The intonation of polar questions in Cypriot Arabic: Prosodic contact in an endangered language

Mary Baltazani[1], *Spyros Armostis*[2], *Elinor Payne*[1]

[1]University of Oxford, UK
[2]University of Cyprus, Cyprus
mary.baltazani@phon.ox.ac.uk, armosti.spyros@ucy.ac.cy, elinor.payne@phon.ox.ac.uk

Abstract

We analyse the intonation of polar questions (PQs) in Cypriot Arabic (CYA), a severely endangered peripheral variety of Arabic spoken by Cypriot Maronites, the intonation of which has not been formally examined before. We compare the patterns in CYA with those in Cypriot Greek (CYG) and in Syrian Arabic (SYA) to search for possible differences with the latter, as a result of centuries of isolation, and similarities with the former, as a result of centuries-long contact on the island of Cyprus.

We elicited PQs in CYG and CYA through a map task and analysed them combining quantitative modelling of intonational contours with conventional Autosegmental-Metrical tools of tune-text alignment.

The results reveal CYA questions to be phonologically very similar to those in CYG, albeit with differences in their fine phonetic detail. Polar questions in the Cypriot varieties of both languages have a L* nucleus followed by H-L% edge tones. This pattern is phonologically different from the pattern in SYA questions, which has been reported as a L* or a L*+H nucleus followed by a final rise. We discuss the implications of these findings for a theory of prosodic contact.

Index terms: Cypriot Arabic, Cypriot Greek, intonation, polar questions, intonation contour modelling

1. Introduction

Recent work on language contact provides evidence of its effects on intonation even when the languages in question are from different families with distinct systems of syntax, morphology and, more significantly, phonology [1-3]. Here, we investigate the intonation of polar questions (PQs) in Cypriot Arabic (CYA), a severely endangered variety of Arabic spoken by Cypriot Maronites of the village of Kormakitis, and officially recognised as a minority language of Cyprus since 2008.

The ancestors of CYA speakers migrated to Cyprus from Greater Syria in several waves from the 7th to the 13th century [4]. CYA has been described as one of the "peripheral varieties" of Arabic [4-7]. While many Arabic varieties have come into sustained contact with non-Semitic languages, what distinguishes peripheral varieties is their early separation, and sustained isolation, from the core Arabic-speaking area very early on [4, 8, 9]. Hence, in the case of CYA, sustained contact with Cypriot Greek (CYG) is coupled with centuries of isolation from other varieties of Arabic.

In the context of linguistic evolution and contact-induced change, CYA has undergone several transformations due to contact with CYG, and the internal displacement of the CYA-speaking community following the events of 1974 [10] has served to heighten these effects. Today, CYA speakers are all bilingual with CYG. However, while some individuals, even within the present-day community, acquired CYA as their first language, the balance of usage has shifted, and there is a growing preference for CYG across various linguistic domains. Analyses of CYA, encompassing lexicon, morphology, phonetics, and phonology, reveal a pervasive influence of CYG on CYA ([11] for a thorough review and references). However, scant attention has been directed towards the phonetic and phonological analysis of CYA [11]. Gaps are especially notable in the investigation of its intonation and broader prosody, despite anecdotal reports suggesting potential CYG influence in these domains.

Our study aims to address these gaps. We focus on the intonation of PQs in CYA and CYG to investigate potential similarities between the Cypriot varieties of these two languages, which, in light of differences between *non-Cypriot* varieties, may be attributable to the enduring effects of long-term contact. We additionally explore whether CYA PQs share any traits with SYA PQs, as a possible benchmark comparison with a non-peripheral variety that was geographically proximal, albeit centuries ago. We posit that, despite marked differences between Greek and Syrian Arabic, the prolonged contact between CYG and CYA has resulted in relative convergence between them which is quantifiable through modelling of the PQ intonational contours.

