Regressive Voicing Assimilation: Catalan Native Speakers' Performance in Advanced L2 English

Departament de Filologia Anglesa i Germanística

Grau en Estudis Anglesos

TREBALL DE FI DE GRAU

Carla Nobau Nanclares

Maria Josep Solé Sabater

9th January 2017

Table of Contents

Abstract

L1 processes have been proven to influence the production of a speaker's L2. In some languages such as Catalan, all obstruents in final position are voiceless due to Final Obstruent Devoicing (FOD). However, final obstruents will adapt their voicing to the following sound. This process is called Regressive Voicing Assimilation (RVA). In English, neither of these processes take place. The present paper aims to study the production of final fricative +C/V sequences (e.g. this boy, this tape, this orange) by advanced Catalan speakers of English in order to investigate if RVA is an L1-production strategy persisting in L2 advanced leaners. Two cues (voicing during the fricative, and preceding vowel duration) are examined. 12 Catalan native speakers (4th year students of English Studies at UAB) and 2 English native speakers were asked to read as naturally as possible a list of English sentences containing /s-z/ (near-) minimal pairs (e.g. price-prize) in contexts where RVA would typically occur in Catalan. Results showed that some subjects used one cue fricative voicing, while other speakers used the other cue, vowel duration, to cue consonant voicing. Few speakers seemed to use both cues. Still others did not make use of any of these cues and showed complete RVA. It can be concluded that RVA is an L1 process persisting in L2 English advanced leaners' speech.

1. Introduction

It is widely agreed that L1 categories and processes influence the production of a speaker's L2, supported primarily by Lado's Contrastive Analysis Hypothesis [1](1957). Yet, other factors have proved to have an influence on the production of an L2, such as markedness of the segment (see Eckman 1977). For example, voiced obstruents are said to be marked because it is difficult to maintain vocal fold vibration during an oral closure. In the case of English, native speakers learn to "enlarge the supraglottal cavity to sustain voicing during closure" (Cebrian, 2000: 3), and are thus are able to produce voiced obstruents in word-final position. Yet, final obstruents in English are partially devoiced, and the voicing contrast tends to be indicated by vowel duration. Romance languages' L1 speakers, however, find the production of voiced final obstruents to be difficult, as they need to acquire the articulatory gesture required to produce voicing in word final position. In addition, they need to identify and produce those differences in vowel duration.

Catalan is a language which has a contrast in obstruent voicing except word finally. Before a pause, voiced obstruents are realized as voiceless, due to final obstruent devoicing (FOD). The spelling in this language may show the letters <b d g> word finally, but in Catalan these are always pronounced as voiceless when followed by a pause: e.g 'Fre[t] (Eng. *Cold*).

Nevertheless, there are cases in Catalan where the word-final obstruent is produced with vocal fold vibration. When immediately followed by a voiced obstruent, it will be

[1] Proposed by Lado in 1957, CAH assumes that L1 patterns interfere in the learning of L2. The systematic comparison of the language and culture to be learned with the native language and culture, would permit the prediction of the patterns that will cause difficulty and those that will not.

realized as voiced. This is due to Regressive Voicing Assimilation (RVA). This process rule, though, applies not only to voiced but also voiceless obstruents (e.g. 'Mag dolent' [gd] (Eng. *Bad magician*), but 'Mag terrible' [kt] (Eng. *Terrible magician*). Thus, final obstruents adapt their voicing specification from the following obstruent. RVA applies to stops, fricatives and affricates, although different types of obstruents require different environments, as will be shown in section 2.1.3.

The main focus of this paper is the voicing of final fricatives by Catalan learners of English due to RVA. Voiceless word-final obstruents such as /s/ in Catalan 'després' [dəs'pɾes] (Eng. *Later*), will become 'després de sopar' [ðəs'pɾez͡ðə su'pa] (Eng. *After dinner*) and 'després venim' [ðəs'pɾez͡bə'nim] (Eng. *We are coming later*) due to the following sound being voiced. In English, however, such voicing does not occur; voiceless word-final obstruents do not assimilate the voicing of their environment. For example, 'once again' can only be pronounced as [wʌns əˈgɛn]. RVA, though, has been proven to be transferred into English by L1 Catalan speakers when producing English utterances (Cebrian, 2000; Cuartero, 2001).

Research has been conducted on the age of onset of L2 learning (AOL) and exposure to English to prove the influence of these factors on the production of word-final obstruents (Cebrian, 2000; Flege and Davidian, 1984; Fullana and MacKay, 2008; Fullana and Mora, 2009). According to them, varying amounts of exposure to English did not result in significant differences among the subjects' production of final obstruents in formal learning contexts (against claims made by Flege's Speech Learning Model). This result indicates that FOD and RVA are L1-based processes that persist in Catalan speakers' English production.

The present paper aims to explore the production of final fricative + C/V sequences (e.g., this boy, this tape, this orange) by advanced Catalan speakers in order to investigate if RVA is an L1-production strategy that persists in advanced L2 learners. On the basis of the existing literature, it is hypothesized that RVA will be present in the speech of advanced L2 learners. If his hypothesis is correct, NNS will voice fricatives when followed by voiced segments, and devoice fricatives when followed by voiceless segments. A second hypothesis is that, because voicing of final obstruents in English is not only indicated by presence or absence of vocal fold vibration during the consonant constriction, but also by the length of preceding vowels, advanced Catalan learners may (or may not) have learned to produce vowel duration differences in a native-like manner.

2. Background

The background to my research topic is divided into two subsections. The first section will present, define and illustrate some aspects crucial to the study, such as markedness, Final Obstruent Devoicing (FOD) and Regressive Voicing Assimilation (RVA). The second section will review previous studies on the production of voicing by Catalan speakers of English.

2.1.1 Markedness

Greenberg (1966) defined the notion of markedness as a structural opposition of two entities. The one entity more widely distributed in languages, and more simple articulatorily, is called ‘unmarked’, whereas the less frequent or less simple is called ‘marked’.

It is important to take into account that the notion of markedness can be applied not only to a particular language, but also cross-linguistically. Thus, a marked feature in one language can be an unmarked feature in another language. Considering this issue, Eckman

(1977) defines markedness as 'A phenomenon A in some languages is more marked than B if the presence of A in a language implies the presence of B, but the presence of B does not imply the presence of A'.

Having established the idea that markedness can be applied across languages, Eckman (1977) proposes the Markedness Differential Hypothesis, which helps to predict which areas of the L2 will be more difficult to acquire:

> (a) The relative degree of difficulty of the areas of the target language which are more marked than the native language will correspond to the relative degree of markedness.
> (b) Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.
>
> (Eckman, 1977: 321)

The notion of markedness is relevant to the present paper as the features that will be studied are unmarked in English, but marked in Catalan. RVA does not take place in English, as obsturents do not adapt their voicing specification from their environment. Catalan leaners of English might find the avoidance of RVA a difficult feature. Then, it can be predicted that the L1 process will be transferred to L2.

2.1.2 Final Obstruent Devoicing

The voicing contrast in obstruents is present in most languages. However, certain languages do not show such contrast in word-final position. The neutralization of the laryngeal contrast word-finally results in the phonetic devoicing of the final obstruent. Some Slavic and Germanic languages, such as German, Polish and Dutch (Kohler 1990; Grijzenhout, J. and Krämer, M., 1998; Grijzenhout, 2000) and some Romance languages, such as Catalan, are examples of such phenomenon. In these languages, the spelling as is, show <b d g>, but before a pause, these will be pronounced voiceless.

(1)

Catalan	Cor[p] ‘raven’	*Dutch*	Kwa[p] ‘lobe’	*Polish*	Jaku[p] ‘Jacob’
	Fre[t] ‘cold’		Hon[t] ‘dog’		Jo[t] ‘jodine’
	Ma[k] ‘magician’		Vraa[x] ‘question’		Pie[k] ‘freckle’

FOD has been extensively researched (e.g., Rothenberg 1968; Ohala 1983; Westbury 1983; West-bury and Keating 1986). FOD can be explained thanks to the aerodynamic voicing constraint (AVC), which accounts for the difficulty of voicing in obstruents as the ‘oral constriction associated with obstruents creates a build-up of intra-oral air pressure that reduced transglottal airflow and thus inhibits vocal fold vibration. Thus medial and final stops tend to devoice (‘passive devoicing’) after a few tens of milliseconds following the stop closure in the absence of additional articulatory adjustments.’’ (Solé, 2014)

In the case of languages that show a voicing contrast in final position, such as English, this difficulty is circumvented. In fact, the speakers of such languages have learned to enlarge the supraglottal cavity, or use other maneuvers to decrease oral pressure and sustain voicing in coda obstruents.

Taking the example of the voicing contrast in English and FOD in Catalan, the notion of markedness, previously discussed, can be illustrated. The presence of voiced word-final obstruents in English implies the presence of voiceless word-final obstruents. Yet, the presence of voiceless word-final obstruents in Catalan does not imply the presence of voiced word-final obstruents. Thus, voiced stops in final position are a marked feature cross-linguistically.

2.1.3 Regressive Voicing Assimilation (RVA)

Regressive Voicing Assimilation takes place when the voicing of a sound is influenced by the voicing specification of the following sound. RVA implies that voiceless sounds become voiced and vice versa.

