

INTONATIONAL VARIATION AND DIACHRONY: GREEK CONTACT VARIETIES

Joanna Przedlacka, Mary Baltazani, John Coleman

Oxford University Phonetics Laboratory
{joanna.przedlacka, mary.baltazani, john.coleman} @phon.ox.ac.uk

ABSTRACT

Relatively little work exists regarding the impact of long-term language contact on intonation. In this paper we give an overview of the project ‘Intonation and diachrony: a phonetic investigation of the effects of language contact on intonational patterns’. In this project we are investigating the intonational tunes of regional varieties of Greek whose speakers have a history of cohabitation with speakers of Venetian Italian (Cretan and Corfiot Greek) or Turkish (Asia Minor and Cypriot Greek). We compare the shape of the f_0 contours of the contact dialects to the corresponding tunes in both Standard (Athenian) Greek and their respective donor languages. Standard Autosegmental–Metrical analysis is combined with statistical modelling of f_0 curves using Functional Data Analysis. The robustness of contact effects over time is traceable through comparisons of speech corpora spanning a century.

Keywords: language contact, Modern Greek dialects, intonation, historical development

1. INTRODUCTION

In multi-ethnic communities bi- or multilingualism leads to the emergence of contact varieties [19]. The impacts of contact on lexicon and grammar are well-documented [20, 11, 18], but there is little research into its effects on regional variation or diachronic change that may arise out of contact situations. In a large-scale project (ESRC grant number ES/R006148/1) we investigate how contact between Modern Greek and other languages shaped the intonation of regional varieties of Greek, using archival recordings. We use a novel methodology, based on mathematical modelling of intonation curves in speech recordings from the recent and more distant past. In contrast to most intonational work, typically carried out on controlled data sets recorded in the laboratory, this method allows an investigation of large corpora with multiple speaking styles. Besides comparisons of contemporary data, it is applicable to modelling intonation change over time.

2. PROJECT GOALS

The project addresses the question of how contact between different languages influences the intonation of regional varieties. The project has two aspects: (1) a synchronic comparison of the contact dialects with their contextual languages and Standard Athenian Greek (henceforth Athenian) (2) a longitudinal analysis of the effects of contact on several varieties. For (1) we follow the following steps: first, we determine how the dialectal melodies differ from Athenian. Second, we analyse the realisation of the melodies in Turkish and Venetian Italian. Third, we measure the similarity between the corresponding melodies in each Greek variety and the donor language. The overarching goal of this aspect is to examine whether all Greek dialects vary along the same intonational dimensions, such as the f_0 of the tones that compose the melodies, the time alignment of tones with the segmental material. The longitudinal aspect of the research seeks to discover the retention or loss of contact effects on varieties. We shall investigate whether influences of the donor language weaken with time after the contact has ceased. To address this question, we will compare the intonation patterns in archival recordings dating back to the 1920s with more recent ones for each of the dialects under investigation.

3. THE CORPORA AND MELODIES

3.1 The corpora

In view of its historical character and wide geographical spread, the project makes use of existing digital audio recordings of Greek dialects collected over the past 100 years. We have compiled a comprehensive repository of Greek regional speech, including several dialects that have been affected by contact with Turkish (Asia Minor Greek) and Venetian Italian (Cretan and Corfiot); we shall also augment the collection with Cypriot data (contact with Turkish). To carry out cross-linguistic comparisons we have assembled corpora of Athenian and the contextual languages: Turkish and speech from the Veneto region.

The bulk of our data comes from institutional archives and online platforms [1, 7, 8, 10, 12, 13, 21, 23] as well as from linguists¹, who have shared their field recordings made in recent decades. We have secured agreements with the data owners to be permitted to use them in research and to publish selected examples. A small portion of the material comes from freely available on-line resources, such as language courses, film, radio, TV excerpts and YouTube videos.

The corpora (c. 400 hours) include reading passages, lists, dialogues, map tasks and sociolinguistic interviews with single or multiple informants. The speakers vary in age (20-94 years old) and speaking style; the recordings vary in length from a few minutes to an hour. Despite the expected pattern variability for each melody (Zipf's law, [9]), the size of our collection is large enough to allow the most common patterns to be revealed. The use of natural speech corpora has the advantage of being more representative of the population in its variability due to the large number of tokens.

