VOICING ASSIMILATION IN STOP SEQUENCES IN CATALAN AND ENGLISH

Nestor Cuartero
Universitat Autònoma de Barcelona, Spain

ABSTRACT

The aim of the experiment was to determine the direction and temporal extent of voice assimilation in Catalan and English VOICELESS - VOICED stop sequences, and the effect of speaking rate on the process. The material consisted of two-word sequences that contained consecutive stops across word boundaries. The results showed that, in Catalan, C1 was realized as voiced when followed by a voiced stop. In English, on the other hand, C1 was implemented with no vocal fold vibration, and C2 was devoiced. The results also showed that, in Catalan, the process applied in a categorical fashion, so that the closure for C1 was implemented as voiced. English exhibited variability in the degree of voicing during the closure of C2. Degree of voicing in C1 did not co-vary with speaking rate in Catalan, and C1 was voiced independent of speaking rate. In English, there was no effect of speech rate, either.

1. INTRODUCTION

The experiment intended to analyze the process of voicing assimilation in Catalan and English. More specifically, it aimed at observing the process of assimilation when two stops that have a different phonological specification for voicing co-occur across word boundaries within a phonological phrase. The second aim was to analyze the influence of speaking rate on voicing assimilation in both languages.

In summary, the experiment had two main objectives: firstly, it was designed to observe the direction and temporal extent of voicing assimilation across word boundaries in Catalan and English. Secondly, we wanted to study the effect of varying speaking rate on assimilation of voicing across word boundaries.

It has been observed that the process of voicing assimilation in obstruents has a different direction and extent in Catalan and English. This suggests that phonological units with the same feature specification are phonetically implemented differently in both languages. Thus, a sequence like Catalan *pot dur* /ˈpɔt ˈdu/ results in complete regressive assimilation of the voicing feature by C1 ([ˈpɔd du]). Conversely, in the English sequence *back door* /bæk ˈdɔː/ ([bæk ˈd̥ɔː]) C1 remains voiceless and C2 is partially devoiced [1]. The present study aims at studying these cross-linguistic differences in the phonetic implementation of phonological units with the same feature specification with original data, and to review how theories of phonetic implementation model such differences.

We also intend to analyze the influence of speaking rate on voicing assimilation in the two languages. Various studies [2, 8] have shown that the overall tempo of performance affects assimilatory processes and connected speech processes in general. More concretely, there seems to be a trend towards more assimilation in fast speaking rates, and less assimilation in slow careful rates of speech. The tendency for assimilation to co-vary with speech rate shows that assimilation is a gradual process rather than a discrete phenomenon. This poses an additional problem for phonological models that are based on discrete phonological units, where phonological processes like voicing assimilation are represented as discrete changes. Thus, we wanted to test whether the patterns of co-ordination of oral and glottal gestures -if there are any- are categorical or gradient across different speaking rates.

2. METHOD

2.1. Material

The material used in the experiment consisted of meaningful two-word sequences that contained consecutive alveolar or velar stops across word boundaries within a phonological phrase. In the test sequences, the first word ended in an underlying voiceless alveolar/velar stop (e.g., Catalan *pot gal*, English *hot girl*). The second word began with an underlying voiced alveolar/velar stop (e.g., Catalan *sac dolç*, English *back door*). In the control sequences, the final sound of the first word and the first sound in the second word were stops with the same phonological specification for voicing.

The number of syllables and the accentual pattern of the word sequences were kept constant. The sequences contained two monosyllabic words with the tonic stress falling on the second word. The material was read twice by each speaker at three different speaking rates.

The two-word sequences were inserted in the frame sentence ‘digui ___ un cop’ for Catalan, and ‘say ___ twice’ for English. The test sequences included in the stimuli together with the control sequences are listed in Figure 1 below. The same words were used in different combinations when possible. The slightly different frequency of occurrence and predictability of the final combinations may have affected the results.

