STOP-VOWEL COARTICULATION IN CYPRIOT GREEK

Eftychia Eftychiou

Phonetics Laboratory, University of Cambridge
ee227@cam.ac.uk

ABSTRACT

The present paper presents the results of an experimental investigation of the connected speech process of close-vowel lenition in Cypriot Greek (henceforth CG). The process appears to be gradient, with stops whose adjacent vowels have been elided being acoustically different from canonical word-final stops, indicating vowel-to-stop coarticulation. Finally, the study reveals two routes to lenition in CG; one involves a full consonant with a lenited vowel, and the other a lenited consonant with a full vowel, potentially signifying that the laryngeal setting is the same in both cases and that the different acoustic patterns are the result of supralaryngeal imprecision.

Keywords: vowel lenition, coarticulation, Cypriot Greek.

1. INTRODUCTION

Differences in vowel realisation are common in languages and studies on vowel lenition reveal differences in the acoustic properties of full and reduced vowels, such as their intensity, voicing, and spectral characteristics [1].

Like CG, Japanese lenites /i/ and /u/. Close vowels are particularly prone to lenition because the back pressure created by their articulation reduces the transglottal pressure drop, promoting devoicing and lenition [2]. Research on Japanese has revealed six types of vowel realisation, ranging from *fully voiced vowels* – produced with voicing during the entire duration of formant activity, and for which the energy for vocal resonance is provided by vocal fold vibration – to *elided vowels*, where no trace of formant structure or voicing for an expected vowel is found in the acoustic record [3].

Similarly, data on Standard Greek (henceforth SG) suggests that vowels go through various stages before reaching elision [4]. Voiceless consonants favour lenition, and results on vowels preceded by fricatives are particularly interesting because they suggest that, even in the most extreme stages of the vowel realisation continuum, there are traces of the vowel preserved in the preceding consonant, in the form of formants superimposed on the frication, which may cue the presence and identity of the vowel to listeners. This finding might explain the fact that SG listeners claimed to always hear the missing vowel.

This consonant-vowel coarticulation is present in both fricative-vowel and stop-vowel pairs and is manifested via various changes in the spectral characteristics of the consonants [5, 6, 7, 8]. Such changes may aid vowel identification, as suggested by previous studies (e.g. [4], [7]).

The aim of the study was to determine whether consonants that become word-final via vowel lenition differ acoustically from canonical word-final consonants. More specifically, it was predicted that coarticulatory traces of the lenited vowel can be found in the preceding consonant, as is the case in other languages.

2. METHOD

2.1. Materials

The material consisted of 7 tokens containing /u/ or /i/ utterance-finally, and preceded by /t/ (see Table 1). As this paper formed part of a larger study, stress was varied from the penultimate (PEN) to the antepenultimate (ANT) and to the pre-antepenultimate (PREANT) syllable for /u/-tokens, while /i/-tokens were investigated in penultimate and antepenultimate stress conditions only[1]. The non-clitic controls were single lexemes stressed on the same syllable as their clitic counterparts.

Table 1: Test tokens and their glosses

	/u/		/i/
Stress	Clitic	Non-clitic controls	Non-clitic
PREANT	[ɛfːˈɐnisi tu] *appearance its*		
ANT	[plɐstisˈini tu] *play-dough its*	[xˈɐsitu] *person's surname*	[ɐpˈoriti] *classified/secret*
PEN	[ɛfçˈi tu] *wish his*	[mɛsɐrˈitu] *person's surname*	[mɛsˈiti] *estate agent (gen)*

For the purposes of the present study, however, a token with a canonical word-final /t/ was used as the control: [bɪt] (*bit*). As there was great diversity in the material, it was not possible to use the same carrier-phrase, though all tokens were utterance-final. Five repetitions of each token were analysed, resulting in a corpus of 40 productions per speaker (8 tokens X 5 productions).