1.1. Polar questions in Cypriot Greek vs Syrian Arabic

Starting with the benchmark comparison between CYG and SYA, in both varieties PQs are distinguished from statements by intonation only, making the characteristics of the intonational contours especially critical. CYG PQs are reported to have a L* nuclear pitch accent (NPA) on the stressed syllable of the nuclear word, followed by H- L% edge tones on the final word in the question [2, 12-14]. This results in a characteristic rise-fall shape.

In SYA in contrast, PQs are reported to have a "rise shape", consisting of a L* or L*+H nucleus followed by a final rise [15,16]. The nuclear accent is reported to associate with the first syllable of the nuclear word, suggesting that it may not necessarily align with the stressed syllable [16: 192]. Thus, CYG and CYA are markedly distinct in their PQ intonation.

2. Methods

2.1. Participants, materials and procedure

We elicited and analysed 140 tokens of PQs from the Cypriot varieties of interest: 5 speakers of Cypriot Greek (2F, 3M) from Nicosia and 1 speaker of Cypriot Arabic (F) from Kormakitis. Participants were aged 58-78 (mean: 64 years) and fulfilled two criteria. First, as native speakers of CYA are now all over 40 years, participants had to be at least this age. Second, to control for the geographic distribution of languages following the events of 1974 [10], we selected participants born before 1974. Given the complex linguistic situation in Cyprus, a detailed sociolinguistic questionnaire was administered to gather background information.

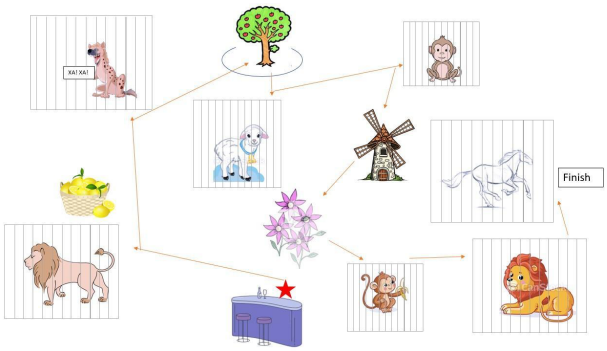

Figure 1: *Zoo map (experimenter's version).*

PQs were elicited through a map task based on the premise that the participant and experimenter (native speakers of their language variety) had lost each other on a visit to a zoo/farm and had to communicate via mobile phone. Both had a map of the zoo/farm, but only the experimenter's map showed the path to be followed (Fig. 1). Participants were instructed to ask yes-no questions to find the experimenter, e.g. "*Do I go towards the lion?*". The animals, plants, and objects in the maps were chosen to elicit tokens that were as close to fully voiced as possible, and with a variety of stress positions.

2.2. Data segmentation and labelling

For each language variety, all instances of PQs were identified, checked for naturalness, orthographically annotated, assigned to one of three stress conditions (final, penultimate and antepenultimate stress in the nuclear word), and translated into English by a native speaker of the language variety in question. Judgments were made on the positioning of the nuclear focus, which, for all tokens examined here, was on the final word. Consonants and vowels were segmented and labelled for the nuclear stressed syllable and any post-nuclear syllables. We analysed PQs over a Region of Interest (ROI), i.e. from the beginning of the stressed vowel in the nuclear word to the end of the utterance. The same ROI was defined in both language varieties for maximal comparability.

2.3. f_0 modelling, hypotheses and comparisons

For each utterance, f_0 was measured every 10 ms using ESPS *get_f0* [17]. 10th-order polynomials (1)

$$\hat{f_0} = \Sigma a_n t^n \quad (1)$$

for $n = 0, \ldots 10$, were fitted to f_0 contours using the GNU Octave [18] *polyfit* function; pitch errors were inspected and manually corrected. The shape of f_0 contours in the ROI was modelled as 4th-order polynomials (2)

$$\hat{f_0} = \Sigma a_n t^n \quad (2)$$

for $n = 0, \ldots 4$, which were then transformed into orthogonal (Legendre) polynomials $\Sigma\, c_n L_n$ (cf. [1]). The five c_n coefficients capture general shape characteristics of the fitted $\hat{f_0}$ contour: c_0 is its average f_0 height; c_1 is its slope; c_2 models it as a parabola, concave up (or down if the sign is negative); c_3 models it as an *N*-like wave (a peak followed by a trough or the reverse if the sign is negative); and c_4 models it as a more complex *M*- or *W*-like wave with more than one peak and trough.

