To appear in: M. Cruz, S. Frota, & P. Oliveira (Eds.), Prosodic variation (with)in languages: Intonation, phrasing and segments. (Studies in Phonetics and Phonology). Equinox. *Author accepted version.*

Text-tune alignment in Tunisian Arabic yes-no questions

Sam Hellmuth, University of York, UK

1 Introduction

Adjustment of the intonational 'tune' to the segmental 'text' has been observed in a number of languages in contexts of tonal-crowding, most commonly in the form of either compression or truncation of the tonal contour (Grabe, 1998; 2004). Nevertheless, a number of cases of the reverse phenomenon have been reported, in which the segmental 'text' is adjusted, through vowel insertion or lengthening, to accommodate the intonational 'tune'.

This paper reports on a previously undescribed apparent 'text-tune' adjustment phenomenon observed in a corpus of Tunisian Arabic (TA) speech data, collected for a wider investigation of the intonational phonology of TA. In our data, the final nuclear accent in yes-no questions is commonly a (delayed peak) rise followed by a complex boundary tone (analysed here as L*+H H-L%). In such tokens, an epenthetic vowel is frequently appended to the last lexical item by some speakers (Fig. 1).

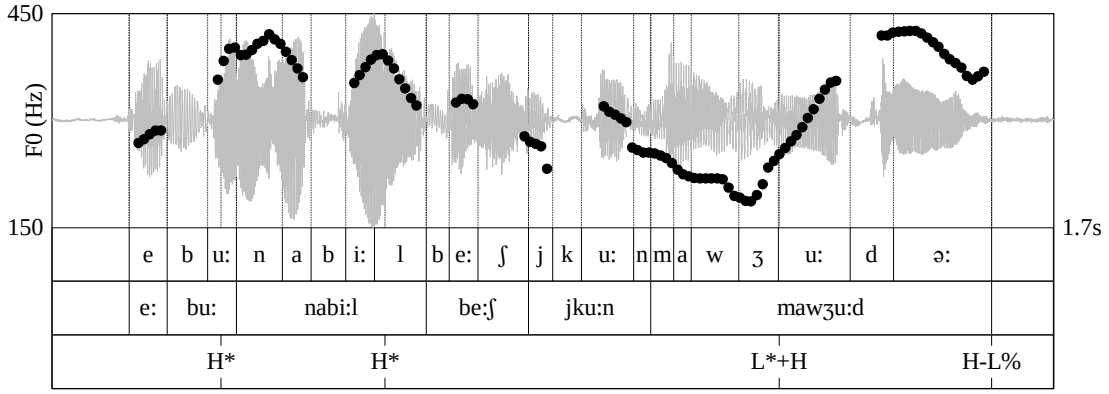

Fig. 1: Yes-no question from read speech data produced with vowel epenthesis [tuns-ynq6-f4].

e:	bu:	naˈbi:l	be:ʃ	jku:n	/mawˈʒu:d/ [mawˈʒu:də:]
um	father-of	Nabil	PART	he will be	present

Will Nabil's father be there?

This pattern of utterance-final vowel epenthesis has not previously been reported in the (small) literature on TA intonation, nor in any other work on the intonation patterns of neighbouring dialects of Arabic, to the best of our knowledge. In the present study we investigate the incidence of utterance-final vowel epenthesis in TA to see if its occurrence varies by discourse context, prosodic contour, speech style, metrical structure or segmental content, as well as by individual speaker. These are all factors which have been reported to condition text-tune phenomenon in one or more varieties of Romance languages spoken across the Mediterranean from Tunisia. In addition we explore a new potential factor, the sociolinguistic variable of gender.

Overall then, the goal here is to explore to what extent the patterns of text-tune adjustment observed in Tunisian Arabic resemble those observed in Romance languages

across the Mediterranean; could the pattern be explained solely in terms of 'north-south' language contact, or should we look elsewhere for explanations?

We start by providing an overview of what is known so far about the prosodic phonology of Tunisian Arabic, and about tune-text adjustment phenomena in other languages, as motivation for the research questions of the paper. The methods employed in the study are then outlined, including details of the corpus data and analysis techniques. The results of the analysis are presented, taking each potential conditioning factor in turn culminating in a statistical analysis which seeks to clarify the relative contribution of each factor to the observed variation in the data. The implications of the findings are then discussed in the light of intonational phonology and intonational typology, as well as our understanding to date of patterns of language contact known to influence Arabic varieties, and our knowledge of potentially comparable patterns in other dialects. A brief conclusion closes the paper, with suggestions for potentially fruitful avenues of further research.

2 Background

2.1 The prosodic phonology of Tunisian Arabic

Tunisian Arabic (TA), defined here as the urban (sedentary) variety spoken in Tunis, has received less attention in the phonological literature than some other dialects of Arabic. In terms of syllable structure, TA permits complex onsets and codas, with geminates permitted in syllable initial and syllable final position (Bouchhioua, 2008), but, as in all Arabic dialects, onsets are obligatory. A wide range of syllable types is thus observed (CV, CVV, CVC, CCV, CCVV, CCVC, CVCC, CCVCC). The distinction between phonological 'long'/'short' vowels in TA is primarily a matter of quality, not duration; nevertheless, in our phonetic transcriptions, 'long' vowels are marked as long (i.e. with [ː]) as they would be in cognate words in other varieties of Arabic, for ease of comparison, but without postulating phonemic vowel length in TA.