All the original sources were acquired as digital data, but they came in a wide variety of audio file formats (mp3, mp4 and PCM wav; two channel or mono), bit rates or sampling rates (44.1 kHz, as for CDs, 22.05 kHz, or 16 kHz). A portion of the data was not immediately suitable for research as some recordings had been made from ¼ inch tape played at different speeds, which required speeding up or slowing down, or in some cases to be reversed in time to be restored. For the purposes of our analyses we converted all the material to 16 kHz, monophonic, uncompressed PCM .wav audio files.

3.2 The target melodies

We focus on the effects of Venetian Italian and Turkish on the development of intonation patterns (target melodies) in four regional varieties of Greek: Cretan, Corfiot, Asia Minor and Cypriot, chosen because their speakers historically interacted with speakers of the Veneto dialect (Corfiot) or Turkish (Asia Minor, Cypriot). Cretan has been in the domain of both Venetian and Ottoman influences in different periods. The target melodies of those varieties are compared to Athenian and to their respective contextual languages. Our goal is to analyse three melodies for the contact varieties: (a) declarative utterances ending in low pitch, indicating finality, (b) continuation tunes ending in high pitch, indicating non-finality and (c) polar questions.

Within the Autosegmental-Metrical model (AM) [16, 14], the *alignment* of a pitch landmark, like a peak (H) or a trough (L), with a particular stressed syllable (*pitch accent*) or a phrase edge (*phrase*

accent for small clauses and *boundary tone* for the whole utterance boundary) has been found to be crucial to the meaning of utterances. The melody of an utterance can be analysed to a large extent as the combination of its pitch accents and boundary tones. The description and classification of these landmarks and their contribution to meaning forms a key part of current models of intonation.

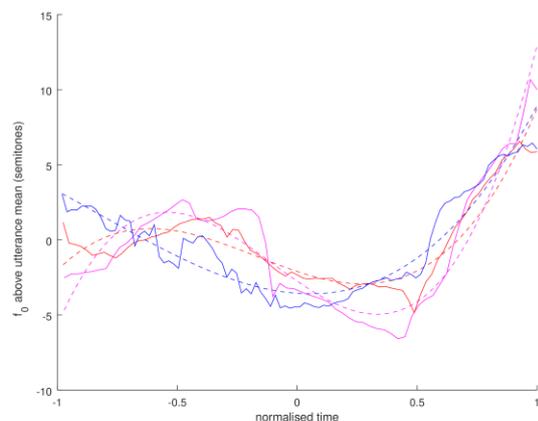
We use phonological landmarks of the standard AM model as a basis for mathematical modelling of the shape of the f_0 curve from the region of interest. The choice of the region of interest is guided by an exploratory impressionistic analysis of the shape differences between the varieties under investigation. It is based on phonological criteria, typically from the annotated start of the pre-nuclear or nuclear vowel (the latter defined as the last stressed syllable of the word in focus) to the utterance boundary. The f_0 was measured every 10ms using ESPS get_f0 and the intonational contours in this region are modelled fitting a 4th order (Legendre) polynomials, for details of the procedure see [6]. A significant advantage of this approach is that it augments the abstract AM analysis with quantitative, empirically testable models of tunes, allowing comparisons of entire pitch curves of utterances rather than merely its pre-categorised components (i.e. pitch accents and edge tones).

4. SYNCHRONIC ASPECTS: TURKISH AND VENETIAN INFLUENCES ON REGIONAL ACCENTS OF GREEK

4.1 Turkish

Our work on Turkish-Greek contact has so far focused on the realisation of continuation tunes in Asia Minor Greek. A more detailed account is presented in [6]. We shall extend our research on Turkish-Greek contact to Cypriot Greek and Cretan. Examples of typical f_0 shapes of continuation tunes are shown in Fig. 2. Here, the Asia Minor Greek contour is a rise-fall-rise, closely resembling the Turkish sample, as opposed to the Athenian Greek, a low level followed by a rise.

Figure 1: Continuation rises in Asia Minor Greek (magenta) vs. Turkish (red) and Athenian (blue). The curves are based on single tokens from female speakers. The continuous lines show measured f_0 , the dashed lines modelled f_0 . The data are time-normalised.