RVA affects all obstruents in Catalan, although the environments differ for different obstruents. The following table summarizes the environments in which FOD and RVA occur in the different final obstruents in Catalan.

Table 1: Environments for FOD and RVA in Catalan.

Process	**Segment type**	**Context**	**Example**
FOD	Stop, fricative	_#	Pot /pɔd/, [pɔt] Vas /baz/, [bas]
RVA	Stops, fricatives	_# $C_{[+voice]}$	Pot donar [dd], vas dir [zd]
		_# $C_{[-voice]}$	pot tenir [tt], vas tornar [st]
RVA	Fricatives (but not stops)	_V	Vas anar [za] *But* pot anar [ta]

Considering Eckman' Markedness Differential Hypothesis, it can be predicted, and has actually been proven (Cebrian, 2000; Major and Faudree, 1996; among others) that the English voicing contrasts in final position is a difficult feature to be acquired by L2 learners.

2.2 Other studies

Fullana and Mora's study (2009) analyzed the perception and production of the voicing contrast in English by Catalan-Spanish bilinguals through a delayed sentence repetition task and an AXB task, where the subjects had to choose which phoneme they heard repeated in a minimal pair. The responses recorded were analyzed acoustically, including vowel duration, consonant duration and presence or absence of voicing during friction and stop closure. The results of the production task, as expected, showed evident differences in the production of segmental length and voicing consistency by the non-native speakers. On the other hand, the results of the AXB test showed that the perception of the minimal pairs was correct.

Fullana and MacKay (2008) studied the effects of age of onset of L2 learning (AOL) and exposure to English on the production of the voicing contrast in English word-final obstruents by Catalan-Spanish bilinguals. Acoustic measurement of vowel length, fricative duration and closure phase was done via Praat. The results of the experiment showed that Catalan-Spanish bilinguals failed to produce voiced segments at native-like levels, and they appeared to resort to L1 production rules such as FOD. Fullana et al. argued that no significant differences for AOL or exposure to English were found.

Cebrian's study (2000) focused on the transferability of L1 rules in the production of word-final obstruents by Catalan native speakers. His study provides a detailed list of Catalan neutralization processes that can be predicted based on the environment in which a segment is found. Cebrian studies the transferring of those processes into English, showing that L1 Catalan L2 English speakers show frequent cases of FOD and RVA. Cebrian also mentions that there was no significant difference on the results depending on the subjects' proficiency in English.

In sum, these studies indicate that RVA (and FOD) are persistent errors in Catalan speakers of English which are found at advanced stages of their learning. The present paper will focus on final fricatives and their realization by advanced Catalan speakers of English. Voicing contrast in final fricatives in English is indicated by a variety of cues (e.g. fricative duration, voicing during the consonant, preceding vowel duration). In this study, only two cues will be examined: fricative voicing and preceding vowel duration.

3. Methodology

3.1 Participants

12 native Catalan speakers (1 male, 11 female) participated in this study. They were 4th year students of the English Studies degree at Universitat Autonoma de Barcelona (mean age =21.8 years). All subjects were asked to respond a questionnaire (see Appendix 8.1) about their level of English. 9 of the subjects had the advanced level required. 3 of the subjects were discarded due to poor level of English and/or Catalan. Subjects differed in amount of exposure to English (school exposure vs language immersion).

Two English native speakers (1 male, 1 female), resident in Barcelona, were also recorded and analyzed, and served as the control group.

3.2 Tasks

The target of this study were the alveolar fricatives /s-z/. The subjects were asked to read as naturally as possible a list of sentences in English five times. The sentences contained (near-) minimal pairs (e.g. price-prize, nice-size) in meaningful sentences (see Appendix 8.2). The fricatives were placed in contexts where Regressive Voicing Assimilation would typically occur in Catalan. Thus, they were followed by a stop, a

voiced nasal and a vowel. The reading sessions were held and recorded in the speech lab at Universitat Autonoma de Barcelona.

The recordings were analyzed with Praat. The duration of voicing (and voicelessness) during the word-final fricatives was measured, in order to see where RVA applied. The duration of the vowel preceding the voiced/voiceless fricative was also measured. These two cues, voicing during the consonant constriction and preceding vowel duration are known to be the main cues to indicate voicing in final obstruents.

Sentences like (1) and (2) were intended to elicit RVA from voiceless to voiced alveolar fricatives. Sentences like (3) were intended to elicit RVA from voiced to voiceless alveolar fricatives.

(1) I think it's a nice movie

(2) I think it was nice of you to come today.

(3) The smaller size T-shirt fits her.

3.3 Segmentation and measurements

The duration of voicing (and voicelessness) during the fricative was measured, as well as the duration of the preceding vowel. All segments were measured both in milliseconds and percentage of voicing. The duration of voicing during the fricative was measured from the beginning of friction and periodic energy in the waveform and spectrogram to the end of periodic energy which coincided with an increase in random friction. Voicelessness during the fricative was measured from the beginning of aperiodic random friction, until the end of high energy. Vowel duration was measured from the first peak of periodic energy, until the beginning of friction in the waveform and the spectrogram.

When followed by a nasal, some of the fricatives were devoiced and voiced again right before the nasal occurred. This was due to anticipatory velum lowering during the end of the fricative.

3.1.1 Unused tokens

RVA in Catalan does not take place if the obstruent is followed by a pause (or major boundary). Tokens in which the speaker paused between the two segments of interest were discarded.

4. Results

The results section will be divided into two subsections. The first one deals with the results for fricative voicing. The second subsection will deal with the results for preceding vowel duration.

4.1 Voicing during the fricative

Although non-native speakers were expected to show relatively consistent results, a great deal of interspeaker variation was found. Close inspection of the measurements for voicing during the fricative constriction suggests that speakers can be categorized into 4 groups: (1) Native speakers (NS), (2) Native-like speakers (FS1), (3) Non-Native speakers who show a difference in the amount of voicing between /s/ and /z/, but with a greater amount of voicing in /z/ than NS (that is to say, they exhibit some voicing assimilation) (FS2), and (4) NNS who show no difference in voicing between /s/ and /z/, i.e., full voicing assimilation (FS3).

The results are shown in Figs. 1-8 and Table 2. We will first focus on the difference in voicing between final /s/ and /z/ in the various contexts for the 4 groups of speakers.

The NS group is composed by native speakers, J and AM, who serve as control. For native speakers (NS) the median voicing values for voiceless /s/ when followed by a voiceless consonant, vowel and nasal were 11.4, 10.5, and 13.9ms, respectively. The median voicing values for /z/ were 15.7, 34.3, 29.8 ms respectively. Note that the range 10.5-13.9 percentage of voicing during the initial portion of the /s/ can be attributed to 'voicing continuation' from the preceding vowel. That is, the vocal folds do not stop vibrating instantaneously, but take a few ms to open and cease vibrating for the /s/.

The individual results for NS are presented in Figs. 1 and 2, which show percentage of voicing on the Y axis when followed by a vowel (V) (e.g. pri**c**e of, pri**z**e is), a voiced Nasal (N) (e.g. pri**c**e never, pri**z**e money), and voiceless stop (T) (e.g. pri**c**e could, pri**z**e should). For Figures showing the absolute values for fricative voicing, please see Appendix 8.3.

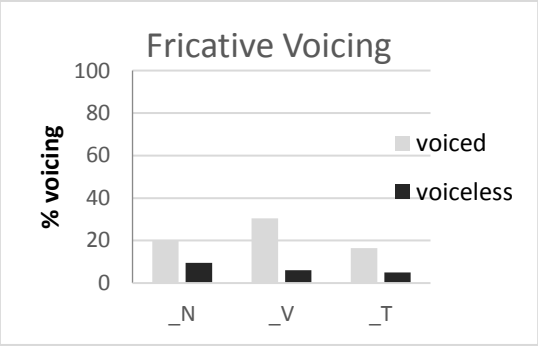

Figure 1. Mean % of fricative voicing for Native Speaker J.

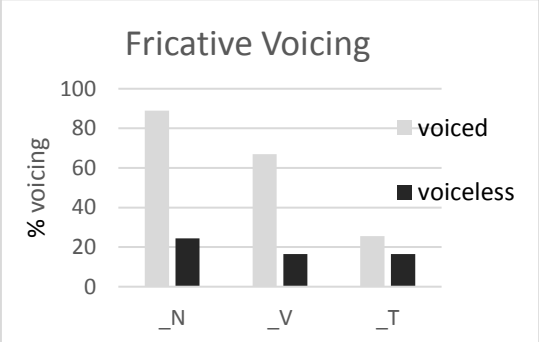

Figure 2. Mean % of fricative voicing for Native Speaker AM.

Although one of the native speakers (AM) shows longer voicing in fricatives than the other (J), we see differences in voicing between a voiced and voiceless fricatives in both native speakers.

The data for /sT/ sequences indicate that voicing continuation in native speakers ranges between 5-24.5 ms. Longer values than approximately 20ms can be considered purposeful voicing of the fricative.