The time of the f_0 minimum and maximum, corresponding to the L* and H- tones respectively, was determined by root-finding (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* as the difference between the L* trough time and the start time of the nuclear vowel and *peak alignment* as the difference between the H- peak time and the start time of the word-final vowel. NB: In the final-stress condition, the nuclear vowel *is* the word-final vowel.

Given the unequal distribution of speech data with six speakers of CYG and only one speaker of CYA, inferential statistical comparisons were considered methodologically inappropriate, and we therefore focus on descriptive statistics. The language varieties were compared though the 7 parameters as the dependent variables (the alignment of the L trough and the H peak as well as the shape of f_0 contours in the ROI through the five c_n coefficients). Stress position (final, penultimate and antepenultimate) and language (CYG, CYA) were the independent variables.

Based on impressionistic descriptions of CYA and reports on SYA [15,16], our hypotheses were: 1) CYA questions resemble CYG polar questions phonologically, that is, they share the L* NPA plus H- L% tune characteristic of CYG. Such similarity will be determined by (a) the alignment of the L and H tones—in CYG, the former is within or slightly before the stressed vowel in the nuclear word and the latter occurs within the final vowel of this word; and (b) the f_0 shape, revealed mainly through coefficients c_2 and c_3. The CYG pattern is a shallow trough followed by a peak and a final fall (i.e., a L* NPA followed by H-L% edge tones), so its c_2 and c_3 coefficients are expected to be negative: a negative c_2 corresponds to a downward parabolic shape (i.e., the rise-fall of the edge tones) and a negative c_3 to a fall-rise-fall movement. 2) CYA PQs are different from SYA ones phonologically, i.e. the former ends in a rise-fall (L* H-L%), in contrast to the final rise (with H edge tones) reported for the latter.

3. Analysis and Results

Overall, the comparisons between CYA and CYG polar questions revealed similarities both in f_0 shape and in alignment characteristics. Only comparisons of interest will be reported below.

Figure 2 gives representative examples of PQs, from top to bottom, in CYG, CYA and SYA. The examples of the two Cypriot varieties come from the corpus we collected, while the SYA example was downloaded from the [19] data collection.

Red boxes in the three examples show the location of the stressed vowel in the nuclear word.

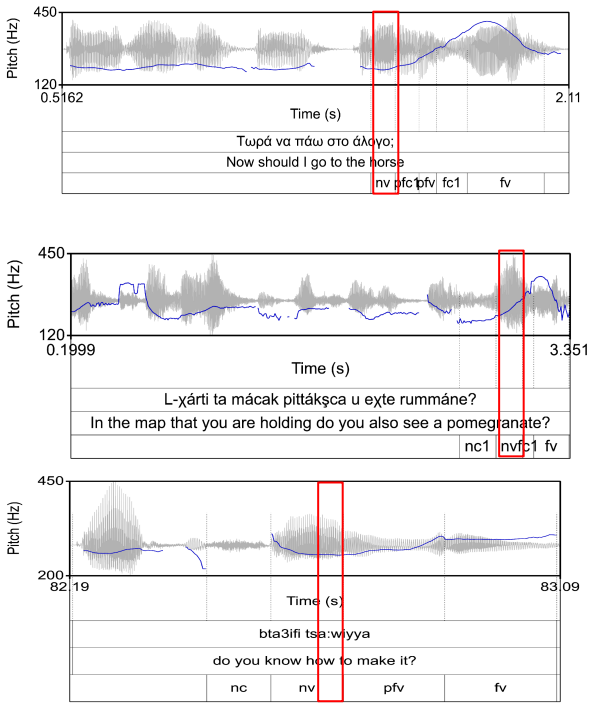

Figure 2: *Examples of PQs for CYG (top), CYA (middle) and SYA (bottom; from [19]).*

