CONTINUATION RISES IN PRE-1974 CYPRIOT GREEK

Joanna Przedlacka[1], Spyros Armostis[2], Mary Baltazani[1], Özlem Ünal-Logačev[3], John Coleman[1]
[1]Phonetics Laboratory, University of Oxford; [2]University of Cyprus; [3]Istanbul Medipol University

{joanna.przedlacka, mary.baltazani, john.coleman}@phon.ox.ac.uk, armosti.spyros@ucy.ac.cy, unaozlem@gmail.com

ABSTRACT

Before the 1974 partition of Cyprus most areas were predominantly Greek-speaking. At the same time there were also many mainly Turkish-speaking settlements and a number of localities with ethnically mixed populations. In this paper we examine the continuation rise intonation tune in archival recordings of the speech of nine Cypriot Greek-speaking males born between 1894 and 1934, from three towns. Using a functional data analysis, we compare their speech patterns to those of their Standard Modern Athenian and mainland Turkish-speaking contemporaries, and to Asia Minor Greek, another Turkish contact dialect in different social circumstances. Our analysis reveals the presence of two patterns in the f_0 curve shape and time alignment of the continuation rise tunes, demonstrating that even the inhabitants of relatively unmixed localities in Cyprus had adopted a Turkish-like intonation pattern in a proportion of their utterances.

Keywords: intonation; mixed populations; Cypriot Greek; Turkish-Greek contact varieties; continuation rise tune; curve fitting.

1. INTRODUCTION

Recent literature ([1-4]) supports the idea that ongoing language contact as experienced by bilingual speakers results in intonational variation and change, giving rise to novel patterns which may combine elements from both contextual languages. Speakers of Asia Minor Greek (AMG), a group of contact dialects originally spoken on the territory of modern Turkey, exhibit a mixture of Greek-like and Turkish-like patterns in continuation rise utterances, both in the f_0 curve shape and the time alignment ([3]). [3] also found that prosodic characteristics of Turkish have persisted in AMG for about a century after the cessation of its contact with Turkish.

Here, we extend the study of continuation rises to Cypriot Greek (CyG), a Greek–Turkish contact variety found in predominantly Greek-speaking communities, while AMG was used in a predominantly Turkish-speaking context. We examine whether the Greek–Turkish sociolinguistic contact in Cyprus resulted in CyG speakers exhibiting a mixture of patterns from the two source languages in the intonation of their continuation rise tunes, as we previously found in AMG. Standard Modern Greek as spoken in Athens (henceforth Athenian) and mainland Turkish are used as controls. We show that CyG continuation rises indeed have two patterns of alignment, one similar to Athenian, the other to Turkish, and we compare the bimodal CyG pattern to that of AMG.

Athenian is the standard variety used in Greece for official communication, in education and in the media. It does not display markedly regional characteristics of any traditional Greek dialect [5], but constitutes an amalgamation of different varieties as a result of internal migration in the latter half of the 20th century [6]. The variety of Turkish examined here is the standard variety as spoken in Istanbul and Western Anatolia. Though not all the Turkish speakers in the corpus come from Istanbul, it was shown in [3] that there were no significant differences in their continuation rise tunes. Cyprus remained under Ottoman rule from 1571 until it was annexed by Britain in 1914 [7], acquiring independence in 1960. When the island was partitioned in 1974, the population was 78.9% Greek and 18.4% Turkish [8]. Very little is known about CyG intonation [9, 10]. Regarding the effect of Greek–Turkish contact on intonation patterns, [11] examined the influence of Cypriot Greek on Cypriot Turkish but we know of no prior work on Turkish influence on CyG.