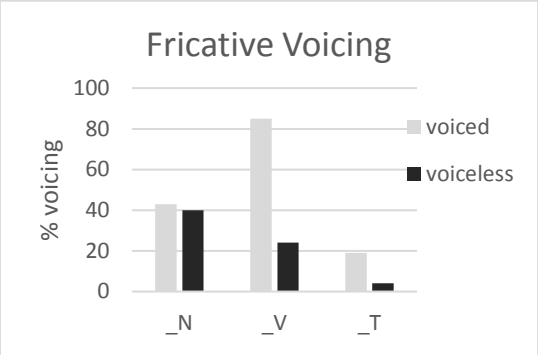

Figure 3. Mean % of fricative voicing for Catalan Speaker ER.

Group FS1 is composed only by one subject, ER, who showed a fricative voicing pattern similar to native speakers. As seen in Fig. 3, she shows a clear difference in voicing between voiced and voiceless fricatives, except when an N followed with a smaller difference, but the expected direction.

The second group, FS2, is composed by 5 Catalan speakers (JUD, ST, CL, GR, JC),[2] who show a somewhat ambiguous pattern. The two Figures presented (Figs. 4 and 5) represent the two main patterns in this group. Subjects in the FS2 group implement differences between a voiceless and a voiced fricative in some, but not all contexts. Still, their /z/s and /s/s have considerably more voicing than those of NS.

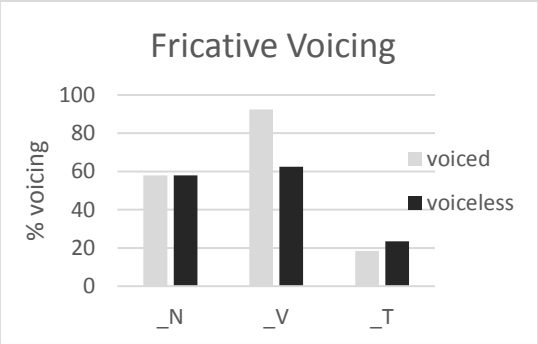

Figure 4. Mean % of fricative voicing for Catalan Speaker ST.

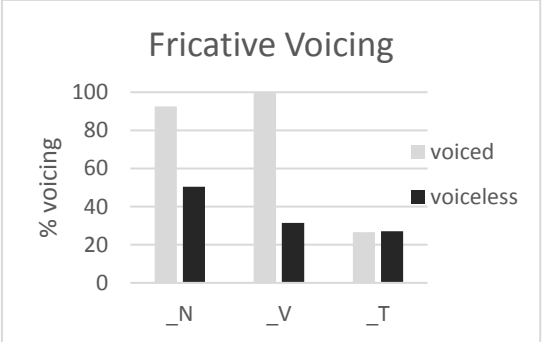

Figure 5. Mean % of fricative voicing for Catalan Speaker JUD.

Comparing the fricatives in a /_T/ environment in Fig. 4, we see that the subject's values for voicing in voiceless tokens are higher than those of the natives. More importantly, there is no difference in voicing between /s/ and /z/ when followed by a voiceless consonant. All subjects in this group, therefore, show RVA in this context. All subjects

[2] Due to space constraints, Figures for the individual speakers can be found in Appendix 8.3

in this group make a difference between /s/ and /z/ in the _V environment, with /z/ showing more voicing than /s/.

The two main patterns in this group are the following: some subjects show the same amount of voicing during the fricative in SN and ZN (see Fig. 4), showing RVA. Others (Fig. 5), show the voiced fricative to have more voicing, though both their /s/ and /z/ have more voicing than NS do. These subjects seem to attempt to make a difference between voiceless and voiced fricatives, but unlike the natives, they voice [z] and [s] for a longer period of time.

The last group, FS3, is composed by Catalan native speakers (AB, MM, MX), who showed clear and complete RVA. Comparing their voiced and voiceless fricatives, it is evident that there is no significant difference in amount of voicing, and their voicing values are up to six times higher than those of native speakers.

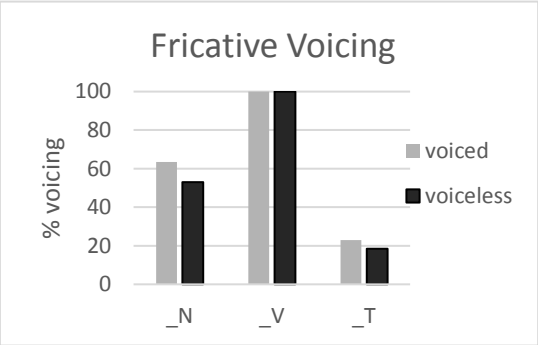

Figure 6. Mean % of fricative voicing for Catalan Speaker MM.

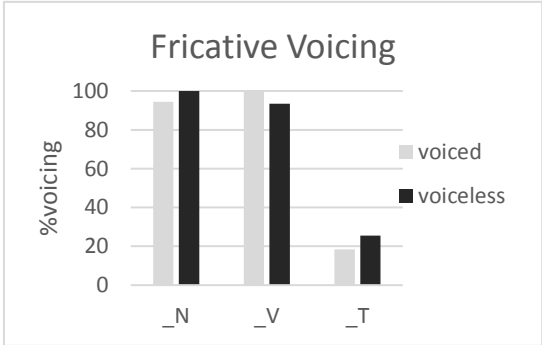

Figure 7. Mean % of fricative voicing for Catalan Speaker AB.

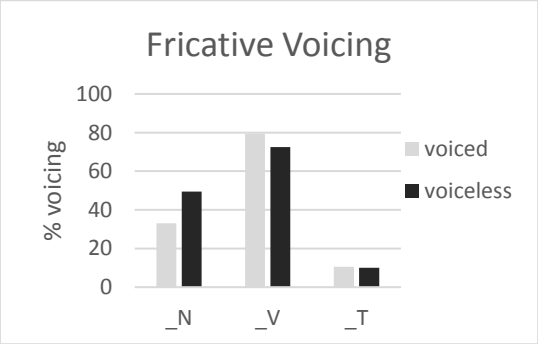

Figure 8. Mean % of fricative voicing for Catalan Speaker MX.

For this group of speakers, the extent of voicing is virtually the same for /sT/ and /zT/ sequences (see Figs. 6, 7, 8). That is, they do not make a difference between a voiceless and a voiced fricative in terms of voicing. In

this group, the difference in voicing between SN-SV and ZN-ZV is inexistent. These subjects show much higher voicing than natives for voiced and voiceless fricatives in the two environments, with median values of 100% of voicing in all contexts (zV, sV, zN and sN) for FS3 as opposed to median values of 34.4% (zV), 10.5% (sV) and 29.8% (zN), and 14% (sN) for natives. The fact that the voicing is the same for a voiceless fricative and a voiced fricative, indicates that there is complete voicing assimilation.

These observations were corroborated by the results of statistical tests. The descriptives per speaker and group and the results of the statistical tests are presented in Table 2. Non-parametric tests were used because the data were not normally distributed. Wilcoxon sum rank tests examining differences in voicing between /s/ vs /z/ when followed by voiceless consonants, vowels, and nasals separately were used for each of the 4 groups.

Table 2. Descriptive statistics and results of Wilcoxon sum rank tests comparing percentage of voicing in sT-zT, sV-zV, and sN-zN sequences for the different groups of speakers. *= p<0.05, **= p<0.01, ***= p<0.001

Group	context	Mean voicing (ms)	Mean voicing %	N	Speaker	Pooled Mean/Median voicing %	Comparisons	Wilcoxon sum rank test	p-value	significance
Native Speakers	sT	4.31	5	10	J					
	sV	6.67	6	10	J					
	sN	9.44	9.5	10	J					
	zT	12.58	12	10	J					
	zV	16.25	23.5	9	J					
	zN	19.5	20	10	J					
	sT	11.68	16.5	9	AM	10.7/11.4				
	sV	16.95	16.5	9	AM	11.2/10.5				
	sN	28.87	24.5	10	AM	17/14				
	zT	21.65	23	10	AM	17.5/15.7	sT-zT	W =120	< 0.05	*
	zV	72.64	100	5	AM	61.7/34.3	sV-zV	W = 45	< 0.001	***
	zN	55.01	89	10	AM	54.4/29.8	sN-zN	W = 51	< 0.001	***

Group	context	Mean voicing (ms)	Mean voicing %	N	Speaker	Pooled Mean/ Median voicing %	Comparisons	Wilcoxon sum rank test	Wilcoxon sum rank test	significance
FS1	sT	3.25	4	10	ER	4/0				
	sV	16.43	24	10	ER	24/11.8				
	sN	26.91	40	10	ER	40/25				
	zT	13.22	19	9	ER	19/11.6	sT-zT	W =17	< 0.05	*
	zV	53.46	85	10	ER	85/100	sV-zV	W = 9	< 0.01	**
	zN	28.81	43	9	ER	43/38.3	sN-zN	W = 34	< 0.05	ns