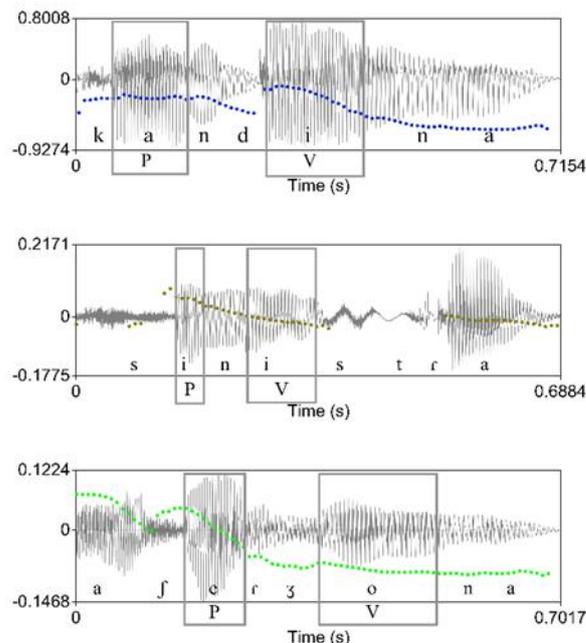
4.2 Venetian Italian

We compare the declarative tune in Cretan Greek to its equivalents in Athenian and Venetian Italian. We hypothesise that the Cretan intonation will show similarities to Venetian rather than Athenian speech, given the five centuries of Venetian occupation of Crete (1204-1699) and resultant contact between the varieties. We examine the Cretan intonation of declaratives [17], example utterances are shown in Fig. 2.

While declarative tunes in all the three varieties end in a fall, there are important differences in the timing, scaling of the fall, as well its alignment with the nuclear vowel. In Athenian both the beginning and end of the fall H*L align with the beginning and end of the nuclear syllable [di] respectively, followed by low boundary tone (L-L%). Within that region, the f_0 falling contour is steep, in contrast to Venetian and Cretan. In contrast to Athenian in both Cretan and Venetian the fall (HL*) starts in the pre-tonic syllable ([fe] in Cretan, [si] in Venetian) and ends typically around the middle of the nuclear vowel ([zo] in Cretan, [ni] in Venetian) and it is followed by a more gradual low plateau from the middle of the stressed vowel to the end of the utterance. This impressionistic description, exemplified by the tokens presented in Fig. 1 was confirmed by the statistical analysis of the modelled curves. The slope of the Athenian fall reflected by the larger negative coefficient of the linear term of the polynomial (μ : -5.7; σ : 13.4), is significantly steeper than the slope in Venetian (μ : -.29; σ : 13.4; $p < .01$), while the slope of the Cretan fall (μ : -3.7; σ : 29.2) is not significantly different

from Venetian, $p=.451$), for detailed analysis of shape of the declaratives in those three varieties see [5]. Given the long-term past contact between Cretan and Venetian, the result is in line with our hypothesis.

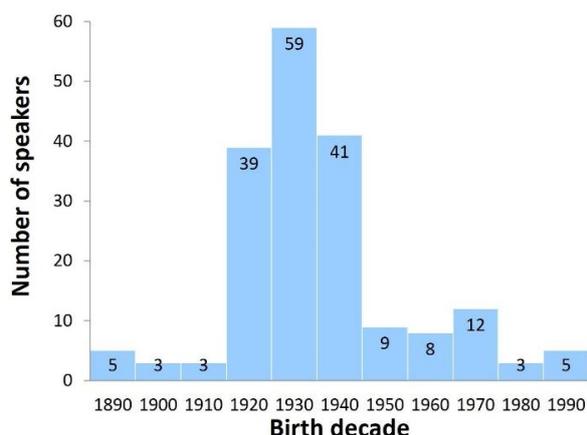
Figure 2: Illustrations of declaratives in Athenian (top), Cretan (middle) and Venetian Italian (bottom); P = pre-nuclear vowel; V = nuclear vowel.