Frame sentence		digui ___ un cop	say ___ again
Test sequences	/t#g/	pot gal	hot girl
	/k#d/	sac dolç	back door
Control sequences	/t#k/	tot coix	hot cop
	/k#t/	sac tort	Czech town
	/d#g/	fred gal	sad girl
	/g#d/	mag dolç	big door

Figure 1. Stimuli for the Catalan and English speakers

The grammatical status of the sequences used in the experiment was a factor that could have an influence on the results. Therefore, the syntactic status of the two-word sequences was kept constant: all of them were noun phrases consisting of a noun and a modifying adjective. The natural syntactic order of these two elements in the two languages was respected: thus, the adjective followed the noun it modified in the Catalan utterances, whereas the modifying adjective preceded the head noun in the English sequences.

Only consonants whose articulatory constriction involved tongue-palate contact –and were consequently observable in the EPG frames – were used. This means that bilabial stops were not included. Homorganic stop sequences and sequences that showed complete assimilation of place of articulation (e.g. Catalan 'pot gal' /ˈpɔt ɡal/ [ˈpɔɡɡal]; English 'hot girl') were not analyzed in this experiment because it was impossible to identify the boundary between the two consonants in the EPG frames.

2.2. Subjects

Four adult subjects participated in the present experiment, two native speakers of the eastern variety of Standard Catalan – MJ, DA-, and two native speakers of English, labeled AL and ME. MJ and DA were born and bred in the Barcelona and Tarragona areas respectively. They both lived permanently in Catalonia and used Catalan for all their daily linguistic exchange. As for the English subjects, AL spoke the variety of British English known as RP, whereas ME spoke General American. The dialectal difference was assumed to be irrelevant for the experiment, since neither AmE nor BrE have been reported to differ in the realization of stop sequences in terms of voicing. All the subjects were linguistically trained. No subject reported a history of speech or hearing disorders.

2.3. Experimental procedure

Subjects were asked to sit in front of a Samtron 486 computer at the UAB Phonetics Laboratory in Bellaterra. The lights were turned off in order to reduce external noise to the minimum. The electroglottograph and palatograph were adjusted to each of the subjects, who were then asked to read the test and control sequences inserted in a frame sentence.

All the utterances were stored in a 486 computer. Electropalatography (Reading EPG3 System) was used to trace the movement of the oral articulators. Electroglottography (EGG) was used to trace the activity of the vocal folds during the production of the stop sequences. EGG traces were obtained with an electroglottograph EG 830, F-J Electronics.

The measurements were carried out with the help of Reading EPG3 software, which provides a simultaneous multi-channel representation of the EPG, the EGG and the acoustic waveform. The duration measurements were double-checked by means of spectrograms (Computerized Speech Lab, Kay Elemetrics).

Each sequence was read twice at different speaking rates. Firstly, the subjects were asked to read the sequences at a normal speed, as if they were talking to a friend. After that, they were asked to read the sequences more slowly, as if they were talking in a very formal context to a formal audience. Finally, the subjects read the sequences as fast as possible. In a pilot experiment, the subjects were asked to read the sequences fast, and then as fast as they could. However, it was found that none of the speakers could produce a distinction between a "fast" and a "fastest" speaking rate. Thus, three speaking rates were obtained from each subject, which were labeled 'slow', 'normal' and 'fast'.

2.4. Analysis procedure

The experiment had both a between-subjects and a within-subjects design. We wanted to test whether there are cross-linguistic differences in the implementation of voicing in stop sequences. Additionally, we intended to measure the effect of speaking rate on voicing assimilation in the two languages studied, namely English and Catalan. Voicing during the consonantal closure was measured. Voicing was considered to be actual vocal fold vibration as reflected in the laryngograph (Lx). Since in English other cues are relevant to voicing, stop closure duration and Voice Onset Time were also measured.

The EPG signal and the waveform were used to measure the duration of the frame sentence, which served as an index of speaking rate (SR). The EPG contacts were used to determine the onset and offset of the articulatory constriction for C1 and C2 in the stop sequences that were the object of this study. A frame-by-frame representation of the articulation of the stop sequences in each utterance was obtained with the Reading EPG3 software. Thus, the tongue-palate contacts at intervals of 10 milliseconds could be observed.

The following points were identified in the frame-by-frame analysis of the stop sequences [7]:

1) Approach to C1 (AC1): the onset of movement towards the articulation of C1.
2) Onset of C1 (OC1): the point in time at which the total constriction for the first consonant in the sequence began.
3) Release of C1 (RC1): the point at which the articulators started moving apart for the release of C1.
4) Approach to C2 (AC2): the onset of movement towards the articulation of C2.
5) Onset of C2 (OC2): the point in time at which the total constriction for the second consonant in the sequence began.
6) Release of C2 (RC2): the point at which the articulators started moving apart for the release of the second consonant in the sequence.
7) Point of Maximal Constriction (PMC): used to determine the onset of a consonant when no complete constriction was present.