2. MATERIALS AND METHODS

The data come from 42 male speakers: 8 Greek speakers from Athens, Greece, 20 Turkish speakers from Turkey, 5 speakers of AMG from Cappadocia and 9 CyG speakers from the towns of Morphou, Rizokarpaso and Paphos in Cyprus (Figure 1). Before 1974, the ethnic makeup of these towns was different ([8]): while Paphos had a mixed population (6232 Greek, 2851 Turkish), the other two were mostly inhabited by Greek Cypriots (Morphou 6480 Greek, 123 Turkish; Rizokarpaso 3151 Greek, 2 Turkish). There are reports that the Paphos intonation is perceived as rather distinct from other CyG accents ([12, 13]). The Cypriot speakers were recorded in 1969, eight of them were born between 1894 and 1914 and one in 1934. The Morphou and Rizokarpaso speakers were recorded in interviews, the Paphos informant was recorded narrating a fairy tale.

Figure 1: A contemporary map of Cyprus, showing the locations of speech samples analysed: Paphos, Morphou and Rizokarpaso. Source: Wikimedia Commons

The Athenian data come from popular movies made in 1954–1961 and the Turkish data from radio and TV recordings from 1930–1989 and the Doegen collection (1917). The AMG data are narratives from holdings of Bibliothèque nationale de France (1927–1930). Using spontaneous and semi-spontaneous speech was motivated by reports that contact influences are more evident in less formal speech styles [14].

We analysed 1161 continuation rise tokens (208 Athenian, 479 CyG, 77 AMG, 397 Turkish), varying in length, lexical makeup and syntactic structure. An utterance is defined as a continuation rise if it: 1) is part of a broad focus declarative, 2) is non-final in a speaker's turn, 3) ends in a high boundary optionally followed by a short pause. The Athenian tune is a simple f_0 rise, where a L* nuclear pitch accent typically aligns with the stressed vowel, followed by an H- phrase accent ([15, 16]; Figure 2 top left). The Turkish tune is a rise-fall-rise, where a H* accent is followed by a L H- phrase accent ([17, 18]; Figure 2 top right). Impressionistically, the f_0 movements in Cypriot Greek exhibit two patterns, sometimes with Athenian-like and sometimes with Turkish-like characteristics (Figure 2 bottom left and right).

Native speakers of each language (three of the co-authors) identified the relevant utterances from the corpus of recordings, orthographically transcribed them and translated them into English.

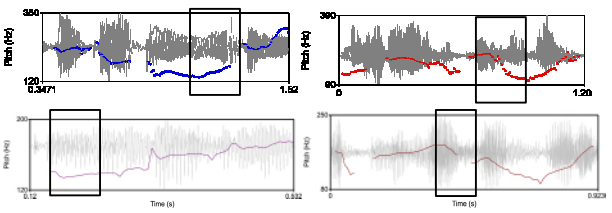

Figure 2: Continuation rise tune examples. Top left: Athenian [triˈada ˈatoma muˈ**i**pane] 'Thirty people told me'. Top right: Turkish [maˈsaja oˈ**tur**madan] 'Before sitting at the table'. Bottom left: Athenian-like Cypriot [iˈ**ka**mari] 'the room'. Bottom right: Turkish-like Cypriot [epluˈ**mi**zando] 'they adorned it'. A rectangle indicates the nuclear vowel, transcribed in bold.

They also located the nuclear vowel in each utterance and manually annotated its beginning and end, using Praat ([19]). We analysed the stretch of the continuation rise over a Region of Interest (ROI), from the beginning of the nuclear vowel to the end of the utterance. Irrespective of any differences in syntactic structure, the same ROI was defined prosodically across the four language varieties for comparability.

2.1. Modelling of f_0 hypotheses and comparisons

For each utterance, f_0 was measured every 10 ms using ESPS *get_f0* [20]. 10th-order polynomials $\hat{f_0} = \Sigma a_n t^n$ for $n = 0, \ldots 10$, were fitted to f_0 contours using the GNU Octave [21] *polyfit* function; pitch errors were inspected and manually corrected. The shape of f_0 contours in the ROI was modelled as 4th-order polynomials $\hat{f_0} = \Sigma a_n t^n$ for $n = 0, \ldots 4$, which were then transformed into orthogonal (Legendre) polynomials $\Sigma\ c_n L_n$ (cf. [22]). The five c_n coefficients capture general shape characteristics of the fitted $\hat{f_0}$ contour: c_0 is the average f_0 height of the contour; c_1 is its slope; c_2 models the shape as a parabola, concave up (or down if the sign is negative); c_3 models the shape as an *N*-like wave with a peak followed by a trough (or the reverse if the sign is negative); and c_4 models the shape as a more complex *M*- or *W*-like wave with more than one peak and trough.

