. Therefore, subjects began by reading the nine sentences in the list at the very fast rate, then they continued to read the list a second time at the fast rate and so forth, following the order outlined above (very fast, fast, normal, slow, very slow). For the "sentence method" the task consisted on reading each single sentence at five rates. Most subjects reported this rapid changing of rates to be more difficult.

From the total production materials, only the sentences containing the *tapo* items were analyzed. A total of 420 samples were collected (3 items x 5 rates x 2 methods x 7 subjects). The production materials were later digitized at 22.0 kHz with 16-bit accuracy using the Syntrium Software Corporation waveform editor *Cool Edit*.

#### 3.4. Data analysis

A total of six temporal intervals were measured from each of the *tapo* tokens: /t/ closure interval, /t/ VOT, /a/ duration, /p/ closure interval, /p/ VOT, and the duration of the second vowel /o/.

## 4. RESULTS

#### 4.1 Word duration

	<i>very fast</i>	<i>fast</i>	<i>normal</i>	<i>slow</i>	<i>very slow</i>
CATALAN	250	284	316	367	441
ENGLISH	307	347	382	431	493
CSE	305	332	370	422	449

Table I: /tapo/ durations in msec for each speaking rate produced by Catalan and English monolinguals [4] and a group of Catalan speakers of English (CSE).

Table I shows how the total duration of the word /tapo/ varies across the different rates, having longer durations in the slow rates and shorter durations in the fast rates. A comparison of the data obtained for the EM with those obtained for the CSE reveals similar changes in overall word duration for all the rates except for the very slow. At this rate, CSE did not show quite the same changes of rate as the English monolinguals.

#### 4.2. Voice Onset Time

<i>SUBJECT</i>	<i>very fast</i>	<i>fast</i>	<i>normal</i>	<i>slow</i>	<i>very slow</i>
1	32	32	43	42	36
2	37	31	43	44	53
3	30	27	32	31	30
4	44	43	54	51	58
5	56	70	71	66	62
6	27	23	16	43	20
7	28	26	26	29	36
average EM	<b>36</b>	<b>48</b>	<b>56</b>	<b>65</b>	<b>74</b>
average CM	<b>14</b>	<b>15</b>	<b>14</b>	<b>15</b>	<b>16</b>

Table II: VOT durations in msec for the five rates produced by a group of CSE and two groups of English and Catalan monolinguals

As it can be seen from table II, the average VOT obtained for EM in our previous study [4], was 36, 56 and 74 msec. at the very fast, normal and very slow speaking rates. The changes of /t/ VOT across the five rates showed by the CSE are not as systematic as the changes showed by the EM group. CSE seem to find it hard to produce different VOT durations for the very fast and fast rates and also for the slow and very slow rates. Nevertheless, if we compare the VOT values obtained for the CSE with the values obtained for both EM and CM, we can say that although CSE do not show quite the same rate effects as EM, they can produce long-lag VOT durations that approximate EM's.

	/t/ VOT /w-dur	/t/ clos / w-dur
Sub. 1	0,51**	0,01
Sub. 2	0,60**	0,62**
Sub. 3	0,35*	0,66**
Sub. 4	0,62**	0,13
Sub. 5	0,12	0,20
Sub. 6	0,26	0,81**
Sub. 7	0,25	0,46**

Table III: Pearson correlations between word duration (w-dur) VOT (2nd column) and word duration and closure interval (3rd column). W-dur includes all segments of the word tapo except the closure interval (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ )

Pearson correlation coefficients were obtained between word duration / VOT and word duration / closure interval. Examination of table III suggests three different patterns in the CSE production of tapo at five speaking rates:

(1) For subjects 3, 6 and 7 closure interval duration was significantly correlated with word duration (subject 3:  $r = 0.66$   $n = 29$   $p = 0.00$ ; subject 6:  $r = 0.81$   $n = 30$   $p = 0.00$ ; subject 7:  $r = 0.46$   $n = 30$   $p = 0.005$ ). These talkers seem to follow the production pattern found in the Catalan monolingual speakers, which consisted of increasing the closure interval duration systematically as speaking rate decreased. The adjustment between speech rate and duration affects the closure-interval rather than VOT. At the normal speaking rate the longest VOT value is 30 msec (inf. 3) and 16 msec. (inf. 6). These are values appropriate for a short-lag /t/ like Catalan.

(2) Subjects 1 and 4 showed a tendency to imitate the EM's performance. The correlation coefficients between word duration and VOT duration were highly significant ( subject 1:  $r = 0.51$   $n = 29$   $p = 0.002$ ; for subject 4:  $r = 0.62$   $n = 29$   $p = 0.00$ ). These two talkers increased VOT systematically as a function of speaking rate., the correlation coefficients The range of VOT values obtained for these two subjects across the five rates is, nevertheless, narrower (16 msec.) than for the EM (38 msec.).

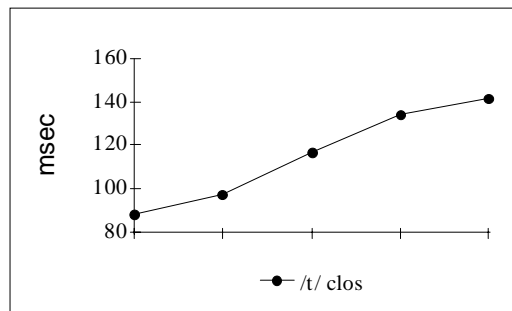


Fig. 3: subject 6 shows a Catalan pattern of performance: the duration of the closure interval increases as speaking rate decreases.