TA displays a typical, rule-governed stress assignment pattern for Arabic dialects, characterised by the following algorithm: stress falls on a final superheavy syllable if present, else on the penultimate syllable (Ghazali, 1973; cited in Bouchhioua, 2008). The phonetic correlates of stress in TA have been investigated in some detail by Bouchhioua (2008) who found word-level lexical stress to be marked primarily by means of spectral balance and F1 lowering, and phrase-level post-lexical stress (pitch accent) marked primarily by duration.

There has been little prior work on the intonation of Tunisian Arabic, limited to several Masters dissertations carried out in Tunisia (Aloulou, 2003; Knis, 2004; Saadi, 2014. According to Aloulou (2003, cited in Bouchhioua 2008) discourse-non-final declaratives bear a rising tone, whereas discourse-final declaratives bear a falling tone; yes-no questions bear a rising tone, but wh-questions generally bear a falling tone (a subset show a complex rise-fall or fall-rise). Knis (2004; summarised in Ghazali, Hamidi, & Knis, 2007) presents sample utterances from two speakers each of a range of Arabic dialects, including TA, but the dataset is too small to permit meaningful generalisations. The most recent study of these, Saadi (2014), describes the contour in yes-no questions as a rise-fall contour with a final H-L% boundary, partially resembling the findings in our study here.

However, none of these previous small-scale studies of TA intonation mention any pattern of vowel epenthesis associated either with yes-no questions or with an interrogative rise-fall contour. To the best of our knowledge, therefore, the present study is the first description of this pattern, which is however robustly observed in the dataset examined. It is possible that the pattern is a relatively recent innovation in TA; alternatively, it may be the case that the small sample sizes used in prior studies failed to capture the full range of intonational expression in TA, which we are now able to explore with the aid of the present, somewhat larger corpus.

2.2 A typology of tune-text alignment

One of the first studies to explore tune-text alignment phenomena in detail was Grabe (1998), who observed a difference between English and German in resolving the problem of how to realise a complex intonational contour, in contexts of tonal crowding. A typical tonal crowding context would be a monosyllabic word in utterance-final position, such that all and any tonal targets in the nuclear pitch accent and following edge tones must be realised on a single syllable. Grabe noticed that, in general, in Southern Standard British English (SSBE) in such contexts, the slope of the intonational contour is increased, resulting in a steeper fall or rise, than is observed when the same prosodic contour is realised on a polysyllabic word or with non-final stress; this strategy is described as *compression* of the contour. In contrast, in parallel contexts, in Standard German, the slope of the intonational contour is unchanged, and instead the rise or fall fails to be fully realised, but is cut off before the fall or rise is complete; this alternative strategy is described as *truncation* of the contour. In subsequent work, on varieties of British English spoken in different cities in the UK, Grabe and colleagues observed variation across dialects. Some British dialects display truncation of the contour (e.g. Newcastle), and others show compression of the contour (e.g. Leeds), in parallel contexts. A stylised example is shown in Figure 2 below, based on Grabe (2004:11).

	Newcastle	Leeds
long word:	 limousine	 limousine
short word:	 lip showing compression	 lip showing truncation

Fig. 2: Stylised representation of compression vs. truncation, as observed in varieties of British English spoken in Newcastle vs. Leeds (based on examples in Grabe 2004:11).

A logical alternative to the problem of tonal crowding - instead of adjusting the intonational 'tune' to fit the available segmental 'text'- is to increase the amount of segmental material available, so as to be able to realise the full intonational contour, without either compression or truncation. Adjustment of the segmental material could in principle take the form either of lengthening of existing segments or insertion of additional segments. Both types of text-tune

adjustment have been reported in Standard European Portuguese (SEP), but we focus here on reported cases of word-final vowel insertion, which may provide clues to the factors conditioning utterance-final vowel epenthesis in TA, which is the focus of the present study.

Word-final vowel 'insertion' (i.e. epenthesis), in the form of blocking of a more general rule of final-vowel deletion, is observed in SEP in yes-no questions and in vocative chants (Frota et al., 2015), and is restricted to cases where the word in question has a final sonorant consonant (examples show [ɾ], [l] and [n]). This contrasts with observations of the same phenomenon in the variety of Portuguese spoken in the Alentejo region of southern-central Portugal, which displays word-final epenthesis after a sonorant consonant in all sentence types. Epenthesis is thus analysed as a general marker of the right-edge of the intonational phrase in that variety (Cruz, 2013), mirroring a similar pattern reported for Galician. Word-final vowel epenthesis is seen in SEP in contexts with the nuclear contours listed in (1) below:

(1)	yes-no questions	H+L* LH%
	greeting vocative call	H* !H%
	insistent vocative call	H* L%

The yes-no question contour has a complex bitonal boundary tone (see arguments against analysis of this as L-H% in (Frota, 2002)), which could be argued to require additional segmental material to be realised. Frota et al (2015) also note that the !H% boundary tone in SEP greeting vocatives displays 'spreading qualities', and that such spreading might plausibly be seen as a hallmark of both types of vocatives, suggesting that the relative 'weight' or complexity of the edge tones is a potential unifying factor across the contours and contexts in (1). Most of the cited examples in the work on SEP are taken from read speech, but there is at least one example reported from spontaneous speech map task data (Frota et al 2015 p 256 Fig. 7.13), suggesting that the phenomenon is not confined to read speech in SEP.