2.2. Speakers

Five male and three female native speakers of CG, aged 21-26, were recorded. All were university students at the time of the recording and had been living in the UK for at least a few months prior to the recordings. None of the speakers reported any hearing or speech defects, and all were naive to the objectives of the experiment.

2.3. Procedure

CG does not have a standard orthography, making it difficult to present the material to the speakers in written form. For want of a better technique, the sentences were presented in written form, but the speakers were not allowed to read them straight away. Instead, they were asked to read each sentence silently and then to look away from the paper. The experimenter would then ask a question to which they had to respond with the token in question. It was hoped that this technique would limit any possible effects of the non-conventional orthographic version of the sentence, as it required that speakers process the sentence before production.

2.4. Recordings

The recordings took place in a sound-treated room. The stimuli were recorded onto DAT tapes, digitised at a sampling rate of 16 kHz, and analysed using the PRAAT speech analysis package.

2.5. Analysis

(a) vowels: spectrograms were examined visually and vowels were deemed to be present when a periodic waveform corresponded to formant structure and vertical voicing striations in the spectrogram, together with an F0 trace;
(b) stop duration: calculated in milliseconds and measured from the last periodic pulse preceding silence in the waveform, to the next periodic pulse following burst in the waveform; when consonants were realised as approximants, their duration was measured from the point where amplitude dropped in the spectrogram to the point where amplitude rose to mark the start of the following vowel.
(c) F2: measured at the offset of the preceding vowel via DFT spectra taken over a 25ms window.

Statistical analysis was done in SPSS using ANOVA. For the F2 analysis, only tokens exhibiting fricated or elided vowels were used to examine whether consonants whose adjacent vowels had been lenited differ from canonical word-final consonants. Not all speakers produced fricated or elided vowels, so the results are based on data by 6 speakers (122 out of 210 tokens were fricated/elided). Similarly, it was not possible to control for how speakers produced their stops, so the set for some speakers does not include all types of stop realisation. As it was not possible to merely delete speakers, since that would create a very small data set on which to run statistical tests, the cell value for these speakers' productions was the mean of the other speakers' productions (there was a total of 10 empty cells).

For the F2 test, V_1F2 onset was the dependent variable and the factor was whatever follows the stop in the VCV sequence; in other words, what is varied is the underlying environment of the stop, i.e. whether there should be a vowel following or whether the stop is actually word-final. For the duration test, consonant duration was the dependent variable and the type of /t/-realisation was the factor.

3. RESULTS

3.1. Vowel Realisation

A range of vowel realisations occurred, which were subsequently classified into four categories according to their common attributes: *full vowels* (50.17% of cases), produced with clear voicing and formant-structure; *F1 vowels* (4.65%), for which there is sufficient energy in the voicing for only the first formant to be visible in the spectrogram; *fricated vowels* (40.14%), produced with no voicing but with their F2 and F3 following the aspiration of the preceding consonant (see Fig. 1); and *elided vowels* (5.01%), for which there is no visual evidence in the acoustic record (see Fig. 2).

Figure 1: Spectrogram of fricated V_2 in /itu/.

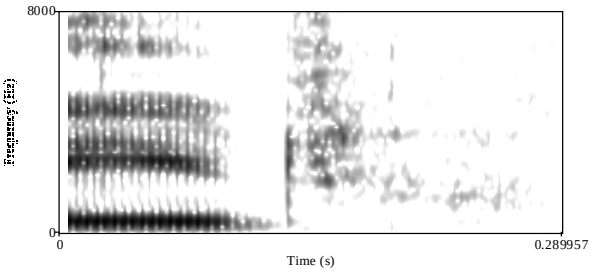
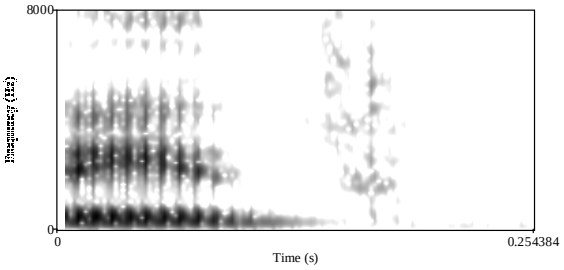


Figure 2: Spectrogram of elided V_2 in /itu/.