Table 1: *Means (standard deviations) of c_2, c_3, and alignment of the L and H tones for CYA and CYG.*

	c_2	c_3	L align (cs)	H align (cs)
CYA	-1.69 (1.15)	-0.07 (0.32)	3.58 (6.04)	8.89 (9.43)
CYG	-1.53 (1.37)	-0.06 (0.39)	5.98 (4.9)	7.39 (5.9)

Comparisons of the shape parameters between CYA and CYG tentatively confirmed hypothesis 1): the lack of evidence of a difference in them indicates that they share the L* NPA and H-L% edge tones. This similarity can best be seen in the mean values of the c_2 and c_3 coefficients (two left columns in Table 1 and Figure 3). As expected, c_2 and c_3 are negative in both CYA and CYG, with similar mean values. The cubic coefficient (c_3) models the pitch contour shape as a wave and its negative values are modelled as fall-rise-fall movements. The parabolic shape modelled by the negative quadratic coefficient (c_2) indicates the very salient rise-fall pitch contour of the edge tones.

Turning to the alignment of the L and H tones, as expected, the L* NPA appeared within the nuclear vowel in both CYG and CYA and the H tone appeared within the final vowel. Table 1 (two rightmost columns) and Figure 4 show the distance of L (in cs) from the start of the stressed vowel in the nuclear word (blue in Figure 4) and the distance of H from the start of the final vowel (red).

Although the distribution of values in CYG and CYA is similar, there are phonetic alignment differences between the languages, which appear to depend on lexical stress placement. First, in the antepenultimate stress condition, the L* NPA in CYA appears either before the start of the stressed vowel or within it, while in CYG it occurs typically within it. Second, in proparoxytone words the H tone typically appears before the start of the final vowel in CYA, but mostly within it in CYG.

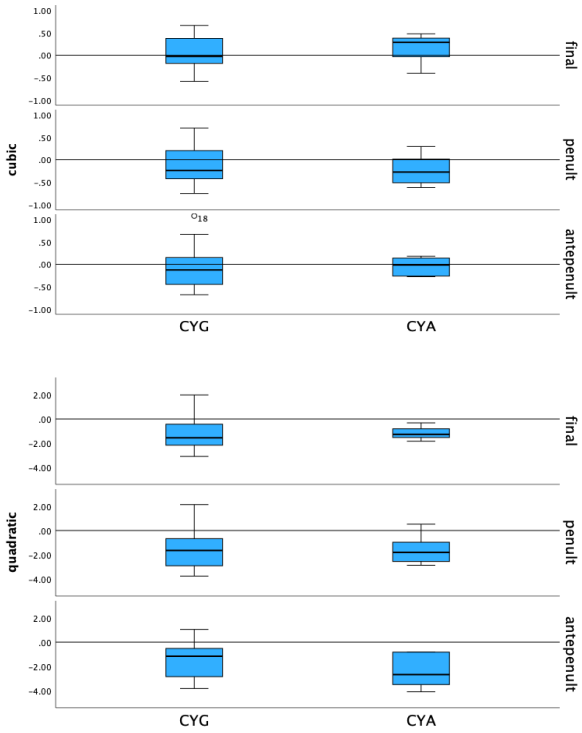

Figure 3: *Distribution of coefficient values in CYG and CYA. Top: cubic c_3. Bottom: quadratic c_2.*