Group	context	Mean voicing (ms)	Mean Voicing (%)	N	Speaker	Pooled Mean/Median voicing %	Comparisons	Wilcoxon sum rank test	p-value	significance
FS2	sT	23.17	27	9	JUD					
	sV	30.81	31.5	10	JUD					
	sN	44.06	50.5	10	JUD					
	zT	40.18	26.5	10	JUD					
	zV	88.05	100	9	JUD					
	zN	81.2	92.5	10	JUD					
	sT	14.68	23.5	10	ST					
	sV	28.66	62.5	9	ST					
	sN	37.39	58	10	ST					
	zT	26.46	18.5	10	ST					
	zV	53.3	92.5	10	ST					
	zN	32.96	58	10	ST					
	sT	15.82	19.5	9	CL					
	sV	26.32	44	10	CL					
	sN	46.83	58.5	10	CL					
	sT	15.78	20.5	10	CL					
	sV	48.64	91.5	8	CL					
	sN	39.16	55	10	CL					
	sT	14.43	16	10	GR					
	sV	25.11	30	10	GR					
	sN	31.12	48.5	10	GR					
	zT	47.77	15.5	6	GR					
	zV	67.25	91	9	GR					
	zN	57.91	70.5	10	GR					
	sT	23.17	27	10	JC	22.6/18.9				
	sV	30.81	31.5	10	JC	39.9/30.5				
	sN	44.06	50.5	10	JC	53.2/40.6				
	zT	40.18	26.5	10	JC	21.5/18.7	sT-zT	W =1293	> 0.05	ns
	zV	88.05	100	10	JC	95/100	sV-zV	W =343	< 0.001	***
	zN	81.2	92.5	10	JC	73.7/78.3	sN-zN	W=1050	> 0.05	ns

Group	context	Mean voicing (ms)	Mean voicing %	N	Speaker	Pooled Mean/ Median voicing %	Comparisons	Wilcoxon sum rank test	p-value	significance
FS3	sT	14.04	25.5	10	AB					
	sV	44.16	93.5	10	AB					
	sN	57.21	100	9	AB					
	zT	18	18.5	8	AB					
	zV	67	100	9	AB					
	zN	67	94.5	9	AB					
	sT	10.95	18.5	9	MM					
	sV	40.29	100	9	MM					
	sN	28.45	53	6	MM					
	zT	23.41	23	8	MM					
	zV	68.07	100	7	MM					
	zN	54.01	63.5	10	MM					
	sT	9.49	10	9	MX	18/16.7				
	sV	41.93	72.5	9	MX	88.6/100				
	sN	37.81	49.5	10	MX	67.5/100				
	zT	11.47	10.5	8	MX	17.3/13.4	sT-zT	W =380	> 0.05	ns
	zV	45.87	79.5	7	MX	93.1/100	sV-zV	W =306	> 0.05	ns
	zN	25.7	33	9	MX	63.6/100	sN-zN	W =387.5	> 0.05	ns

Wilcoxon tests for the NS group indicated significantly more voicing for /z/ than /s/ in the three contexts: when a voiceless consonant ($W = 120$, $p < 0.05$), a vowel ($W= 45$, $p < 0.001$) and a nasal followed ($W= 51$, $p < 0.001$). That is, Native Speakers show a significant difference in amount of vocal fold vibration between /s/ and /z/ in all contexts.

Group FS1 had just one speaker who sounded near-native to English ears. Her results are similar to those for the NS group. FS1 shows differences in voicing between /s/ and /z/ in a voiceless context ($W= 17$, $p < 0.05$) and before a vowel ($W= 9$, $p < 0.01$), but the differences do not reach significance before a nasal ($W= 34$, $p > 0.05$).

Wilcoxon tests for the FS2 group showed no significant differences in voicing between /s/ and /z/ when followed by a voiceless consonant ($W = 1293$, $p > 0.05$) or a nasal ($W = 1050$, $p > 0.05$) – the latter most likely due to large variability in the data –

but /z/ exhibited significantly more voicing than /s/ when followed by a vowel (W = 343, p > 0.0001).

Wilcoxon tests for the FS3 group examining voicing differences between /s/ and /z/ in the voiceless (_T) and voiced (_V, _N) contexts showed no significant differences in any of the three contexts (p >0.05). The median values indicate that both /s/ and /z/ were voiceless when a voiceless consonant followed and fully voiced when a voiceless sound followed.

4.2 Results for preceding vowel duration

Table 3. Vowel duration ratios before voiced and voiceless fricatives in final position and grouping of the speakers according to these ratios. Ratios > 1.3 are indicated in grey.

		_fricN	_fricV	_fricT
NS	J	1,48	1,37	1,3
	AM	1,35	2,22	1,22
FS1	AB	1,33	1,21	1,39
	MX	1,31	1,11	1,31
	JUD	1,37	1,60	1,36
FS2	MM	1,1	1,48	1,02
	ST	1,15	1,54	0,88
FS3	GR	1,09	1,24	1,12
	CL	0,82	1,22	1,04
	ER	1,05	1,09	0,97
	JC	1,04	1,16	1,01

Examination of the vowel duration differences before final /s/ and /z/ for the individual speakers again showed large variation between non-native speakers and suggested, overall, 4 different patterns (see Tables 3 and 4). First, the ratio of vowel duration before voiced and voiceless consonants was calculated for each speaker. Second, the criterion to group speakers was whether the ratio was larger than or smaller than 1.3:1, the ratio typically quoted for native speakers. In addition, speakers showed a natural division between those who had a ratio smaller than 1.3 – that is, whose vowel duration did not clearly differ before voiced and voiceless fricatives – and those with a ratio larger than 1.3. Thus FS3 was defined as speakers who showed a voiced-to-voiceless consonant vowel duration ratio *smaller* than 1.3 in *all* contexts. FS2 as speakers with a vowel

duration ratio *larger* than 1.3 in at least *one* context. Finally, FS1 was defined as speakers with a ratio similar to Native Speakers (that is, ratios > 1.3 in at least two contexts).

The fact that the grouping of speakers according to consonant voicing (i.e., negative VOT) differs from the grouping for vowel duration, and that there is no overlap of speakers for FS1, FS3 and little for FS2 (only one speaker, ST) in the two dimensions, suggests that some non-native speakers seem to use VOT to cue consonant voicing while other speakers seem to use vowel duration to cue consonant voicing, but only a few speakers appear to use both cues.

The results for individual speakers are shown in Figs. 9-19 and Table 4. First, the differences in vowel duration before /s/ and /z/ in the various contexts will be discussed for the 4 groups of speakers.

J and AM compose the NS group, who serve as control. The individual results for NS are presented in Figs. 9 and 10 which show vowel length in ms (Y axis) when the fricative is followed by a voiced consonant (_fricN) (e.g. price never, prize money), a vowel (_fricV) (e.g. price of, prize is), and a voiceless consonant (_fricT) (e.g. price could, prize should).

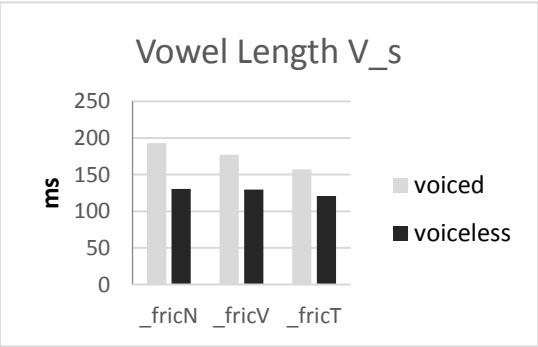

Figure 9. Mean vowel duration in ms for Native Speaker J.

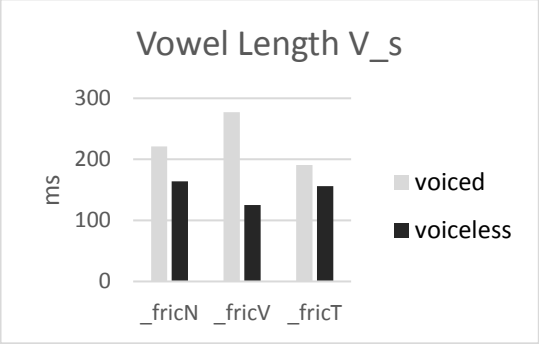

Figure 10. Mean vowel duration in ms for Native Speaker AM.

As expected, both native speakers make a difference in vowel duration before /s/ and /z/ (Figs. 9 and 10), such that vowels are shorter before /s/ than /z/. As mentioned above, this

group shows a vowel duration ratio larger than 1.30:1 in all contexts (except when followed by a voiceless consonant for speaker AM).

Statistical test were used to validate these observations. Because the data were normally distributed, one-factor ANOVAs were used. In order to examine if vowel duration differed before voiced and voiceless fricatives in the three contexts, one-factor ANOVAs were performed for each group of speakers separately. For Native Speakers, the results showed significantly longer vowels before /z/ than /s/ in the three contexts, when a voiceless consonant (mean= 173.9ms vs 138.9 ms, respectively; $F_{(1, 37)} = 15.1$, $p < 0.001$), a vowel (mean= 216.7ms vs 126.3 ms, $F_{(1, 37)} = 37.82$, $p < 0.001$) and a nasal followed (mean= 207.0ms vs 147.1ms $F_{(1, 38)} = 26.63$, $p < 0.001$), with a significantly smaller difference in the voiceless than the vowel and nasal contexts, which do not differ between them.

The second group, FS1, composed by those speakers who behaved native-like (AB, MX, JUD), showed the expected ratio in at least two contexts. This group of speakers showed a smaller difference in vowel duration in _fricV than in other contexts (Figs. 11 and 12), except for JUD (Fig. 13), who performed native-like in all contexts.