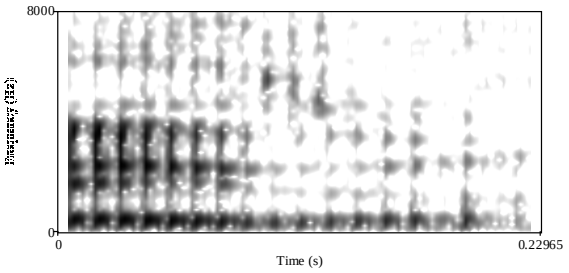


3.2. Stop Duration

Close inspection of waveforms and spectrograms showed that stop realisation is variable in CG. /t/ realisations varied from approximants (10.74% of cases), to voiced stops (30.37%), to partially voiced stops (47.66%), and to canonical stops (silence+burst) (10.28%). Interestingly, stop duration was at its minimum when the stop was lenited to an approximant (43ms) and at its maximum when produced canonically (82ms). There was a highly significant difference between the durations of the different types of /t/-realisation ($F(3, 15)=32.05$; $p<0.001$), with a very robust distinction between approximant and canonical duration as revealed by post-hoc tests ($p=0.001$).

In addition, stop and vowel realisation appear to be in complementary distribution. Full vowels were preserved more often when the consonant was lenited (e.g. Fig. 3), and elided more often when the target for the consonant was fully achieved; interestingly a canonical stop was never followed by a full vowel and an approximant was never followed by an elided vowel.

Figure 3: Spectrogram of approximant and full V_2 in /itu/.



3.3. V_1 F2 Offset

The difference in F2 offset for the vowel preceding /t/ was highly significant ($F(2, 14)=11.81$; $p=0.001$). The F2 value for /t(u)/ averaged at 1945 Hz, the value for /t(i)/ averaged at 1975 Hz and for /t#/ at 2112 Hz, and post-hoc tests revealed that there exist significant pair-wise comparisons only between /t(u)/ and /t#/ ($p<0.001$) and /t(i)/ and /t#/ ($p<0.001$).

4. DISCUSSION AND CONCLUSION

The study has presented evidence for the common occurrence of vowel lenition in CG and for the fact that realisations of vowels cluster around four different categories. This finding suggests that CG follows a pattern similar to other linguistic systems such as Japanese [3] and SG [4] with respect to the vowel realisation continuum, ranging from realisations of full vowels to realisations of elided vowels. As the analysis made it essential for productions to be classified in a category according to the majority of their features, arguing for discreteness is implausible. Therefore, unless as many categories are postulated as there are productions, the process can be interpreted as being gradient.

The gradience of the process might be a consequence of articulatory imprecision. According to Articulatory Phonology, the gestures for vowel and consonant overlap in time, with the resulting production manifesting properties of both to varying degrees according to the extent of overlap [9]. Although further articulatory studies, e.g. EPG, could provide another angle on this and investigate the claim that underlying gestures are never deleted or reduced to zero magnitude, the acoustic results comprise proof for the gestural interpretation; the measurements suggest that traces of the vowel are retained in the preceding consonant, indicating vowel-to-stop coarticulation in CG and suggesting that the underlying gesture

for the vowel is present in the consonant, even in the extreme stage of complete elision of the vowel.

This claim is borne out by the results of the experiment, as both /u/ and /i/ appear to leave coarticulatory traces on the preceding stop, to distinguish it from canonical word-final stops. This result is not in agreement with previous perceptual studies (e.g. [5]), which found high rates of identification of /i/ in the context of /t/ and low rates in the context of /u/, suggesting that /i/ leaves more coarticulatory traces on /t/ than /u/. The reason for this discrepancy, however, might lie in the fact that these studies used burst stimuli which did not contain any transition information. Since /t/ is a dental sound, its burst will exhibit high-frequency spectral peaks. It is possible that listeners are predisposed to restoring an elided /i/ in such contexts, unless information to the contrary is presented. Moreover, the fact that the position of the lips during /t/ production does not differ greatly from that of /i/, might also contribute to /i/ identification. It appears, then, that coarticulatory information in the transitions is very important for /u/ identification.