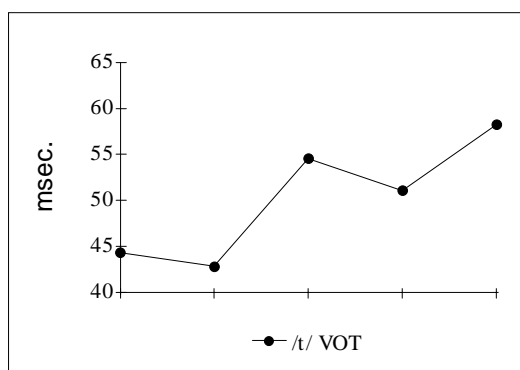


Fig. 4: subject 4 shows the EM pattern of performance: the duration of VOT increases as speaking rate decreases.

(3) The cases of subjects 2 and 5 have to be treated apart. The non significant correlation coefficients obtained for subject 5 ( $r = 0.20$ ,  $n = 24$ ,  $p = 0.242$  between closure interval and word duration;  $r = 0.12$ ,  $n = 30$ ,  $p = 0.257$  between VOT and word duration) tell us that he did not perform the task well. On the other hand, the correlation coefficients for subject 2 were equally significant for both VOT ( $r = 0.60$   $n = 29$   $p = 0.03$ ) and closure interval ( $r = 0.62$   $n = 29$   $p = 0.00$ ). This means that this subject's adjustment for rate is reflected on the duration of these two parameters together. That is, both VOT and closure interval increase in duration as speech rate decreases.

## 5. DISCUSSION

The rate-adjustment strategy, as reflected in word duration, used by CSE was similar to that used by English monolinguals. The duration of the word *tapo* increased as rate decreased. However, speaking rate did not equally affect VOT duration for the CSE as compared to the EM of our previous study [4]. The CSE might have been unable to control the speech production mechanism precisely enough to adjust VOT duration to the five different rates.

One of the claims in L2 research is that age is a determining factor that influences the accuracy of production of an L2 sound. This means that a person who starts learning a second language in adulthood hardly ever will succeed in producing the sounds of the target language like a native speaker does. Our results are, therefore, in accordance with this claim: although CSE managed to produce /t/ with VOT values that approximate those obtained for EM, they did not adjust the duration of this parameter to speaking rate to the same degree than EM did. The influence of the L1 is obvious: if VOT is not affected by speaking rate in Catalan, one can understand that CSE are reluctant to adjust the duration of this parameter when speaking rate requires it.

The production pattern of subjects 1, and 4 is the closest to the EM's obtained in the previous study [4]. This finding suggests that L1 patterns and structures can be modified at some stage of the L2 learning process. This is an important finding if we consider that the subjects who participated in our study had an average LOR (length of residence) in an English-speaking country of only one year and the percentage of L2 use was very low (20% or less). The pattern observed in subjects 3, 6 and 7 suggests that these talkers transferred the L1 production rule into the L2 varying the duration of the closure duration as a function of rate like Catalan monolinguals had done. Subject 2 mixed both the English and Catalan strategies. Rate adjustment was equally reflected in both VOT and closure interval duration. This suggests that this talker is simultaneously applying the L1 and the L2 production rules.

One of the hypothesis of Flege's Speech Learning Model is that bilinguals can establish a new phonetic category for an L2 sound that differs phonetically from the closest L1 sound, such as /t/. The present findings are insufficient to determine if any of the subjects have acquired a new category for English /t/. To prove that, we should test rate effects in perception. Nevertheless, we can predict that the chances for category formation are higher for subjects 1 and 4. These two talkers closely imitated the EM's pattern. It is very unlikely that subjects 3, 6 and 7 develop a new category for English /t/ because they transferred the L1 rate-adjustment strategy to the L2.

The marked individual differences in L2 production patterns found in the present study suggest, on the one hand, a lack of homogeneity in L2 background. It is difficult to establish which criteria should meet the subjects when they participate in a L2 study. Unlike the participants of most of Flege's studies, mostly an immigrant population living in the L2-speaking country, the subjects who participated in our study have limited exposure to the L2. On the other hand, perhaps we need to test a larger group of subjects in order to know which of

the three production patterns is most dominant in a CSE population.

## ACKNOWLEDGEMENTS

This research was supported by grant FP95 from the DGICYT of the Spanish Ministerio de Educación y Cultura to the author.

## REFERENCES

- [1] Flege, J.E. 1992. Speech Learning in a Second Language. In C. A. Ferguson, L. Menn and C. Stoel-Gammon (eds.), *Phonological Development. Models, Research, Implications*. Timonium, Maryland: York Press, 565-604.
- [2] Flege, J.E. 1995. Second language speech learning: Theory, findings and problems. In W. Strange (ed.) *Speech Perception and Linguistic Experience*. Baltimore: York Press, 233-277.
- [3] Miller, J. and Volaitis, L. 1989. Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*, 46, 505-512.
- [4] Rallo Fabra, L. 1998. Speaking rate effects in Catalan and English. A cross-language study. *Estudios de Fonética Experimental*, vol. IX, 243-272.
- [5] Schmidt, A. M. & Flege, J. E. 1995. Effects of Speaking Rate Changes on Native and Nonnative Speech Production. *Phonetica* 52, 41-54.
- [6] Schmidt, A. M. & Flege, J. E. 1996. Speaking Rate Effects on Stops Produced by Spanish and English Monolinguals and Spanish / English Bilinguals. *Phonetica* 53, 162-169.