A similar pattern of tune-text adjustment, depending in part on the shape of the prosodic contour, has also been reported in an experimental study of loanwords into Bari Italian (Grice, Savino, Caffo, & Roettger, 2015). Amid a general picture of considerable variation by speaker and by item, the authors report that vowel insertion was conditioned by a number of factors including 'prosodic contour' (defined as question vs. statement), number of syllables in the word, and voicing of the final consonant in the phrase-final word. The incidence of word-final vowel epenthesis was higher in read speech than in spontaneous speech (Grice p.c.).

In summary therefore, utterance-final vowel epenthesis (or blocking of final vowel deletion) has been observed in certain varieties of two Romance languages spoken in southern Europe:

Table 1: Summary of factors reported to condition utterance-final vowel epenthesis.

language	*discourse function*	*segmental and metrical context*	*prosodic contour*
Standard European Portuguese	yes-no questions vocatives	if final consonant is sonorant	H+L* LH% H* !H% H* L%

Alentejo European Portuguese	all sentence types	if final consonant is sonorant	H+L* L% L* L%
Bari Italian	yes-no questions continuation rises	more common on monosyllables more common if final consonant voiced	L+H* L-H% L* L-H% H* H-^H%

2.3 Research questions

This paper seeks to determine whether or not the factors which have been shown to condition word-final epenthesis in varieties of European Portuguese and Italian also condition word-final epenthesis in TA. The study thus investigates a range of internal and external factors which might explain the observed patterns of word-final epenthesis in TA, in a small dataset extracted from a larger corpus of TA speech.

The key factors explored are variation due to individual speaker or lexical item, alongside discourse function (e.g. question, vocative, continuation), tonal contour (choice of nuclear accent and/or boundary tones), segmental and metrical context (number of syllables in the word, position of stress in the word, type of final segment) and speech style (spontaneous vs. read/scripted speech). Speech style is included due to the observation by Ng (2013) that word-final vowel epenthesis (or, *paragoge*, to use her term) is frequently observed in language contact scenarios, such as second language acquisition or loanword adaptation; Ng thus ascribes word-final epenthesis to hyperarticulation ('reduced gestural overlap') in effortful speech. If this characterisation is correct, then the incidence of word-final vowel epenthesis in TA is predicted to be higher in read speech than in spontaneous speech, as is indeed reported for Bari Italian (Grice, p.c.). Furthermore, if Ng's (2013) hypothesis that paragoge is a diagnostic of language change is correct, then we might reasonably expect that change to follow typical patterns observed in diffusion of linguistic innovations, and to be led by women (Labov, 2001; Eckert & McConnell-Ginet, 2003), resulting in a potential effect of gender as a conditioning factor. An additional factor explored here, therefore, which was not discussed in prior parallel studies on Italian or European Portuguese, is a potential role of gender, as a sociolinguistic factor.

3 Methods

The study exploits data from the Intonational Variation in Arabic (IVAr) corpus (Hellmuth & Almbark, 2017), rather than from an experiment designed specifically to target this construction; not all potentially relevant variables are systematically varied in the dataset under investigation therefore. The key potential conditioning factors can be investigated, however, and use of corpus data allows us to explore the extent to which vowel epenthesis is observed in both read and semi-spontaneous speech, as well as the degree of variation across speakers.

The data in the Tunisian Arabic portion of the IVAr corpus (which bears the short code 'tuns') was collected in Tunis, Tunisia, in April 2014. Recordings were made with 12 speakers (6 female, 6 male). They were aged between 20-24 at the time of recording, and all were born and raised in Tunisia, to parents who were both born and raised in Tunisia. All were first language speakers of TA, though all were also fluent in French, which is taught in schools in Tunisia from the age of 10 (Grade 5). Two speakers (f2 and f3) were born to

parents who were born in the south of Tunisia; the remainder were born to parents born in the north of Tunisia.

Declarative statements with varying information structure, together with a range of different question types, were embedded in a scripted dialogue which was read aloud as a role play by all 12 speakers, working in pairs. The metrical structure of the last lexical item in each target sentence was systematically varied, with stress on the antepenult, penult or final syllable. We extracted from these dialogues a set of read speech yes-no questions (*ynq*, N=68) and a set of control sentences (N=55) of other types (including declaratives and wh-questions) in which the last lexical item is the same as in one or more of the *ynqs*. We also extracted from the scripted dialogue a set of read speech vocatives (N=12). In addition, a search was made for tokens of vowel epenthesis in list items elicited in a Dialogue Completion Task (cf. Frota & Prieto, 2015) and for yes-no questions in spontaneous speech collected using a Map Task (Anderson et al., 1991). All of the data are from the IVAr corpus and details of the data elicitation materials are available from the IVAr database webpage.[1]

The extracted subsets of data provide tokens were submitted to qualitative impressionistic auditory analysis to identify various phenomena of interest, including the presence/absence of an utterance-final vowel and the shape of the prosodic contour. All data analysis was performed by the author, with qualitative analysis based on auditory impression with reference to the spectrogram and fundamental frequency contour using Praat (Boersma & Weenink, 2015). Criteria used for identification of an epenthetic vowel were the presence of periodic vibration and/or formant structure after the consonantal release, or a visible change in the formant structure and/or intensity after a word-final lexical vowel (following Grice et al 2015). Acoustic measurements of the epenthetic vowels was performed using a Praat script written by the author, to extract the duration, mean intensity and f1/f2 values (at the midpoint) of each labelled vowel.