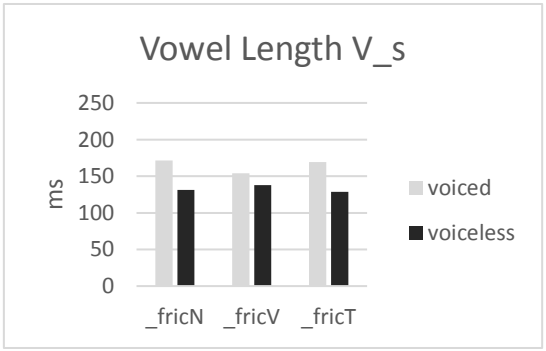

Figure 11. Mean vowel duration in ms for Catalan Speaker MX.

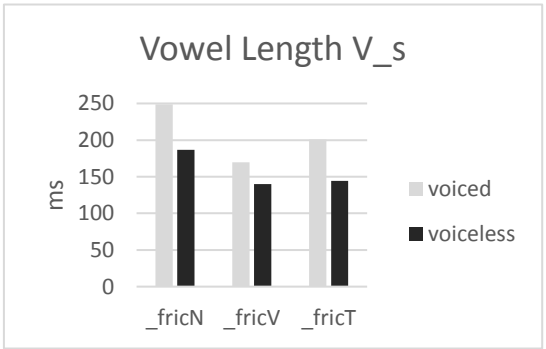

Figure 12. Mean vowel duration in ms for Catalan Speaker AB.

The statistical tests indicated that the FS1 group showed differences in vowel duration comparable to those of native speakers. The results of the ANOVA for FS1 speakers also

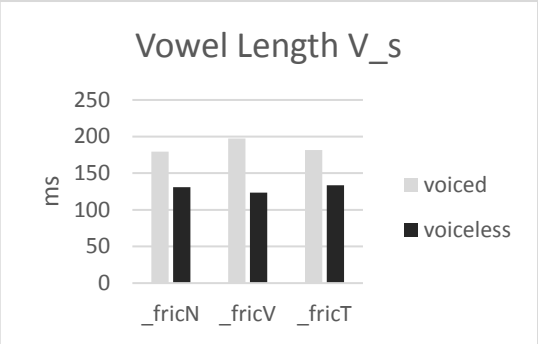

Figure 13. Mean vowel duration in ms for Catalan Speaker JUD.

showed longer vowels before /z/ than /s/ when a voiceless consonant (mean= 179.6ms vs141.9ms, respectively, ($F_{(1, 52)}$ = 18.63, p < 0.001), a vowel (mean= 173.92ms vs 133.24ms, $F_{(1, 51)}$ = 19.40, p < 0.001) and a nasal followed (mean= 197.98ms vs 148.72ms, $F_{(1, 56)}$ = 20.29, p < 0.001). As for Native Speakers, differences in vowel duration before /s/ and /z/ were smaller before voiceless consonants than before vowels and nasals.

FS2, composed by two Catalan speakers (MM, ST), showed the expected vowel duration ratios in only one context. When the fricative was followed by a vowel, the ratio was larger than 1.3, but this was not the case when a N or a voiceless consonant followed.

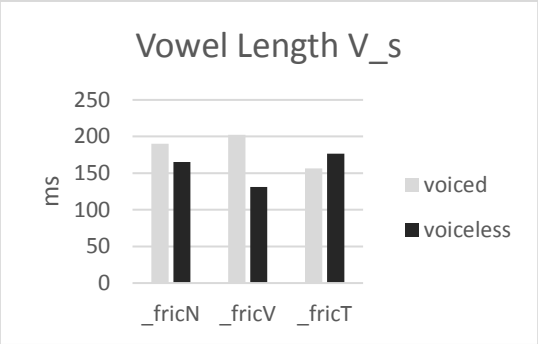

Figure 14. Mean vowel duration in ms for Catalan Speaker ST.

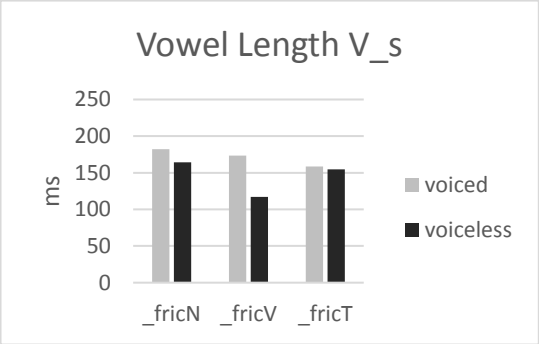

Figure 15. Mean vowel duration in ms for Catalan Speaker MM

This is partly corroborated by the statistical results. The one-factor ANOVAs for Group FS2 did not show significant differences in vowel duration before /z/ and /s/ when a voiceless consonant follows (mean= 155.29ms vs 163.70ms, $F_{(1, 35)}$ = 0.99, p>0.05) – suggesting Regressive Voicing Assimilation to [s] and no difference in vowel duration. This group, however, exhibited significantly longer vowels before /z/ than /s/ when a vowel (mean= 189.06ms and 124.33, respectively; $F_{(1, 33)}$ = 42.20, p < 0.001) and a nasal

followed (mean= 184.52ms and 162.05 ms, respectively; $F_{(1, 34)} = 15.56$, $p < 0.001$). This indicates that FS2 speakers make a difference in vowel duration before voiced and voiceless fricatives in two contexts (before a vowel and a nasal), although the extent of the difference is not native-like.

Finally, FS3 is composed by those speakers who showed a voiced-to-voiceless consonant vowel duration ratio *smaller* than 1.3 in *all* contexts (ER, JC, GR, CL). Generally, this group tended to perform better in a _fricV context (ratios ranging from 1.09 to 1.24). However, they do not make a systematic difference in vowel duration in the other two contexts (Figs. 16-19).

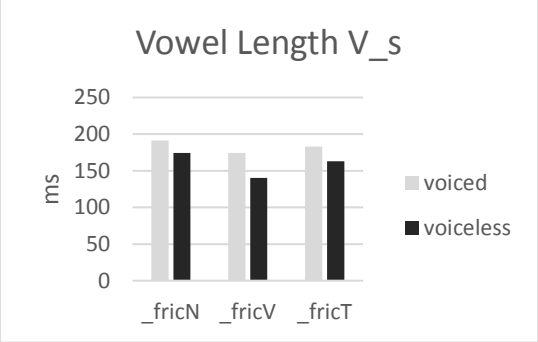

Figure 16. Mean vowel duration in ms for Catalan Speaker GR

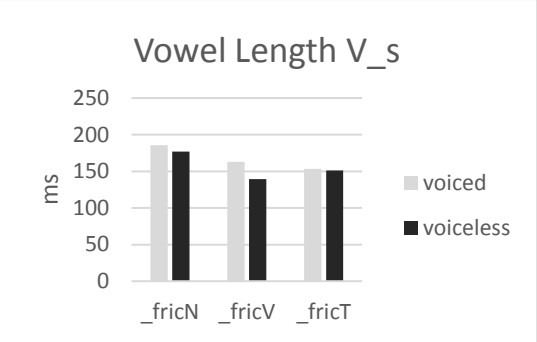

Figure 17. Mean vowel duration in ms for Catalan Speaker JC

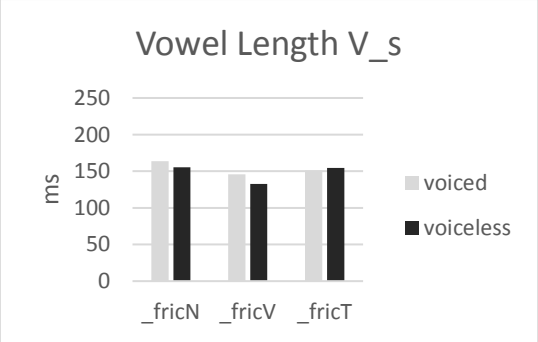

Figure 18. Mean vowel duration in ms for Catalan Speaker ER

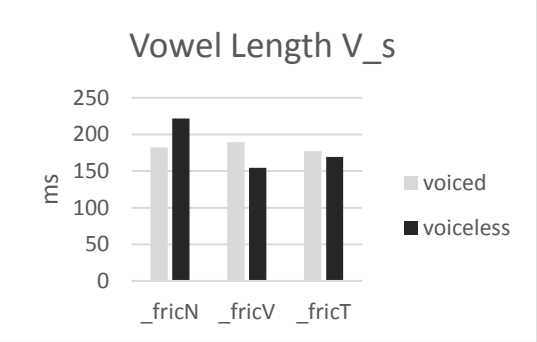

Figure 19. Mean vowel duration in ms for Catalan Speaker JUD

The statistical tests indicated that group FS3 only exhibited significantly longer vowels before /z/ (mean= 167.37ms) than /s/ (mean= 141.72ms) when a vowel followed ($F_{(1, 75)}$

= 24.33, p < 0.001). The difference in vowel duration, however, was not as large as the one for Native Speakers.