Our hypotheses, informed by visual inspection of the curves and an impressionistic auditory analysis of the annotated CyG tokens are: 1) we expect an influence of Turkish on CyG (as in AMG) to be revealed in the f_0 shape through the similarity in coefficients c_2 and c_3: the Athenian f_0 curve is expected to be a low plateau followed by a final rise, so its c_2 and c_3 coefficients are expected to be near zero, while the Turkish pattern involves a rise-fall-rise f_0 movement, so its c_2 and c_3 coefficients are expected to be positive[1]; 2) based on our findings for AMG we expect two patterns of alignment in the AMG and CyG varieties, one Athenian-like (the L* trough before or within the nuclear vowel) and one Turkish-like, where the L trough is part of the edge tones and occurs after the nuclear vowel. The time of the f_0 minimum (i.e., L) in the Athenian L* H- or Turkish H* L+H- tune was determined by root-finding (i.e. calculating when the first differential of the modelled contour equals zero), using the GNU Octave/Matlab function `real(roots(polyder(a)))`. We define *trough alignment* τ as the difference between the L trough time and the end of the stressed vowel.

Since the AMG and Cypriot data is not unimodal (see Results), many standard parametric statistical models are not appropriate. Therefore, we fitted a 2-

component Gaussian mixture to the AMG and CyG data using the GNU Octave/Matlab `fitgmdist` function, to factor apart the Athenian-like and Turkish-like components of the bimodal AMG and CyG distributions. This function implements an iterative expectation–maximization (EM) algorithm to estimate the parameters of the components. Since the Athenian and Turkish distributions *are* unimodal ([3]), Gaussian distributions were fitted to the Athenian and Turkish data using the GNU Octave/Matlab `normpdf` function. The means of coefficients c_2, c_3 and τ were then compared across the four language varieties using three sets of t-tests.

1) We compared Athenian vs Turkish c_2, c_3 and τ, to check that their patterns are indeed different.

2) Similarly, for AMG and CyG, we tested whether those coefficients are different in the Athenian-like vs. the Turkish-like Gaussian components.

3) We tested whether the Athenian-like components of the CyG and AMG patterns are significantly different from the Athenian controls, and whether the Turkish-like components of the CyG and AMG patterns are significantly different from the Turkish controls. (This is the main question of this paper.)

3. RESULTS

Comparisons between the four varieties revealed that the intonation patterns in CyG are a mixture of Athenian-like and Turkish-like patterns, as we had also found in the AMG contact variety. In the CyG data, there was no significant difference between ethnically mixed and non-mixed towns, so we pool all the Cypriot data together. We present histograms of c_2, c_3 and τ in the four varieties (Figure 3) and the means and standard deviations of the Gaussian distributions (Table 1).

3.1. Continuation rise tune: shape

Mean c_2 is close to 0 in Athenian, indicating a shallower, broader parabolic component than in Turkish (cf. Figure 2), with a more dynamic rise-fall-rise, with a narrower trough than in Athenian and hence a greater c_2. The difference is highly significant ($p < 0.001$). CyG and AMG both show more spread-out c_2 distributions than Athenian or Turkish, and the two components of the Gaussian mixture models have highly significantly different means ($p < 0.001$), even more distinct in CyG than in AMG. In both CyG and AMG, the first component of c_2 is close to 0, similar (but not identical) to Athenian, whereas the second component is >2, similar but not identical to Turkish.