For annotation of TA prosodic contours, the Autosegmental-Metrical framework is adopted (Ladd, 2008), in which intonational tunes are comprised of phonological tonal targets which are phonologically associated with certain positions in the metrical structure of the utterance. Like most – though not all – Arabic dialects described so far (Hellmuth, 2013), TA displays postlexical intonational marking of both the heads and edges of prosodic domains. An intonational tune in TA is thus formed of one or more pitch accents (tones associated with prominent syllables) and boundary tones (associated with the right edge of phrases). For the present study, relevant portions of the data were prosodically annotated by the author using a prototype annotation system for Arabic which is currently under development (Hellmuth & Almbark, in preparation), but which is based on the ToBI annotation system proposed for American English (Beckman & Elam, 1997; Beckman, Hirschberg, & Shattuck-Hufnagel, 2005). A key difference between the original ToBI system and the transcription system used here, is that in the original ToBI system, the high phrase accent (H-) in a final H-L% boundary combination (phrase accent + boundary tone) is assumed to have an effect of upstep on the following low boundary tone (L%) such that the H-L% sequence is realised with level pitch. In the transcription system used here, instead, a H-L% boundary combination is realised as a slight fall at the boundary (as seen in the example illustrated in Fig 1 above).

[1] http://ivar.york.ac.uk/

This analysis is motivated by the fact that, in this contour, the final H peak is aligned quite consistently at short distance before the phrase boundary, and is thus analysed as a phrase accent (H-). In contrast, the position of the elbow at the start of the rise towards this peak does vary according to prosodic structure of the utterance, with a somewhat earlier start to the rise in cases where the stressed syllable in the target word is earlier in the word (i.e. on the antepenult or penult syllable), but without showing tight alignment to the start of the accentual syllable, which is generally realised with low pitch; the post-accentual rise is thus analysed as a trailing tone which is part of a bitonal pitch accent associated with the accented syllable (L*+H). The specific detail of the labels used in annotation is not critical to the investigation here, though identification of the presence or absence of a complex rise-fall (L*+H H-L%) contour, as opposed to a plain rise (L* H-H%) is important, as will be seen. Additional quantitative analysis of the fundamental frequency contour was performed, therefore, in a subset of the data (the read speech data only), using a Praat script written by the author, to support the qualitative analysis. Descriptive statistics were produced using Excel or R (R Development Core Team, 2008) and a Classification by Regression Trees (CART) analysis (Baayen, 2008) was carried out using R.

4 Results

The vowel epenthesis pattern was initially observed during prosodic annotation of yes-no questions in read speech. Our goal in this paper is to establish the distribution of the pattern, by inspecting a wider range of available corpus data, in both read and semi-spontaneous speech, and exploring a range of potentially relevant factors which we treat in turn.

4.1 Discourse function and speech style

The first task is to determine whether vowel epenthesis is primarily conditioned by discourse function; that is, whether vowel epenthesis in TA marks yes-no questions only, or not. Table 2, below, shows the results of auditory analysis carried out to determine the presence or absence of a word-final vowel, in read speech yes-no questions. Just over half (54%) were produced with vowel epenthesis, suggesting that, although vowel epenthesis is common in yes-no questions, it is not a necessary or consistent cue to yes-no question status.

Table 2: Incidence of vowel epenthesis in read speech yes-no questions, by target utterance.

target	*last lexical item*		*number of tokens in which vowel epenthesis was observed*	
ynq1	[ˈja.ma.ni]	‘Yemeni’	6 (N=12)	50%
ynq2	[ˈba.la.di]	‘traditional’	7 (N=11)	64%
ynq3	[ˈze:.na]	‘Zena’	7 (N=12)	58%
ynq4	[la.ˈja:.li]	‘nights’	4 (N=11)	36%
ynq5	[ʔa.ˈmi:n]	‘Amin’	6 (N=10)	60%
ynq6	[maw.ˈʒu:d]	‘present’	7 (N=12)	58%
Total			37 (N=68)	54%

In contrast, when the set of control utterances from read speech was inspected, containing the same lexical item in utterance-final position as in one or more of the yes-no questions, but with a different discourse function (including focus statements and wh-questions), no cases of vowel epenthesis were observed, as shown in Table 3. This contrasts with the pattern observed in the Alentejo variety of European Portuguese, in which utterance-final vowel epenthesis is observed across a range of sentence types (Cruz, 2013).

Table 3: Incidence of vowel epenthesis in read speech control sentences, by target utterance.

target	*last lexical item*		*number of tokens in which vowel epenthesis was observed*	
coo1	[ˈja.ma.ni]	'Yemeni'	0 (N=11)	0%
idf1	[ˈja.ma.ni]	'Yemeni'	0 (N=12)	0%
whq1	[ˈja.ma.ni]	'Yemeni'	0 (N=11)	0%
con4	[ˈze:.na]	'Zena'	0 (N=12)	0%
con6	[maw.ˈʒu:d]	'present'	0 (N=12)	0%
Total			0 (N=58)	0%

Since vowel epenthesis was reported in vocatives in SEP, a set of read speech vocative utterances, extracted from the scripted dialogue were inspected also. Of the 12 vocatives inspected, only one was produced with utterance-final vowel epenthesis (tuns-voc-f3, see Fig. 4 below), which we assume would be interpreted with interrogative force, as well as vocative function, and which we will discuss further in the next section. Similarly, since vowel epenthesis was reported in continuation rises in Bari Italian, we inspected examples of semi-spontaneously produced lists, elicited using a Dialogue Continuation Task (DCT, see Methods). In one task, speakers were asked to list the days of the week, and in a second task, to list what they had eaten so far that day. DCT data was collected with all 12 speakers in our sample, but no instances of vowel epenthesis were observed in any of the lists produced.