In the other two contexts (when a voiceless or a voiced consonant followed), the differences in vowel duration before voiced and voiceless fricatives did not reach significance. Indeed, the mean values for Group 3 suggest that the duration of the vowel is not clipped before voiceless consonants (mean vowel duration before /z/= 162ms; before /s/ 166.93ms in a voiceless context (_fricT); mean= 178,19ms vs 184.61ms before voiced and voiceless fricative in a nasal context (_fricN)). In sum, the results for this group of speakers suggest that they do not make a difference in vowel duration to cue voicing in the following obstruent except between vowels, and when they do make a difference, the extent of those differences is much smaller than in Native speakers.

Table 4. Descriptive statistics and results of ANOVAs comparing vowel duration before /s/ and /z/ in sT-zT, sV-zV, and sN-zN sequences for the different groups of speakers. *= p<0.05, **= p<0.01, ***= p<0.001

Group	context	Mean vowel duration (ms)	Pooled mean vowel duration (ms)	N	Speaker	Voiced-to-voiceless C ratio	Comparisons	ANOVAs	p-value	significance
Native Speakers	_sT	120.9		10	J					
	_sV	129.79		10	J					
	_sN	130.3		10	J					
	_zT	157.22		10	J	1.3				
	_zV	179.07		9	J	1.37				
	_zN	192.96		10	J	1.48				
	_sT	158.61	139.75	9	AM					
	_sV	122.37	126.08	9	AM					
	_sN	163.97	147.13	10	AM					
	_zT	190.61	173.91	10	AM	1.22	sT-zT	$F_{(1, 37)}$ = 15.1	< 0.001	***
	_zV	273.1	226.08	6	AM	2.22	sV-zV	$F_{(1, 37)}$ =37.82	< 0.001	***
	_zN	221.11	207.03	10	AM	1.35	sN-zN	$F_{(1, 38)}$ = 26.63	< 0.001	***

Group	context	Mean vowel duration (ms)	Pooled mean vowel duration (ms)	N	Speaker	Voiced-to-voiceless C ratio	Comparisons	ANOVAs	Wilcoxon sum rank test	significance
FS1	_sT	159.52		10	AB					
	_sV	139.99		10	AB					
	_sN	188.08		9	AB					
	_zT	187.61		8	AB	1.39				
	_zV	165.91		9	AB	1.21				
	_zN	242.82		9	AB	1.33				
	_sT	135.16		9	JUD					
	_sV	123.54		10	JUD					
	_sN	130.91		10	JUD					
	_zT	181.66		10	JUD	1.36				
	_zV	199.1		8	JUD	1.60				
	_zN	185.73		10	JUD	1.37				
	_sT	129.21	141.29	9	MX					
	_sV	136.52	133.35	9	MX					
	_sN	131.3	150.09	10	MX					
	_zT	169.1	179.45	8	MX	1.31	sT-zT	$F_{(1, 52)} = 18.63$	< 0.001	***
	_zV	155.44	173.48	7	MX	1.11	sV-zV	$F_{(1, 51)} = 19.40$	< 0.001	***
	_zN	153.83	194.12	9	MX	1.31	sN-zN	$F_{(1, 56)} = 20.29$	< 0.001	***

Group	context	Mean vowel duration (ms)	Pooled mean vowel duration (ms)	N	Speaker	Voiced-to-voiceless C ratio	Comparisons	ANOVAs	p-value	significance
FS2	_sT	154.95		9	MM					
	_sV	117.36		9	MM					
	_sN	157.8		6	MM					
	_zT	155.2		8	MM	1.02				
	_zV	171.62		7	MM	1.48				
	_zN	182.02		10	MM	1.1				
	_sT	171.58	163.26	10	ST					
	_sV	131.3	124.33	9	ST					
	_sN	164.55	161.17	10	ST					
	_zT	155.37	155.28	10	ST	0.88	sT-zT	$F_{(1, 35)} = 0.99$	>0.05	ns

	_zV	201.27	186.44	10	ST	1.54	sV-zV	$F_{(1, 33)}$ =42.20	< 0.001	***
	_zN	187.02	184.52	10	ST	1.15	sN-zN	$F_{(1, 34)}$ = 15.56	< 0.001	***

Group	context	Mean vowel duration (ms)	Pooled mean vowel duration (ms)	N	Speaker	Voiced-to-voiceless C ratio	Comparisons	ANOVAs	p-value	significance
FS3	_sT	202.61		9	CL					
	_sV	154.48		10	CL					
	_sN	221.86		10	CL					
	_zT	177.15		10	CL	1.04				
	_zV	189.45		8	CL	1.22				
	_zN	182.43		10	CL	0.82				
	_sT	154.53		10	ER					
	_sV	132.75		10	ER					
	_sN	155.21		10	ER					
	_zT	152		9	ER	0.97				
	_zV	145.75		10	ER	1.09				
	_zN	165.35		9	ER	1.05				
	_sT	151.11		10	JC					
	_sV	139.39		10	JC					
	_sN	176.75		10	JC					
	_zT	153.27		10	JC	1.01				
	_zV	162.97		10	JC	1.16				
	_zN	185.51		10	JC	1.04				
	_sT	163.06	167.82	10	GR					
	_sV	140.29	141.72	10	GR					
	_sN	174.14	181.99	10	GR					
	_zT	166.31	162.18	6	GR	1.12	sT-zT	$F_{(1, 72)}$ = 0.40	>0.05	ns
	_zV	176.65	168.70	9	GR	1.24	sV-zV	$F_{(1, 75)}$ =24.33	< 0.001	***
	_zN	191.18	181.11	10	GR	1.09	sN-zN	$F_{(1, 77)}$ = 0.003	>0.05	ns

5. Discussion and Conclusions

This paper focused on the production of the voicing contrast in final fricatives by advanced Catalan speakers of English. As noted in the introduction, although the voicing contrast in fricatives exists in Catalan word-medially, it does not occur in final position where Final Obstruent Devoicing and Regressive Voicing Assimilation take place. In order to cue voicing distinctions in English, the subjects were expected to utilize two parameters not used in their native language in final position: fricative voicing and preceding vowel duration. The fact that when grouping the subjects for their production of the two dimensions – fricative voicing and preceding vowel duration – there was no overlap of speakers for FS1, FS2 (except for speaker ST) and FS3, suggests that some subjects used one cue (e.g. vowel duration difference) over the other (e.g. vocal fold vibration) to cue consonant voicing and other subjects did the reverse. Only a few speakers used both (e.g. ST, JUD). Other speakers did not seem to master either of the two cues. Thus, we can conclude that RVA is an L1 process persistent in advanced L2 English speakers.

As we have seen, an advanced level of English does not translate into a good production of final fricative voicing contrasts. In fact, regarding fricative voicing, only one subject behaved native-like, showing the same amount of voicing and devoicing as the native speakers. FS2 did not exhibit a difference in voicing between /s/ and /z/ in a voiceless context (which indicates RVA), but /s/ and /z/ differed in voicing in a voiced context (although it did not reach significance in a nasal context most likely due to large variability in the data), with higher voicing values than those exhibited by native speakers. FS3 showed clear Regressive Voicing Assimilation in English.

There also seems to be variation in the mastering of vowel duration differences before a voiced and a voiceless fricative. Only three Catalan speakers, group FS1, showed the vowel duration ratio expected in at least two contexts. The extent of the vowel duration differences were comparable to those of Native Speakers. FS2 showed significantly longer vowels before /z/ than /s/ when followed by a nasal or a vowel but the extent of those differences was smaller than for English speakers. FS2, on the other hand, did not exhibit significant differences when followed by a voiceless consonant, which suggests that there is Regressive Voicing Assimilation to [s] in this context. Finally, the results for FS3 only showed significant differences in vowel duration, again of a smaller size than NS's, when the fricative was followed by a vowel. In the other contexts, there was no difference in vowel duration, which suggests RVA.

RVA is a topic that has been extensively researched. However, it seems that most of the work has been done on obstruents in final-position. It would be interesting to further this research by examining RVA when the obstruent appears in other contexts, for example, in initial position, e.g., the snail.

7. References

Cebrian, J. (2000). Transferability and productivity of L1 rules in Catalan-English interlanguage. *SSLA, 22*, 1-26.

Cuartero, N. (2001). *Voicing Assimilation in Catalan and English.*

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*, 315–330.

Flege, J. E., and Davidian, R. D. (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics*, *5*, 323–347.

Fullana, N. and MacKay, Ian R. A. (2008) "'I Said Made (Mate?)': Catalan/Spanish Bilinguals' Production of English Word-Final Obstruents". *Canadian Acoustics / Acoustique Canadienne*, 36, (3): 118-119.

Fullana, N. and Mora, J. C. (2009) Production and perception of voicing contrasts in English word-final obstruents: Assessing the effects of experience and starting age. In M. A. Watkins, A. S. Rauber, & B. O. Baptista (Eds.) *Recent Research in Second Language Phonetics/Phonology: Perception and Production.* (pp. 97-117). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Greenberg, J. H. (1966). *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.

Grijzenhout, J. (2000). Voicing and devoicing in English, German, and Dutch; evidence for domain-specific identity constraints. *Theory des Lexikons; Arbeiten des Sonderforschungsbereichs, 282*, 1-22.