The extent of this coarticulation is not surprising given the nature of the sounds in question. According to the *principle of articulatory compatibility* [7, 8], the raised tongue positions required for the constriction of the alveolar consonants and for the articulation of close vowels are not contradictory, so the vowel can be anticipated by overlapping gestures.

The gradience of the process might result in facilitating perception; that is, the existence of vowel-consonant coarticulation might aid vowel identification. Experiments on English fricatives, for instance, have shown that spectral changes caused by overlapping gestures are sufficient cues for the identification of the vowel, thus establishing a link between production and perception [7]. Of course, this interpretation implies that at some point the process might be categorical, as listeners become more adept at picking out the appropriate perceptual cues even in the extreme cases of elided vowels, without needing to have contact with intermediate stages.

Surprisingly, the experiment on F2 offset showed that /u/ and /i/ are not significantly different from each other. One reason for this might be that, although there might be enough information in the V_1F2 offset to indicate when there is a vowel following and when there is not, via vowel-to-vowel coarticulation, there might not be enough information to show exactly which vowel follows the consonant. In other words, the non-significant result might be a function of the fact that the measurement was taken too early in the acoustic record to show the identity of the vowel.

Finally, there has been evidence for a continuum of productions not only for vowels but, also, for consonants, with consonant duration at its minimum in lenited cases and at its maximum in canonical cases. Thus, there exist two routes to lenition in CG. It seems that vowel realisation is dependent on consonant realisation. If there is a complete closure, voicing ceases and the vowel is voiceless; if the closure is not achieved, voicing is preserved and the vowel is voiced. This suggests an aerodynamic explanation for voicelessness in the final syllable in terms of passive devoicing as a consequence of the supralaryngeal articulation.

5. ACKNOWLEDGEMENTS

The author would like to thank Professor Francis Nolan, Dr. Mark Jones and the anonymous referees for their comments on this paper.

6. REFERENCES

[1] Gordon, M. 1998. The phonetics and phonology of non-modal vowels: a cross-linguistic perspective. *Berkeley Linguistics Society* 24, 93-105.

[2] Ohala, J. J. 1983. The origin of sound patterns in vocal tract constraints. In MacNeilage, P. (ed.) *The Production of Speech*. New York, USA: Springer.

[3] Varden, J. K. 1998. On high vowel devoicing in Standard Modern Japanese: implications for current phonological theory. Unpublished PhD Dissertation, University of Washington.

[4] Dauer, R. M. 1980. The reduction of unstressed high vowels in Modern Greek. *JIPA* 10, 17-27.

[5] Bonneau, A. 2000. Identification of vocalic features from French stop bursts. *Journal of Phonetics* 28, 495-502.

[6] Sereno, J. A., Baum, S. R., Marean, G. C., Lieberman, P. 1987. Acoustic analyses and perceptual data on anticipatory labial coarticulation in adults and children. *JASA* 81 (2), 512-519.

[7] Soli, S. D. 1981. Second formants in fricatives: acoustic consequences in fricative-vowel coarticulation. *JASA* 70 (4), 976-984.

[8] Yeni-Komshian, G. H., Soli, S. D. 1981. Recognition of vowels from information in fricatives: perceptual evidence of fricative-vowel coarticulation. *JASA* 70 (4), 966-975.

[9] Browman, C. P., Goldstein, L. 1992. Articulatory Phonology: an overview. *Phonetica* 49, 155-180.

[1] /i/ does not occur in enclitics so it was not possible to investigate it in the pre-antepenultimate condition.