Finally, to determine whether speech style conditions vowel epenthesis in TA (as was observed in Bari Italian, Grice, p.c.), we identified a set of yes-no questions produced in semi-spontaneous speech, in map task data. In total, 40 yes-no questions were identified in the map task data, and of these, 25 were produced with utterance-final vowel epenthesis (63%). This suggests that utterance-final vowel epenthesis in TA is not restricted to read speech. The incidence of vowel epenthesis is slightly higher in the spontaneous speech sample than in the read speech data, however, more of the tokens in the spontaneous speech sample were produced by female speakers (N=27) than by male speakers (N=13), and, as we shall see below, in the read speech sample female speakers tended to use vowel epenthesis more than male speakers. A sample token of vowel epenthesis in spontaneous speech is illustrated in Figure 3 below. The response of the interlocutor in this instance was 'yes' [ʔe:], which provides independent 'next-turn proof' (Hutchby & Wooffitt, 2008) of our auditory impression that the opening utterance can be classified as a yes-no question, since it was treated as such by the interlocutor in the original conversation.

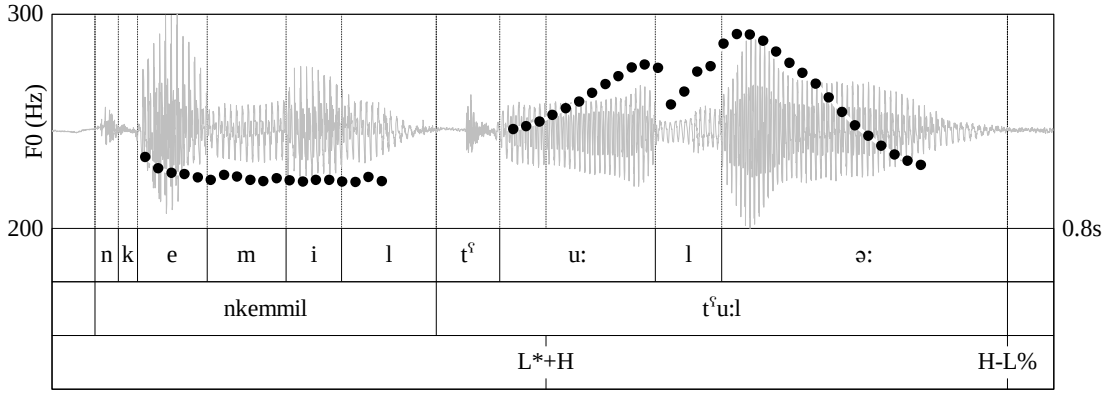

Fig. 3: Yes-no question from map task data produced with vowel epenthesis [tuns-mp1-f3.109-110].

ˈnkem.mil /t ˤu:l/ [ˈtˤu:lə:]

I-continue straight ahead

Should I go straight ahead?

Figure 4 shows the spread of values of measures of vowel duration and mean intensity in the labelled epenthetic vowels, as well as a plot of F1/F2 taken at the midpoint of each labelled vowel. Duration and intensity are normally distributed, but there is considerable variation in the quality of the inserted vowel, an issue to which we return below.

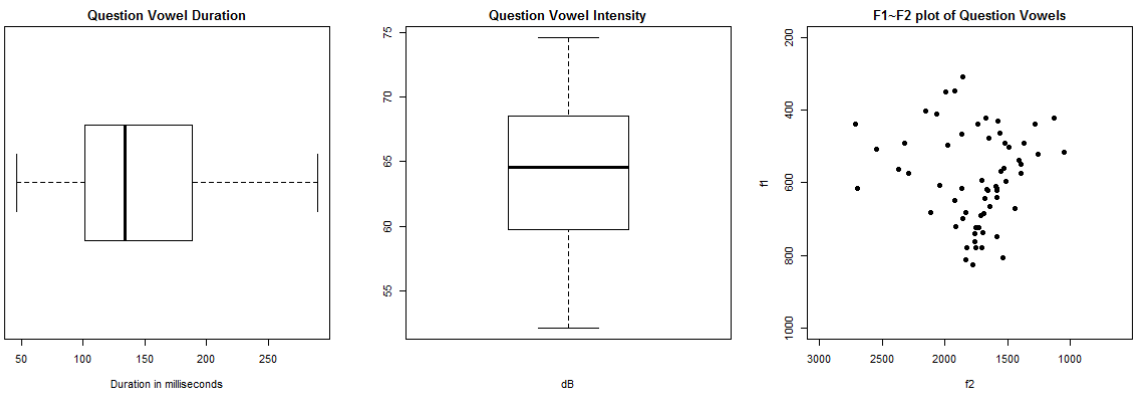

Fig. 4: All epenthetic vowels: boxplots of duration (left) and intensity (centre), and f1/f2 plot (right).