Grijzenhout, J. and Krämer, M. (1998*). Final devoicing and voicing assimilation in Dutch derivation and cliticization*. Düsseldorf: Univ. Düsseldorf, Sonderforschungsbereich 282.

Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In *Speech production and speech modelling* (pp. 69-92). Springer Netherlands.

Lado R. (1957). *Linguistics across cultures.* Ann Arbor: University of Michigan Press.

Major, R. C. and Faudree, M. C. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition*, *18*, 69–90.

Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The Production of Speech*, 189–216. New York: Springer-Verlag.

Rothenberg, M. (1968). *The breath-stream dynamics of simple-released-plosive production*. Basel: Karger.

Solé, M. (2014). The perception of voice-initiating gestures. *Laboratory Phonology*, 5(1), 37-68. doi:10.1515/lp-2014-0003

Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73. 1322–1336.

Westbury, J. R., & P.A. Keating. (1986). On the naturalness of stop consonant voicing.*Journal of Linguistics* 22. 145–166. Basel: Karger.

8. Appendices

8.1. English level questionnaire

Questionnaire:

Please answer a few questions about yourself.

1. Have you lived in a country where English is spoken? How long?

.....

2. How would you rate your overall current **exposure to English? (Circle one)**

1= no contact ←----------→ 7= everyday, extended conversation

1 2 3 4 5 6 7

3. How often do you watch TV or listen to the radio in English?

1= never←……..→7= daily

1 2 3 4 5 6 7

4. Have you taken any courses in English? If yes, for how long?

……

5. How would you rate you level of spoken English?

1 2 3 4 5 6

Beginner Advanced…beginner…Low Intermediate…Intermediate…High-Intermediate… …Advanced

6. Do you speak English at home?

Yes No

8.2 Task Sentences

Subjects were asked to read a document similar to this one, where there were no titles or indications of where RVA is expected, in order not to influence their speech.

Vlss→Vd

1. I think it's a nice movie.
   I think it was nice of you to come today.
   I think that's a nice purse.

2. The man told the girls the price never varies.
   The man told the girls the price of the box.
   The man told the girls the price could vary.

Vd→ Vlss

1. The prize money bought their ticket to Japan.
   The prize is a trip to Japan.
   The prize she won was a trip to Japan.

2. The smaller size never fits me.
   The smaller size is too big for her.
   The smaller size T-shirt fits her.

8.3 Figures

8.3.1 Figures for fricative voicing values in %

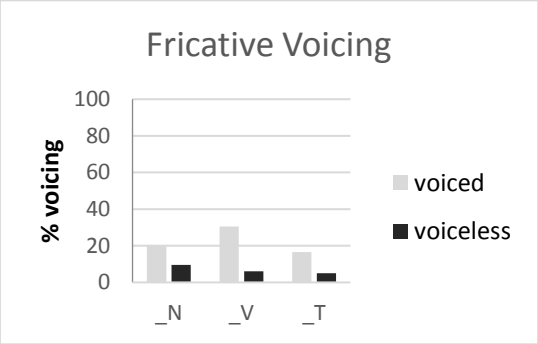

Figure 1. Mean % of fricative voicing for Native Speaker J.

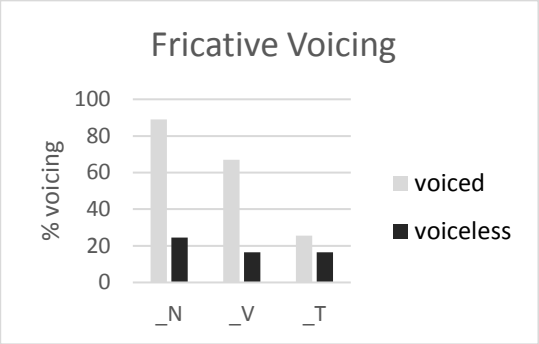

Figure 2. Mean % of fricative voicing for Native Speaker AM.

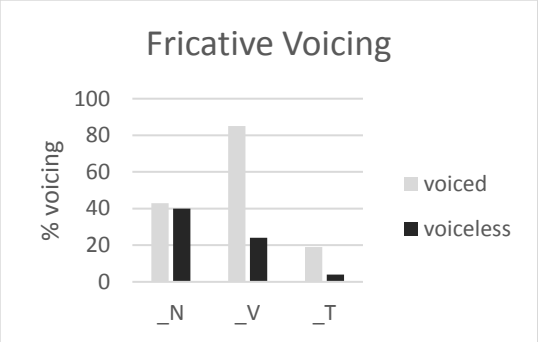

Figure 3. Mean % of fricative voicing for Catalan Speaker ER.

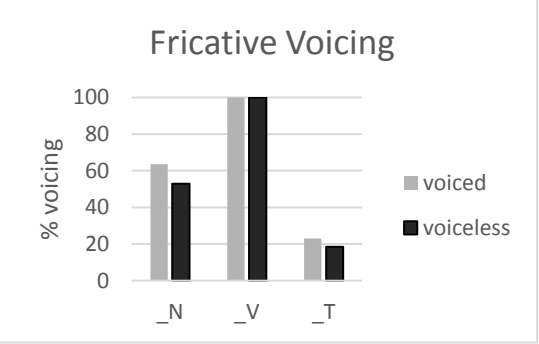

Figure 4. Mean % of fricative voicing for Catalan Speaker MM.

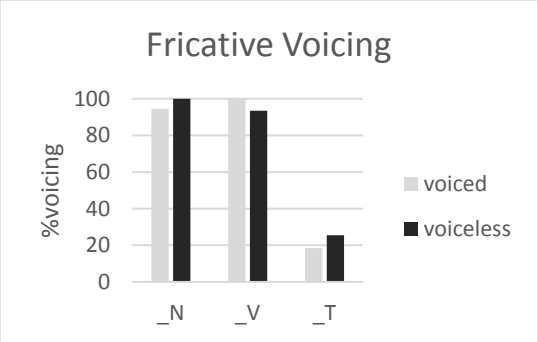

Figure 5. Mean % of fricative voicing for Catalan Speaker AB.

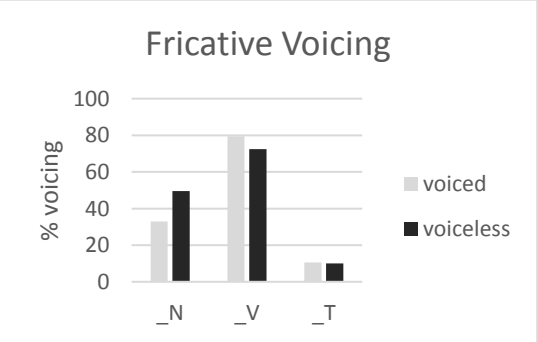

Figure 6. Mean % of fricative voicing for Catalan Speaker MX.

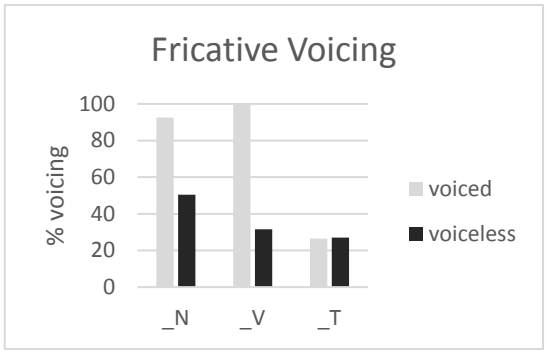

Figure 7. Mean % of fricative voicing for Catalan Speaker JUD.

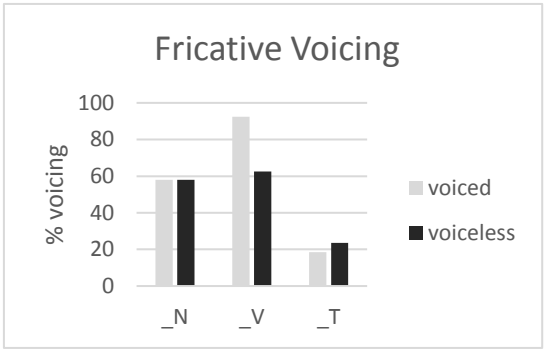

Figure 8. Mean % of fricative voicing for Catalan Speaker ST.

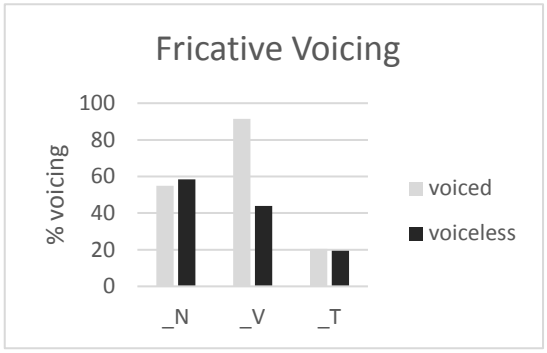

Figure 9.. Mean % of fricative voicing for Catalan Speaker CL.