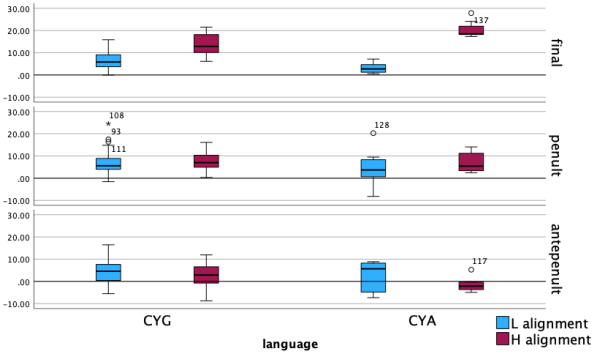

Figure 4: *Distance of the L (blue) and H (red) tones, respectively, from the beginning of the nuclear and final vowel in CYA and CYG.*

In oxytone words, however, these differences are diminished, with the H tone appearing later within the final

vowel (compared to the other two stress conditions) in both CYA and CYG. This is probably due to tonal crowding: where the nuclear vowel is the final vowel, both the L and H tones must be realised within it. Differences are nevertheless observable even in this condition. In CYA the H is pushed further to the right than in CYG. Additionally, the alignment of L in CYA also appears to be affected, occurring earlier, and much closer to the start of the vowel than in CYG. While these findings must be taken as preliminary (because they are based on only one CYA speaker), they suggest possible differences in the underlying structural organisation (reminiscent of reports [16] of stress-mediated variation in the NPA in Damascene Arabic).

4. Discussion

This paper presents a first quantitative comparison between CYG and CYA in the intonation of polar questions. The findings tentatively confirm our predictions that the f_0 shape and tune-text alignment patterns of CYA polar questions are similar to those of CYG and indicate that the polar question tune in these two languages has the same basic phonological structure, with a L* nuclear accent and H-L% edge tones. This tune differs from the L*+H (or L+H*) H- H% tune reported for Syrian [16].

This similarity between CYA and CYG, together with marked differences between CYA and SYA, is strongly suggestive of resulting from sustained contact of CYA with CYG, coupled with centuries of isolation from other varieties of Arabic. Such results provide further support to reports (e.g. [1]) that contact between typologically and 'genetically' distinct languages can influence the intonation of the contact varieties. Specifics of both the intonational phonology and the tune-text alignment appear to be transferred between languages. Intriguingly, while it appears that gross patterns in the tune may 'float' freely enough through sustained contact, apparent differences in the effects of lexical stress position on tonal alignment indicate prevailing influences of typological differences at a higher level. This sheds further light on the complexities involved in prosodic contact, where the integration of tunes and grammar may ultimately clash.

Although the evidence we present tentatively suggests contact-induced convergence, we cannot rule out alternative, non-contact explanations without being more certain about which variety of Arabic to compare CYA to, or indeed the extent of within-Arabic variation for the general area. It is likely that the descendants of Cypriot Maronites migrated from areas to the north of Aleppo [4], which may be quite distinct in their PQ tunes from the Damascene variety we have taken as an interim benchmark (not to mention variation over time).

Our next steps will include further investigation of these parameters, as well as examination of other types of questions. Nevertheless, [15] suggests the rising question tune observed for Damascene Arabic is remarkably pervasive across a wide range of Arabic varieties, with the notable exception of Moroccan, another geographically peripheral variety. The combination of apparent *divergence* from generalised patterns for Arabic with apparent *similarity* with CYG provides plausible evidence for the contact narrative.

5. Acknowledgements

This study is part of a wider research project 'Mapping prosodic convergence in the Eastern Mediterranean' (https://mappingprosody.phon.ox.ac.uk/) for which we gratefully acknowledge the support of John Fell Oxford University Press (OUP) Research Fund, grant 0011309, and of the Economic and Social Research Council (UK), grant ES/Y005767.

6. References

[1] Baltazani, M., J. Przedlacka, Ö. Ünal-Logačev, P. Logačev & J. Coleman. 2023. Intonation of Greek in contact with Turkish: a diachronic study. *Language Variation and Change,* 34(3), 271-303. doi:10.1017/S0954394522000126.

[2] Payne, E., S. Armostis, E. Lombardo, R. Simon & M. Baltazani. 2023. Prosody in Contact: Polar Questions in Cypriot varieties of Greek and Turkish. *Proceedings 20th ICPhS*, Prague.