In summary, although the vowel epenthesis pattern is only observed consistently in yes-no questions, it is not a necessary cue to yes-no question status, since epenthesis is observed in roughly half of ynq tokens only. Other potentially relevant contexts did not yield a similar proportion of vowel epenthesis tokens, but instead only isolated examples, with one case observed in a wh-question (tuns-whq1-f5) and one in a vocative (tuns-voc-f3). In each case, these outlier tokens were produced with a different prosodic contour than that observed in all other parallel utterances (which typically ended in a fall H* L-L% or rise L* H-H%). In the next section therefore we turn to the detail of the prosodic contour, to see whether this is the primary determiner of utterance-final vowel epenthesis in TA.

4.2 Prosodic contour

To determine the extent to which utterance-final vowel epenthesis in TA is conditioned by prosodic context, firstly, the read speech yes-no questions were prosodically annotated. Table A in the Appendix sets out in full the incidence of vowel epenthesis according to nuclear contour (last pitch accent + boundary tones) and by speaker. The most common prosodic contour observed on the last lexical item in read speech yes-no questions was a complex rise-fall pitch contour comprised of a rising nuclear pitch accent L*+H followed by a complex falling edge tone, analysed here as H-L%. There were no instances of vowel epenthesis in utterances produced with a simple pitch contour with no change of direction, such as a rise (L*+H H-H%) or fall (H* L-L%). Among the vocatives, the only instance of a complex pitch contour was found in the one token in which vowel epenthesis was also observed (tuns-voc-f3), illustrated in Figure 5 below.

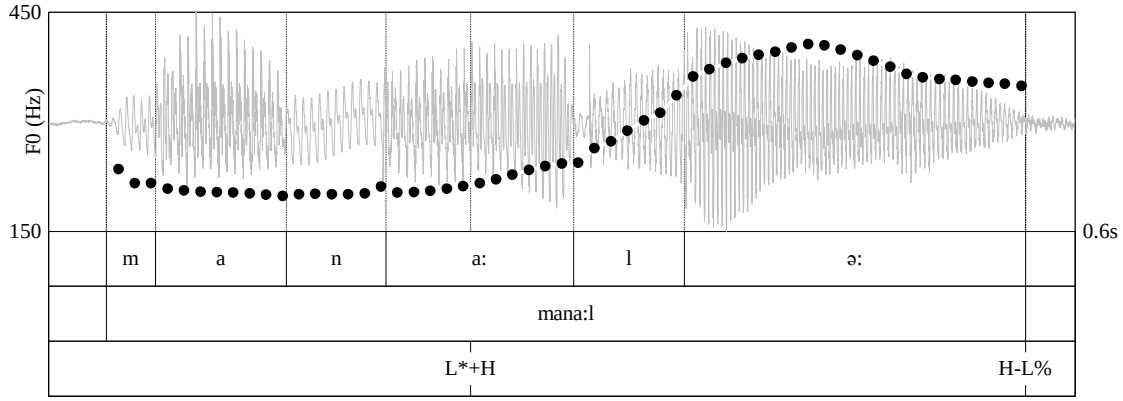

Fig 5. Vocative from read speech data produced with vowel epenthesis [tuns-voc-f3].
/maˈna:l/ [maˈna:lə:]
(name.f)
Manal!?

It seems therefore that vowel epenthesis is primarily conditioned by prosodic contour shape, since no instances at all of vowel epenthesis are observed on utterances which bear a simple rising (H* L-L%) or falling (L* H-H%) contour.

Nevertheless, it is not the case that a complex pitch contour necessarily triggers the presence of vowel epenthesis. Figure 6 below shows plots of mean f0 for female and male speakers and for tokens produced with or without vowel, measured at ten points through the last lexical item, in all read speech utterances transcribed with a L*+H H-L% contour (N=54). The plots show a clear rise to a peak (analysed here as a L*+H accent) following by a small fall at the right edge of the word (analysed here as H-L%), regardless of whether or not a vowel is produced. Note that the position of the H peak relative to the stressed syllable of the accented word does not vary according to the prosodic structure of the target word (here, varied in terms of the position of stress in the word), hence analysis of that peak as a H- phrase accent. In contrast, the position of the elbow at the start of the rise towards that peak does shift somewhat earlier as the position of the stressed syllable moves earlier in the

word, so that the rise is found on the postaccentual syllable, whenever present; this rising movement is thus analysed as due to a trailing H tone which is part of a bitonal pitch accent associated with the accented syllable (L*+H).