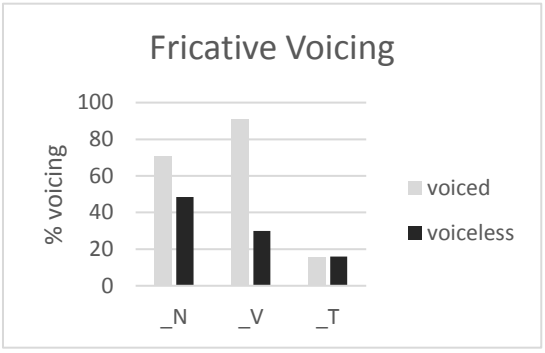

Figure 10. Mean % of fricative voicing for Catalan Speaker GR.

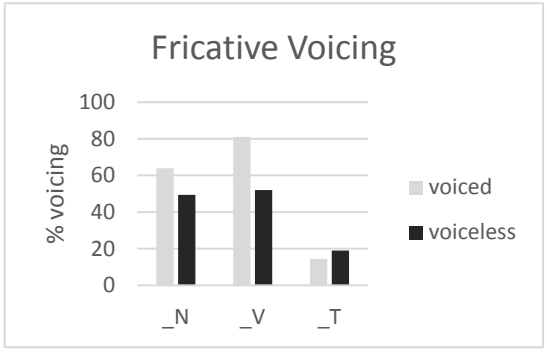

Figure 11. Mean % of fricative voicing for Catalan Speaker JC.

8.3.2. Figures for fricative absolute values

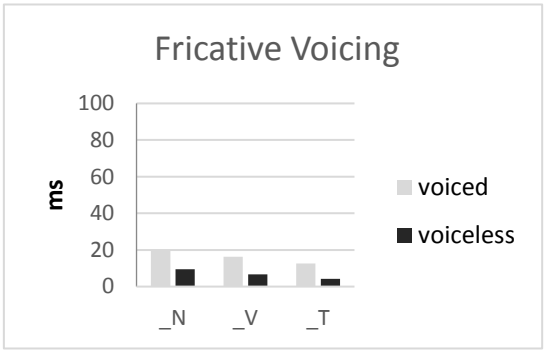

Figure 12. Mean fricative duration in ms for Native Speaker J.

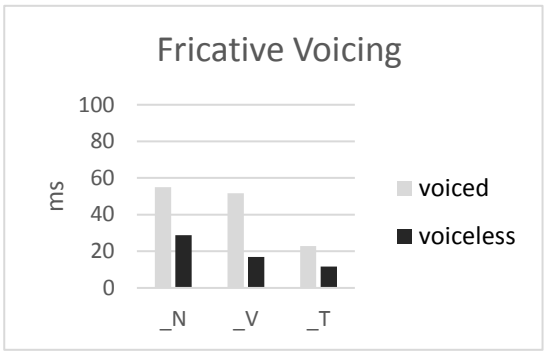

Figure 13. Mean fricative duration in ms for Native Speaker AM.

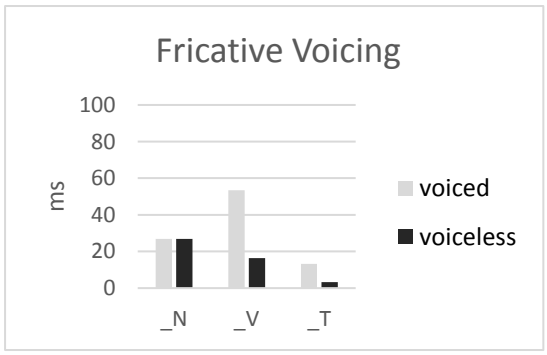

Figure 14. Mean fricative duration in ms for Catalan Speaker ER.

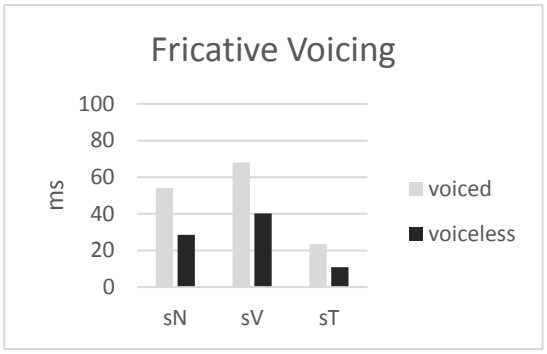

Figure 15. Mean fricative duration in ms for Catalan Speaker MM.

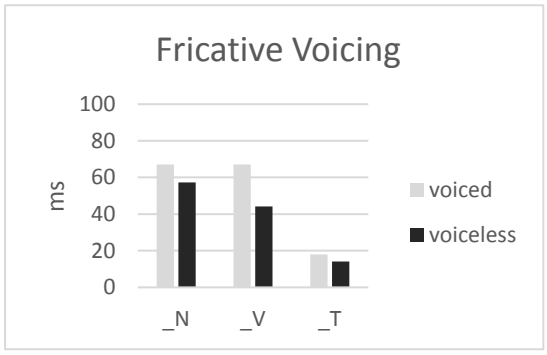

Figure 16. Mean fricative duration in ms for Catalan Speaker AB.

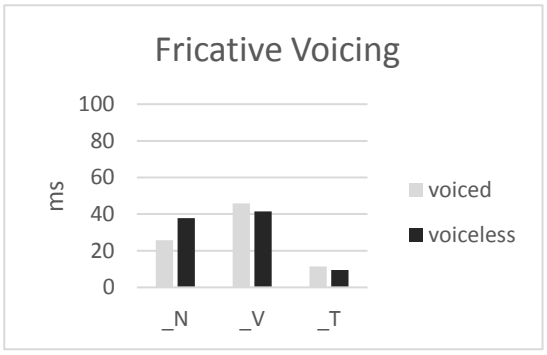

Figure 17. Mean fricative duration in ms for Catalan Speaker MX.

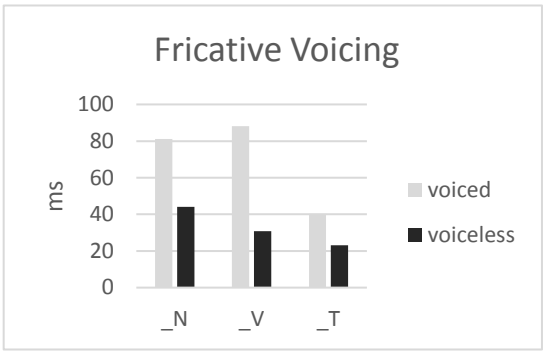

Figure 18. Mean fricative duration in ms for Catalan Speaker JUD.

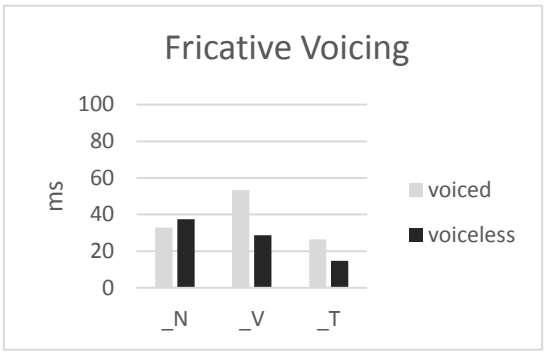

Figure 19. Mean fricative duration in ms for Catalan Speaker ST.

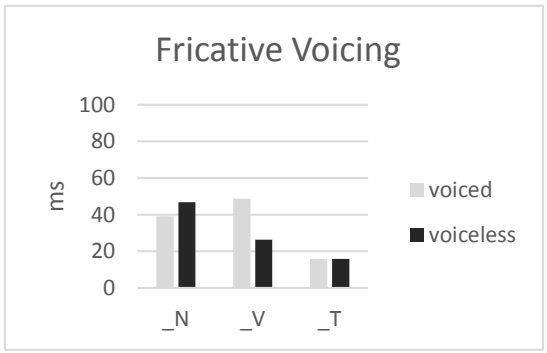

Figure 20. Mean fricative duration in ms for Catalan Speaker CL.

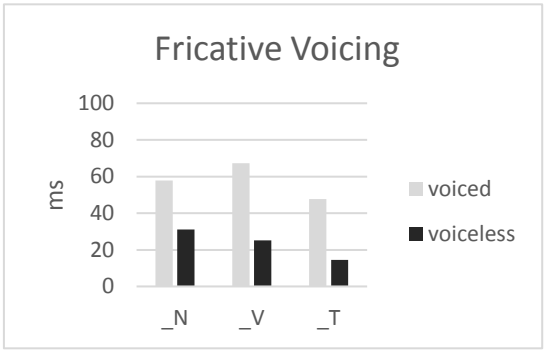

Figure 21. Mean fricative duration in ms for Catalan Speaker GR.

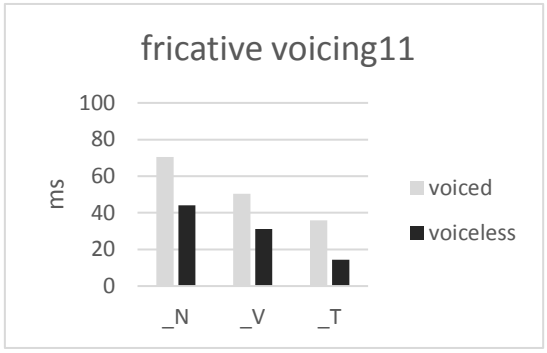

Figure 22. Mean fricative duration in ms for Catalan Speaker JC.