[3] Przedlacka, J., M. Baltazani, S. Armostis, Ö. Ünal-Logačev, J. Coleman. 2023. Continuation rises in pre-1974 Cypriot Greek. *Proceedings 20th ICPhS*, Prague.

[4] Borg, Alexander. 2004. *A Comparative Glossary of Cypriot Maronite Arabic (Arabic-English) with an Introductory Essay*. Leiden: E. J. Brill.

[5] Borg, Alexander. 1985. *Cypriot Arabic: A Historical and Comparative Investigation into the Phonology and Morphology of the Arabic Vernacular Spoken by the Maronites of Kormakiti Village in the Kyrenia District of North-Western Cyprus*. (Abhandlungen für die Kunde des Morgenlandes hrsg. von der Deutschen Morgenländischen Gesellschaft Band XLVII, 4.). Stuttgart: Deutsche Morgenländischen Gesellschaft; Kommisionsverlag Franz Steiner, Wiesbaden Gmbh.

[6] Roth, Arlette. 1973–1975. Le verbe dans le parler arabe de Kormakiti (Chypre). In *Epetiris Tou Kentrou Epistimonikon Erevnon VII* (1973–1975). Nicosia: Kentro Epistimonikon Erevnon, pp. 21–117.

[7] Grigore, George. 2019. Peripheral varieties. In *The Routledge Handbook of Arabic Sociolinguistics.* Edited by Enam Al-Wer and Uri Horesh. London: Routledge, pp. 117–33.

[8] Borg, Alexander. 1994. Some evolutionary parallels and divergences in Cypriot Arabic and Maltese. *Mediterranean Language Review*, 8: 43–73.

[9] Roth, Arlette. 2006–2007. Quelles nouvelles perspectives s'ouvrent avec l'exploration et la description des dialectes arabes dits périphériques? In *Romano-Arabica 6/7: Peripheral Arabic Dialect*. Bucharest: University of Bucharest. Center for Arab Studies, pp. 133–48.

[10] Karyolemou, Marilena. 2018. Language revitalization, land and identity in an enclaved Arab community in Cyprus. In S. Drude, N. Ostler & M. Moser (eds.), Endangered languages and the land: Mapping landscapes of multilingualism, Proceedings of FEL XXII/2018 (Reykjavík, Iceland), 14–18. London: FEL & EL Publishing.

[11] Armostis, S., & Marilena Karyolemou. 2023. Contact-Induced Change in an Endangered Language: The Case of Cypriot Arabic. *Languages* 8: 10. https://doi.org/10.3390/languages8010010

[12] Arvaniti, A. 1998. Phrase accents revisited: comparative evidence from Standard and Cypriot Greek. *Proceedings of the 5th International Conference on Spoken Language Processing*, vol. 7. Sydney. Australia, 2883–2886.

[13] Grice, M., Ladd, R., Arvaniti, A. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(2), 143–185.

[14] Themistocleous, C. 2011. *Prosody and information structure in Athenian and Cypriot Greek* (PhD Thesis). National and Kapodistrian University of Athens [In Greek].

[15] Hellmuth, S. 2018. Variation in polar interrogative contours within and between Arabic dialects. *Proc. Speech Prosody* 2018, 989-993, doi: 10.21437/SpeechProsody.2018-200.

[16] Malek al Hasan & Shakuntala Mahanta. 2022. The Intonational Phonology of Syrian Arabic: A Preliminary Analysis. *Proc. Speech Prosody* 2022, 190-194 [2022], doi: 10.21437/SpeechProsody.2022-39.

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

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

[19] Hellmuth, S., & Almbark, R. 2019. Intonational variation in Arabic Corpus 2011-2017. [Data Collection]. Colchester, Essex: UK Data Archive. 10.5255/UKDA-SN-852878; https://reshare.ukdataservice.ac.uk/852878/