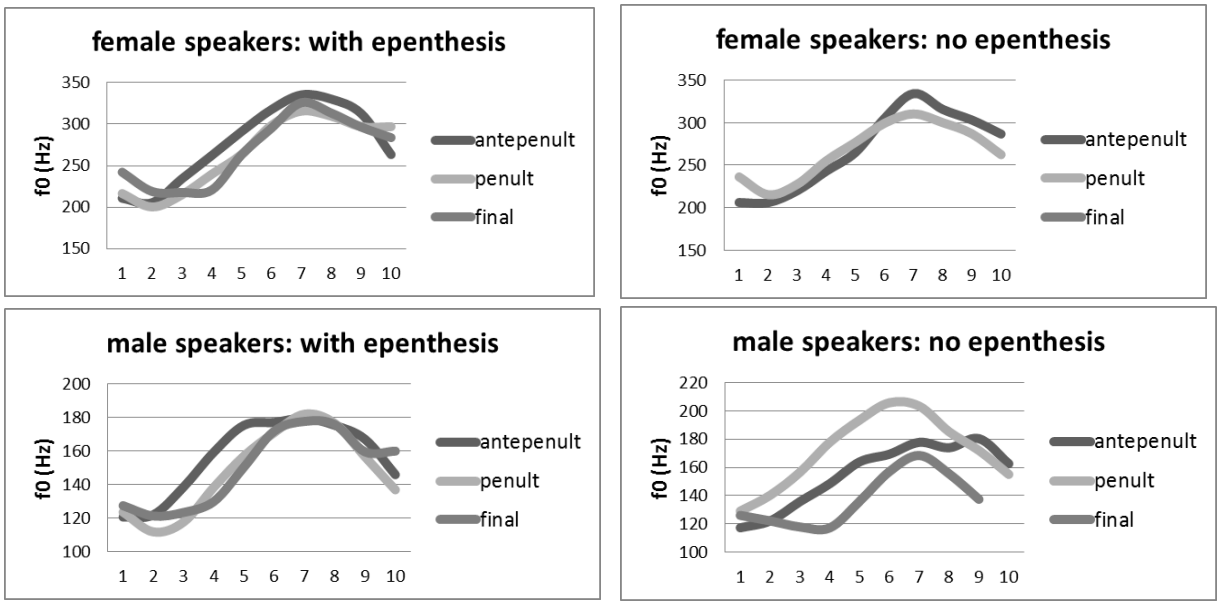

Figure 6: Mean values of f0, at ten measuring points through the last lexical item, in read speech yes-no questions produced with L*+H H-L% contour by female and male speakers, by presence of vowel epenthesis or not, and by position of the stressed syllable in the word.

To summarise, a key conditioning factor for vowel epenthesis in TA yes-no questions appears to be the presence of the complex rise-fall nuclear pitch contour (here, L*+H H-L%), as no epenthesis was observed in tokens bearing other contour types. However, not all utterances produced with the complex pitch contour show vowel epenthesis. In the next two sections we explore other factors which may further influence the occurrence of vowel epenthesis.

4.3 Segmental and metrical context

A number of factors are reported to constrain the incidence of vowel epenthesis, in the prior literature on text-tune adjustment in southern European languages (as summarised in Table 1 above). These include the metrical structure of the target word (with more epenthesis observed on monosyllables) and the segmental content of the target word (with more epenthesis after a sonorant and/or voiced consonant). An exhaustive list was created of all lexical items found in utterance-final position in all yes-no questions analysed (N=108), including data from read speech (N=68) and spontaneous speech (N=40). A count was then made of the incidence of vowel epenthesis in these lexical items, by different relevant factors.

The only instances of monosyllables in the dataset are from spontaneous speech (e.g. as in Fig 3 on the word [tˤu:l] ‘straight ahead’), as no utterance-final monosyllables were elicited in the read speech scripted dialogue task. This results in a small number of tokens (N=5) for analysis here, but the majority of these monosyllables (4 out of 5) show vowel epenthesis. Nevertheless, vowel epenthesis is frequently observed on polysyllabic words, as shown in Figure 7 below. If vowel epenthesis in TA were a pure case of tune-text adjustment, serving to provide sufficient segmental material for realisation of all intonational tones, we

might expect vowel epenthesis to be much less common at the end of polysyllabic words, where plentiful segmental material is available, and this is not the case. The incidence of vowel epenthesis also seems not to be conditioned by the position of stress in the word, as shown in Figure 8 below, and as can also be seen in the left hand plots in Figure 6 above (among the tokens in which vowel epenthesis was observed, there is a roughly equal incidence of vowel epenthesis in tokens regardless of the position of stress in the utterance-final word). Again, if tonal crowding were the motivation for text-tune adjustment, we might expect more vowel epenthesis in words with final stress, and this is not the case. Finally, there appears also to be no categorical effect of final segment type (that is, the type of segment appearing at the end of the utterance final word). Vowel epenthesis is seen just as frequently in tokens in which the utterance-final word ends in an obstruent (and whether voiced or voiceless e.g. [maw.ˈʒuːd] 'present' or [ʕliːk] 'to you') as in words ending in a sonorant or vowel, as shown in Fig 9 below.

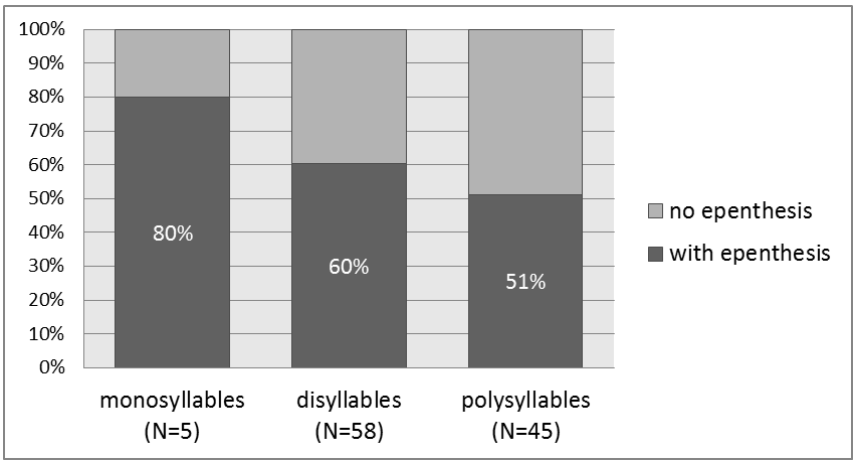

Fig. 7: Vowel epenthesis by number of syllables in the word, in all yes-no questions (N=108).