	c_2	c_3	τ (cs)
Athenian	0.19 (1.0)	-0.032 (0.24)	-15.2 (7.5)
Cypriot component 1	-0.01 (0.89)	0.08 (0.19)	-22.5 (7.7)
AMG component 1	0.56 (1.3)	-0.1 (0.25)	-15.2 (9.6)
Cypriot component 2	2.62 (1.19)	0.18 (0.36)	8.1 (10.1)
AMG component 2	3.1 (1.23)	0.36 (0.27)	8.0 (13.8)
Turkish	2.1 (1.7)	0.19 (0.38)	18.0 (11)

Table 1: Means (standard deviations) of c_2, c_3 and τ. For CyG and AMG, means and standard deviations of the two components in the Gaussian mixture models are given separately, for comparison with the single Gaussian pdfs used to model the Athenian and Turkish controls.

As expected, the Turkish-like component of the CyG pattern is *not* significantly different from the Turkish controls in the shape coefficients c_2 and c_3. However, contrary to our expectation, the L trough alignment in the Turkish-like component of CyG *is* significantly different from the Turkish controls ($p < 0.001$). Furthermore, c_2, c_3 and τ of the Athenian-like components of the CyG pattern *are* significantly different from the Athenian controls (for c_2, $p < 0.05$; for c_3 and τ, $p < 0.001$). Somewhat differently from CyG, the Turkish-like component of AMG is highly significantly different from the Turkish controls in c_3 and τ. The parameters of the Athenian-like component of AMG are not significantly different from Athenian except for c_3, which is significantly different only at the $p < 0.05$ level.

In spite of these significant differences between the AMG and CyG components and the Athenian and Turkish controls, it should be noted that he magnitudes of the differences are very small, showing that the two components of CyG and AMG are *similar* to Athenian and Turkish, consistent with the hypothesis that CyG and AMG intonation is a mixture of the two control varieties.

AMG resembles Turkish more than Athenian in contour shape, according to coefficient c_3, which was usually close to 0 in Athenian, i.e., the trough of the contour preceded the peak. In contrast, c_3 was positive in Turkish and AMG component 2, that is, the peak of the contour preceded the trough.