These points were then used to determine which portions of the stop sequences were voiced. The interval in the EGG waveform representing glottal excitation was measured in all utterances and the percentage of voicing in the assimilating consonant was then calculated.

As illustrated in Figure 2, in Catalan, the percentage of voicing in C1 was measured, whereas in English the percentage of voicelessness in C2 was measured. Voicing in C1 in Catalan was expressed in terms of negative percentages (i.e. voicing for C2 occurred before the articulatory closure for C2), and English values were expressed in terms of positive percentages (i.e., voicelessness for C1 extended during C2).

A second set of measurements that included VOT and duration of the stop sequence closure was also obtained in the English utterances. VOT was considered to be the duration of the interval between the release of C2 and onset of periodic glottal excitation for the following vowel. To measure the duration of the stop constriction, the duration of the interval from OC1 to RC2 was measured.

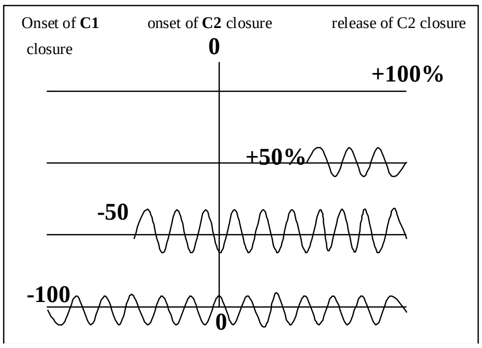

Figure 2. Schematic representation of the measurements of voicing in the stop sequences

3.RESULTS

3.1. Direction and extent

As regards the mixed sequences in English -/t#ɡ/ (‘hot girl’) and /’k#d’/ (‘back door’)- C1 was voiceless in all the utterances, so that there were no cases of regressive assimilation of voicing. As for the production of C2, Figure 3 shows that the English speakers exhibited two markedly distinct patterns. Speaker ME showed complete devoicing of C2 closure in most utterances - although there were three cases of partial assimilatory devoicing in the alveolar/velar sequence /t#ɡ/. Since there was no glottal activity during the production of the underlyingly voiced consonant, the percentage of voicing during the CC closure was the same for mixed and voiceless sequences. The question remains whether the distinction between the mixed and the voiceless sequences was maintained in some other way.

Speaker AL exhibited a voiced C2 closure and mechanical devoicing in the latter portion of the closure. This subject did not show devoicing of C2 but a canonical voiceless/voiced sequence with a synchrony between the glottal gesture and the supraglottal constriction (see Figure 3).

In sum, the voicing pattern in the mixed sequences for speaker ME resembles that of voiceless stops, reflecting the progressive devoicing predicted for English. Speaker AL does not show devoicing of C2, but a synchrony between oral and glottal articulations.

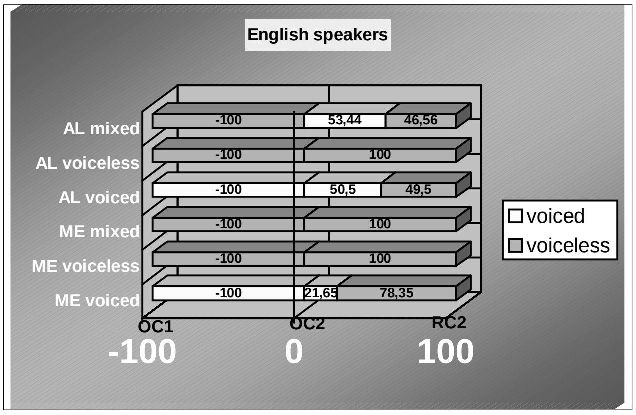

Figure 3. Average percentages of voicing and voicelessness in the English control and test sequences.