5. DIACHRONIC ASPECT OF THE CORPORA

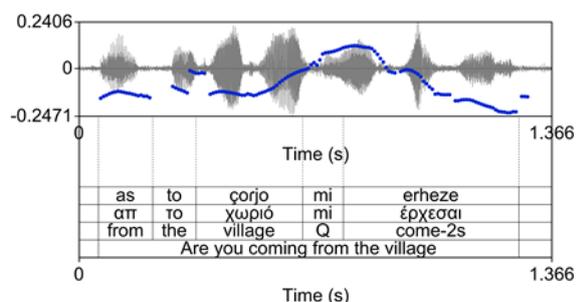
The corpora we have collected together were recorded over the course of the 20th and early 21st century. For diachronic comparisons, the most comprehensive corpus is the Asia Minor Greek material, which spans five generations of speakers. In total, we have 85 hours of audio recordings of 188 speakers of the varieties of Greek spoken in various locations in the Anatolian Peninsula (Cappadocia, Pontus, Istanbul and Smyrna). The oldest material consists of recordings of prisoners of war from the First World War made in 1917, with the informants' birthdates going back to 1894. These recordings form part of the Wilhelm Doegen collection and were obtained from [7]. The audio recordings of the youngest informants, born in the 1990s, come from [13]. They are heritage speakers of Asia Minor Greek, recorded in 2007 in five localities in mainland Greece (Kato Apostoloi, Ksiroxori, Mandra, Neo Agioneri, Plagia), where they were resident at the time. The bulk of Asia Minor Greek data comes from speakers born between the 1920s and 1940s. A histogram of the speaker birthdates is shown in Fig. 3

Figure 3: Asia Minor Greek speakers in our corpus by birth decade.



An early example of Asia Minor Greek comes from a 1927 recording available from [8]. The 23-year old male is a speaker of Cappadocian, a variety of Greek historically spoken in central region of the Anatolian Peninsula (present-day Turkey) which had been in contact with Turkish for almost nine hundred years [22]. Misti, the village he originates from, ceased to exist in 1924 when its ethnically Greek inhabitants were forcefully expelled from their homeland following the Lausanne Treaty on population exchanges between Greece and Turkey. The speaker is recorded in a conversation with his uncle, discussing family matters. The waveform of the utterance with pitch track is shown in Fig. 4. The recording is 2 minutes 46 seconds long and it contains 6 tokens, which display Turkish traits.

Figure 4: An acoustic waveform of a polar question, uttered by a male speaker of the Misti dialect. Time is shown on the horizontal axis. The panels below show segmental transcription, Greek orthography and English translation. Source: <https://gallica.bnf.fr/ark:/12148/bpt6k129226q/f1.media>.



The example utterance shows features of Turkish intonational phonology, as it differs from non-contact Greek dialects in both pitch accent and alignment. In Athenian speech, the polar question tune has been

analysed as L* LH- L% [4, 5, 6]. The nuclear syllable is aligned with a trough. In contrast, in Turkish polar questions, the nuclear word is closely aligned with a high peak [15]. The token is also closer to Turkish than Athenian in its morphological composition by its use of the question particle /mI/, which closely follows the nucleus. Athenian does not use any such morpheme to indicate polar questions. Athenian speech typically shows a downward slope from the nuclear syllable and a concave upwards rise from an L* near the utterance end. Here, the token displays a tune characteristic of Turkish, i.e. an upward slope from the nucleus followed by a final rise-fall.

Regarding the alignment, the pattern displayed here is closer to Turkish (close nucleus-peak alignment) while in Athenian the peak comes up to two syllables after the nucleus. All four other polar question tokens from the same conversation display the Turkish alignment pattern.

In the next phase of our project we shall carry out a series of diachronic comparisons of the Asia Minor Greek material. We shall compare the speech of speakers who acquired their dialect in contact circumstances to that of subsequent generations, who are no longer in daily contact with Turkish. To discover the retention or loss of contact effects on this variety we will be looking for evidence of continuity or disjunction of the intonational shapes at different points in time to their Athenian equivalents. Our goal is to develop a model of a diachronic change of intonation.

Our methodology, which integrates acoustic phonetics, intonational phonology, mathematical modelling and speech processing, is an approach not hitherto applied to historical linguistics. We also envision that it will be useful both to the linguistics and speech processing communities regardless of the language under investigation. In addition to the time dimension, which will be a series of comparisons of intonation curves from recordings gathered at different points in time, our experimental design makes it possible to compare dialects which had been in contact with the same language, though the circumstances of the contact differ. Here Asia Minor Greek, whose contact with Turkish had ceased several generations ago, can be compared to Cypriot, where the contact is ongoing.