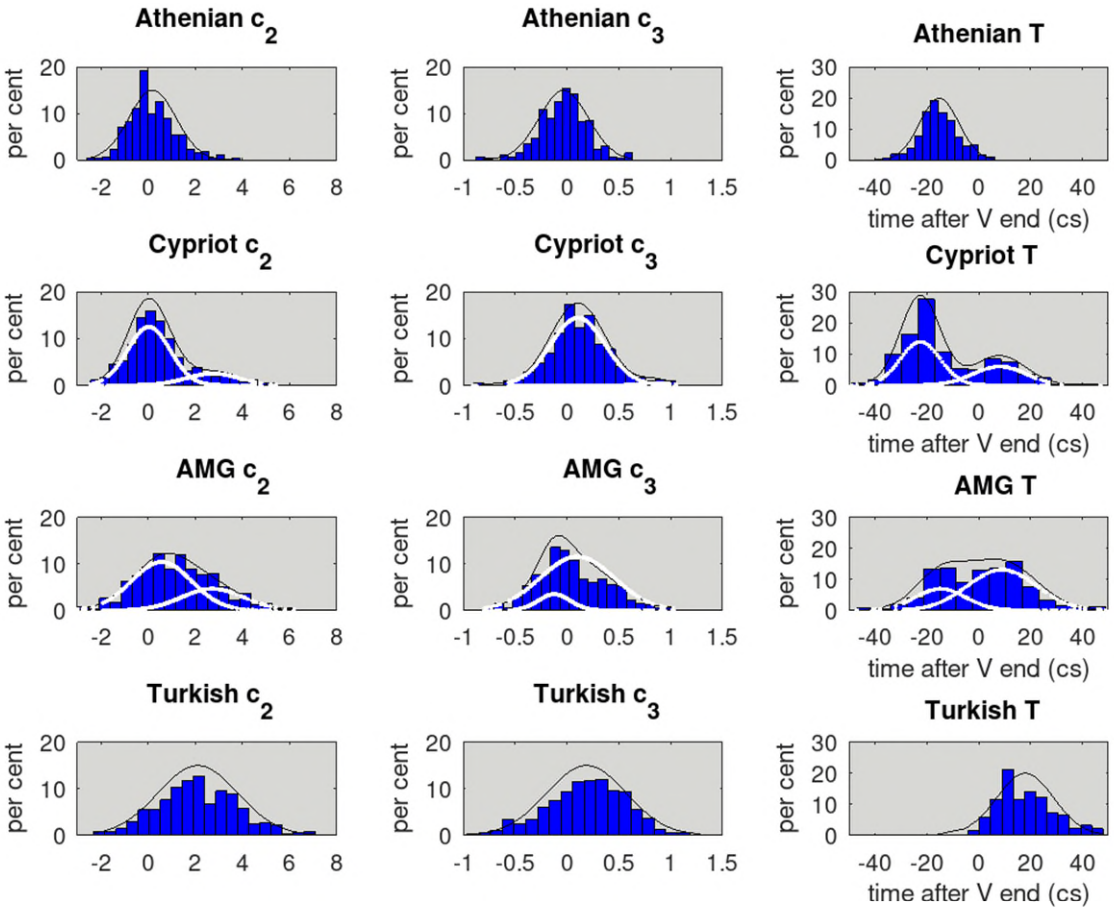

Figure 3: Histograms of *c2*, *c3* and τ (time lag from the nuclear vowel end to the L trough) in continuation rises in Athenian, CyG, AMG and Turkish. Black line: Gaussian distribution fitted to the Athenian and Turkish data, and a 2-component Gaussian mixture fitted to the AMG and Cypriot data. White lines: The two components of the mixture models.

3.2. Continuation rise tune: alignment

The distribution of trough alignment in CyG continuation rises is bimodal, with two peaks, one resembling the Athenian and the other the Turkish distribution. In AMG and CyG, the Athenian-like and Turkish-like Gaussian components of τ are highly significantly different ($p < 0.001$). The Athenian mean is significantly greater i.e. later than the CyG component 1 ($p < 0.001$), by c. 0.7 s, and the Turkish mean is significantly greater than the AMG and CyG component 2 ($p < 0.001$), by c. 0.1 s.

4. DISCUSSION

As reported previously [3, 4], there are differences in the shape and alignment of Athenian and Turkish continuation rises. The nuclear vowel aligns with a trough in Athenian but with a peak in Turkish. In general, CyG continuation rises had variable realisations resembling sometimes the Athenian and sometimes the Turkish intonation patterns. However, the CyG trough in component 1, although similar to Athenian, was aligned significantly earlier than in Athenian, and the trough in component 2 was aligned significantly earlier than in Turkish. The shapes of the two patterns of CyG continuation rises were also similar to (but significantly different from) their Athenian and Turkish controls. These results suggest that the variation in the intonational patterns observed in CyG, as in AMG, can at least in part be accounted for as a result of language contact.

More generally, these results provide further support to previous reports [2-4] that contact between languages from different families which have differences in syntax, morphology and phonology, can influence the intonation systems of the contact variety. Specifics of the intonational phonology and the tune-text alignment seem to be transferred between languages.

5. REFERENCES

[1] Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *J. Phon.* 32, 543–563.

[2] Queen, R. 2012. Turkish−German bilinguals and their intonation: Triangulating evidence about contact induced language change. *Language* 88(4):791–816.

[3] Baltazani, M., J. Przedlacka, Ö. Ünal-Logačev, P. Logačev & J. Coleman. 2022. Intonation of Greek in contact with Turkish: a diachronic study. Language Variation and Change, 34(3), 271-303. doi:10.1017/S0954394522000126.

[4] Baltazani, M., Przedlacka, J., Coleman, J. 2020. Intonation of Greek–Turkish contact: a real-time diachronic study. Proceedings of the 10th International Conference on Speech Prosody 2020, 730-734, DOI: 10.21437/SpeechProsody.2020-149.