The voicing patterns of the Catalan subjects DA and MJ are presented in Figure 4. It can be seen that both speakers exhibited similar voicing timing patterns for the voiced control sequences and for the mixed test sequences. In the mixed sequences /t#ɡ/ (‘pot gal’) and /k#d/ (‘sac dolç’), C1 was fully voiced for speaker DA, who exhibited no difference in voicing patterns between mixed and voiced sequences. As in the voiced control sequences, mixed sequences for speaker MJ showed voicing in approximately the first half of C1 and mechanical devoicing during the second half of C1, which continued into C2. For speaker MJ, the presence of voicing in the first half of the closure of C1 in mixed sequences can only be a result of full regressive assimilation of C2 voicing plus mechanical devoicing.

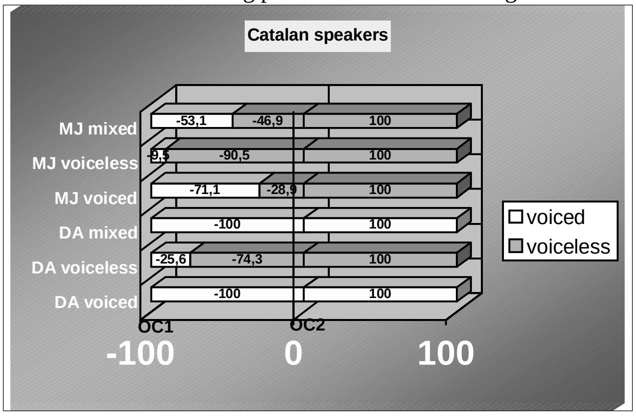

Figure 4. Average percentages of voicing and voicelessness in the Catalan control and test sequences.

3.2. Effect of speaking rate

In English, no co-variation of devoicing during the C2 closure and speaking rate was found. In ME’s mixed sequences, C2 presented no voicing during the closure in nearly all occurrences of /t#ɡ/ and /k#d/ at all speaking rates. Consequently, correlation coefficients for all the mixed sequences were not significant ($r=0.125$, $p<0.05$, $r2= 0.016$). However, three occurrences of /t#ɡ/ uttered at faster speaking rates presented partial assimilatory devoicing into C2. Speaker AL showed full voicing of C2 except for three tokens that exhibited partial devoicing.

Thus, both English speakers showed devoicing of C2. For one of the speakers, devoicing tended to be complete, although some cases were found of partial devoicing. However, whether progressive assimilation was complete or partial was not dependent on speaking rate. The other speaker also presented partial devoicing of C2 at faster speaking rates for one of the test sequences. The fact that there was no devoicing in most of the utterances produced by this speaker might be attributable to the experimental conditions, which may have caused the speaker to utter the sequences overcarefully and ‘unnaturally’. In other words, the sequences that presented partial devoicing tended to be those that were uttered at faster speaking rates, when the speaker did not monitor his speech as much.

All in all, there was inter-speaker variability in the degree of devoicing of C2, as shown by speakers ME and AL; however, there were instances of partial devoicing in both speakers independent of speaking rate. Hence, there was variability both within and across speakers.

In Catalan, the mixed sequences (/t#g/ and /k#d/) from both DA and MJ showed complete voicing of C1 at all speaking rates. Voicing assimilation applied categorically in Catalan. There were no cases of partial spreading of voicing. Sometimes, however, voicing assimilation was not present, mostly at slow rates, possibly because of other factors like speaking style and word boundary effects. Voiced control sequences (/d#g/, /g#d/) also showed two sporadic cases of voiceless C1, which may be interpreted as Final Obstruent Devoicing due to word boundary effects. The lack of voicing assimilation in four mixed sequences uttered by speaker MJ can be explained in similar terms. As regards speaking style, it has to be taken into account that only those sequences that did not show place assimilation were analyzed, so that we left out the most 'naturally' produced items. This might also explain the fact that some sequences presented a voiceless C1.

In brief, regressive assimilation of voicing did not co-vary with speaking rate in Catalan. Sequences of stops where C1 was phonologically voiceless and C2 phonologically voiced were mostly produced with a fully voiced C1, regardless of rate.

3.3. Duration of the sequence and VOT

Since no differences were found between mixed and voiceless sequences in the patterns of voicing during the closure for the English speaker ME, other cues in the data from the English speakers were analyzed in order to see whether the distinction between the two cluster types was maintained in some other way. The two variables analyzed were VOT and duration of the stop sequence closure.