6. REFERENCES

- [1] Academy of Athens <http://www.academyofathens.gr/en>
- [2] Arvaniti, A., Baltazani, M. 2005. Intonational Analysis and Prosodic Annotation of Greek Spoken Corpora. In Jun, S-A. (eds) *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford: OUP, 84-117.

- [3] Baltazani, M., Jun, S.-A. 1999. Focus and topic intonation in Greek. *Proc. 14th ICPHS*, vol. 2, San Francisco, 1305-1308.
- [4] Baltazani, M. 2006. Effects of stress on intonational structure in Greek. In Hoffmann, R., Mixdorff, H. (eds), *Speech Prosody 2006*, Third International Conference, ISCA Archive, <http://www.isca-speech.org/archive/sp2006>, paper 156.
- [5] Baltazani, M., Przedlacka, J., Coleman, J. 2019. Intonation in contact: Athenian, Cretan and Venetian declaratives. Paper presented at *3rd Phonetics and Phonology in Europe*. Lecce, Italy.
- [6] Baltazani, M., Przedlacka, J., Coleman, J. 2019. Intonation in contact: Asia Minor Greek and Turkish. *ICPhS2019*.
- [7] Berliner Lautarchiv at the Humboldt-Universität <https://www.lautarchiv.hu-berlin.de/en/sound-archive/>
- [8] Bibliothèque nationale de France. <http://www.bnf.fr>
- [9] Black, P.E. 2009. Zipf's law. In Pieterse, V., Black, P. E. (eds), *Dictionary of Algorithms and Data Structures*. <https://xlinux.nist.gov/dads/HTML/zipfslaw.html>. Accessed 8/12/2018.
- [10] CLIPS: Corpora e Lessici dell'Italiano Parlato e Scritto. <http://www.clips.unina.it/it/>
- [11] Clyne, M. 2003. *Dynamics of Language Contact. English and Immigrant Languages*. Cambridge: Cambridge University Press.
- [12] Committee for Pontic Studies <http://epm.omegatechnology.gr/top9.aspx?materialid=1>
- [13] Endangered Languages Archive <https://www.soas.ac.uk/elar/>
- [14] Ladd, D.R. 2008. *Intonational Phonology*. 2nd edition. Cambridge: CUP.
- [15] Özge, U., Bozsahin, C. 2010. Intonation in the Grammar of Turkish. *Lingua 120 (1)*: 132-175.
- [16] Pierrehumbert, J. 1980. *The phonology and phonetics of English intonation*. PhD dissertation. Cambridge, MA: MIT.
- [17] Przedlacka, J., Baltazani, M., Coleman, J. 2018. Contact-induced diachronic change in intonation: Greek non-standard varieties. *LabPhon16*, Lisbon 19-22 June 2018.
- [18] Ralli, A. 2012. Morphology in language contact: verbal loan blend formation in Asia Minor Greek. In Vanhove, M. *et al.* (eds), *Morphologies in contact*, Berlin: Akademie Verlag, 185-201.
- [19] Sankoff, G. 2001. Linguistic Outcomes of Language Contact. In Trudgill, P., Chambers, J., Schilling-Estes, N. (eds), *Handbook of Sociolinguistics*. Oxford: Blackwell, 638-668.
- [20] Thomason, S. 2001. *Language Contact*. Edinburgh: Edinburgh University Press.
- [21] UCL Phonetics Collection <https://sounds.bl.uk/Accents-and-dialects/UCL-phonetics-recordings>
- [22] Vassalou, N., Janse, M, Papazachariou, D. 2017. The vowel system of Mišótika Cappadocian. In Georgakopoulos, Pavlidou, T (eds), *Proc. The 12th International Conference on Greek Linguistics. 2: 1139-1154*. Berlin Edition: Romiosini.
- [23] Vivaldi: Vivaio Acustico delle Lingue e dei Dialetti d'Italia© 1998-2018. <https://www2.hu-berlin.de/vivaldi/?id=0001&lang=it>

i We would like to thank Mark Janse, Petros Karatsareas and Dimitris Papazachariou for sharing their field recordings.