[5] Trudgill, P. 2003. Modern Greek dialects. A preliminary classification. *Journal of Greek Linguistics* 4:54–64.

[6] Allen, Peter S. 1986. Positive aspects of Greek urbanization: The case of Athens by 1980. *Ekistics* 53(318/319):187–194. ww.jstor.org/stable/43621978.

[7] Borowiec, A. 2000. *Cyprus: a troubled island.* Greenwood Publishing Group. ISBN 978-0-275-96533-4.

[8] Cyprus Ministry of Interior. 1992. The Demographic Structure of Cyprus. Parliamentary Assembly. p. 6. Archived (PDF) from the original on 7 January 2011. Retrieved 22 January 2011.

[9] Grice, M., D. R. Ladd & A. Arvaniti. 2000. On the place of "phrase accents" in intonational phonology. *Phonology* 17:143-185.

[10] Themistocleous, Ch. 2012. Cypriot Greek Nuclear Pitch Accents (*Ta pirinika tonika ipsi tis kypriakis ellinikis*). In Z. Gavriilidou, A. Efthymiou, E. Thomadaki & P. Kambakis-Vougiouklis (eds), Selected papers of the 10th ICGL. Komotini, Greece: Democritus University of Thrace: 796–805.

[11] Kappler, M. 2011. A tale of two languages. Tracing the History of Turkish-Greek Language Contacts. *Türk Dilleri Araştırmaları*, 21.1: 95-130

[12] Christodoulou, Ch. 2015. *The Local Dialect of Northwest Paphos: Phonological Description* [In Greek]. PhD dissertation, Aristotle University of Thessaloniki.

[13] Fotiou, C., Grohmann, K.K. 2022. A Small Island With Big Differences? Folk Perceptions in the Context of Dialect Levelling and Koineization. *Frontiers in communication* 6, article 770088. www.frontiersin.org.

[14] Barnes, H. & J. Michnowicz. 2015. Broad focus declaratives in Veneto-Spanish bilinguals: Peak alignment and language contact. *Studies in Hispanic and Lusophone Linguistics* 8(1):35–57.

[15] Baltazani, M., & Jun, S-A. 1999. Topic and focus intonation in Greek. *Proc 14th ICPhS*, vol. 2, 1305-1308.

[16] Arvaniti, A. & Baltazani, M. 2005. Intonational Analysis and Prosodic Annotation of Greek Spoken Corpora. In S.-A. Jun, (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: OUP. 84–117.

[17] Özge, U., Bozsahin, C. 2010. Intonation in the Grammar of Turkish. Lingua *120 (1)*: 132-175.

[18] Ipek, C., Jun, S-A. 2014. Distinguishing Phrase-Final and Phrase-Medial High Tone on Finally Stressed Words in Turkish. In *Proc. 7th Speech Prosody International Conference*, Dublin, Ireland.

[19] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from http://www.praat.org/.

[20] Talkin, David. (1995). A robust algorithm for pitch tracking (RAPT). In W. B. Kleijn & K. K. Palatal (eds.), Speech Coding and Synthesis, Elsevier Science B.V. 497–518.

[21] Octave community. 2013. GNU Octave 3.7+. Available online at: http://www.gnu.org/software/octave/index.html

[22] Grabe, E., Kochanski, G., Coleman, J. 2007. Connecting intonation labels to mathematical descriptions of fundamental frequency. *Language and Speech* 50(3), 281–310.

[1] We do not expect any differences in c_0, c_1 or c_4 to be linguistically interesting because (i) normalisation of the utterance mean f_0 to 0 semitones means that c_0 will only indicate whether the ROI happens to be above or below the whole utterance mean; (ii) c_1 is generally positive, because the data are all continuation rises of various gradients; (iii) the c_4 term is only included in the model to improve its overall goodness-of-fit to the data, and is not linguistically interpretable. For brevity, we do not examine c_0, c_1 or c_4 further.