Voiceless sequences had a larger voice lag than mixed and voiced sequences, a difference that was maintained across all speaking rates. Thus, the difference in VOT may be one of the cues to differentiate between mixed and voiceless cluster types.

The duration of the stop constriction was measured in order to find out whether differences in consonant duration could be used to cue the voicing feature. There was no significant difference between mixed and voiceless sequences as far as stop closure duration was concerned.

4. DISCUSSION

The results of our experiment showed that VOICELESS – VOICED stop sequences are phonetically implemented differently in Catalan and English. In Catalan, there is regressive assimilation of voicing, so that a phonologically voiceless C1 is realized as voiced when followed by a voiced stop.

The results also show that the assimilation process applied in a categorical fashion in Catalan, so that the closure for C1, the assimilating consonant, was implemented as voiced. The results obtained by Carbonell [3] and Dinnsen & Charles-Luce [5, 6] suggested that there was incomplete neutralization –i.e., partial voicing assimilation - in mixed sequences, so that they presented less voicing in the closure than their voiced counterparts. In our experiment, it was found that mixed sequences like /t#g/ and /k#d/ usually have the same degree of voicing in the closure of C1 as sequences of voiced stops like /d#g/ and /g#d/, although they may well differ in other factors such as VOT and consonant duration. However, some cases were found in our experiment where voicing assimilation did not apply in mixed sequences due to an intervening pause. This seems to agree with Cebrián's findings [4], who claims that voicing assimilation across word boundaries is not as consistent as Final Obstruent Devoicing.

English, on the other hand, exhibited great variability in the degree of voicing during the closure. The English subjects exhibited three voicing patterns during the closure for C2, and they did not show a consistent pattern of co-ordination of glottal and supraglottal gestures in VOICELESS – VOICED sequences, although the tendency seemed to be to devoice C2.

In sum, the conflict of glottal articulatory demands in the sequences studied is resolved differently in Catalan and English. Catalan speakers showed one common pattern of co-ordination of glottal and supraglottal gestures in such sequences: both subjects implemented the closure of C1 with full voicing if voicing assimilation was present. In English, voicing during the closure does not seem to be relevant, which explains the variability found between and within speakers.

4. CONCLUSION

The present experiment has shown that voicing assimilation is regressive and complete in Catalan. Furthermore, rate of speech does not affect degree of assimilation of voicing into C1. In other words, voicing assimilation is present regardless of speaking rate in this language. The phonetic implementation of VOICELESS-VOICED sequences in English is of a different nature. In English, there was no consistent pattern for the coordination of oral and glottal gestures during the closure of C2 in the stop sequences studied, and there was considerable variability in voicing during the C2 closure. The results showed that VOT was significantly different in voiceless and mixed sequences, but this was not the case for consonant duration. Thus, the relevant phonetic cue to differentiate sequences like /t#g/ and /t#k/ in English is not voicing during the closure of C2 or duration of the closure, but rather Voice Onset Time.

ACKNOWLEDGEMENTS

Part of this research was funded by grant DGICYT PB96-1158 awarded by the Ministry of Education and Culture of Spain.

REFERENCES

[1] Abercrombie, D. 1967. Elements of general phonetics. Edinburgh: Edinburgh University Press

[2] Barry, M. 1985. A palatographic study of connected speech processes. *Cambridge Papers in Phonetics and Experimental Linguistics* 4.

[3] Carbonell, J.F. 1992. Final devoicing and voicing assimilation in Catalan: an acoustic experiment. *Masters Dissertation,* University College London, London.

[4] Cebrián, J. 1997. Transfer of voicing processes in Catalan-English interlanguage. *Proceedings of EUROSLA'7,* Barcelona.

[5] Charles-Luce, J. and Dinnsen, D.A. 1987. A reanalysis of Catalan devoicing. *Journal of Phonetics*, 15, 187-190.

[6] Dinnsen, D. A. and Charles-Luce, J. 1984. Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, 12, 49-60.

[7] Gibbon, F., Hardcastle, W. and Nicolaidis, K. 1993. Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: a cross-linguistic investigation. *Language and Speech*, 36 (2, 3), 261-277.

[8] Kerswill, P. 1985. A socio-phonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge Papers in Phonetics and Experimental Linguistics* 4.