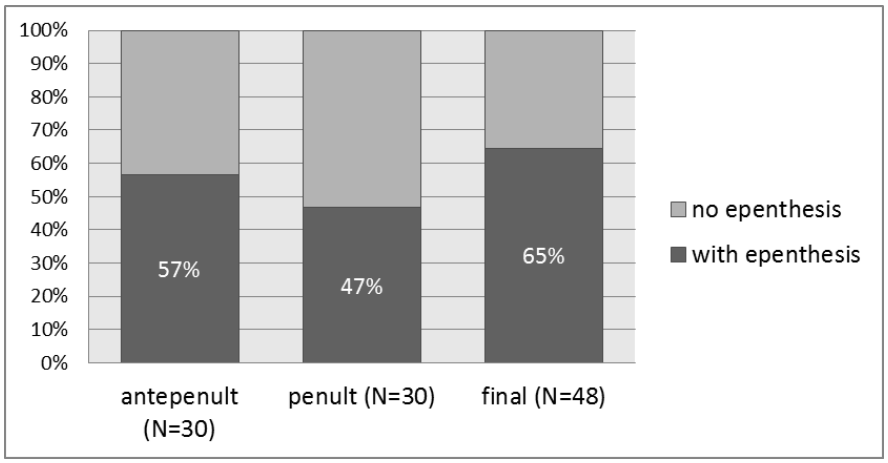

Fig. 8: Vowel epenthesis by position of stress in the word, in all yes-no questions (N=108).

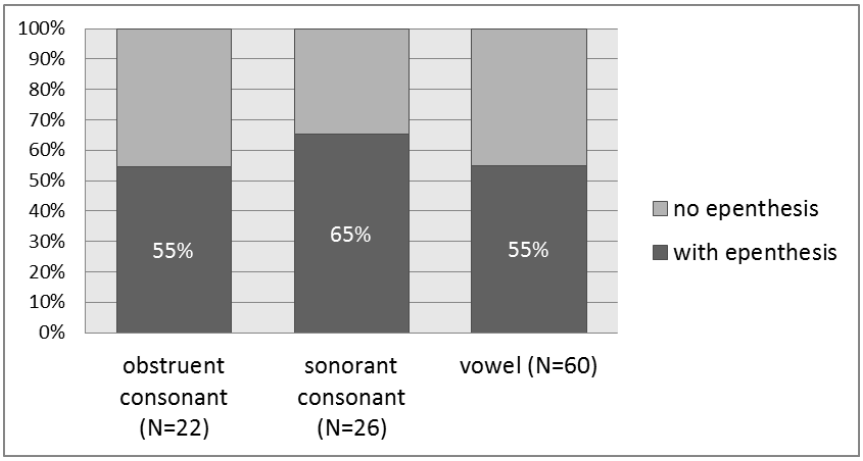

Fig. 9: Vowel epenthesis by final segment type, in all yes-no questions (N=108).

There is, however, an effect of final segment type on acoustic properties of the epenthesised vowel, as shown in Figure 10. Epenthetic vowels following a consonant tend to be longer and louder than those following a vowel. Following a vowel, the quality of the epenthetic vowel is also highly variable, tracking the quality of the vowel it follows (though in all cases some change in formant structure must have been detected for it to be labelled); there is more limited variation in the quality of vowels along an [a]~ [ə]~[u] continuum after a final sonorant, which overlaps fully in distribution with that of vowels after a final obstruent.

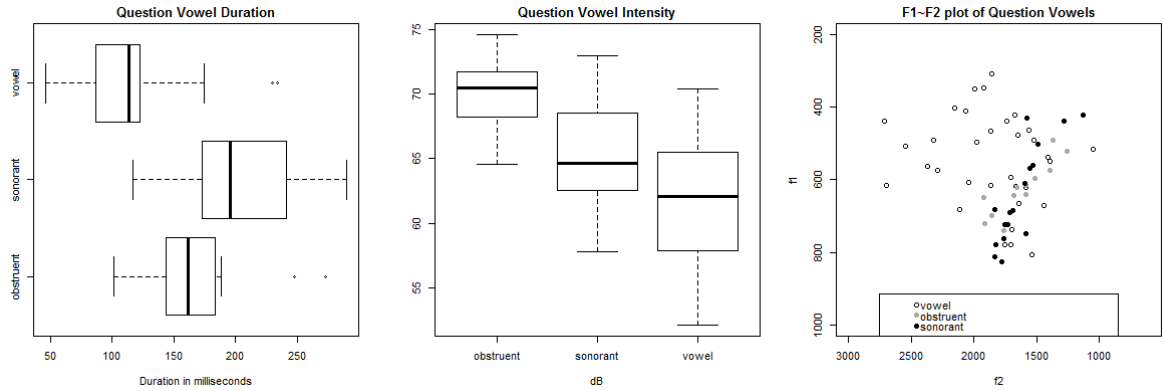

Fig. 10: All epenthetic vowels, split by preceding utterance-final segment type: boxplots of vowel duration (left) and mean intensity (centre), and f1/f2 at vowel midpoint (right).

We explore below whether there are probabilistic effects of these factors on the incidence of vowel epenthesis, as was observed in Bari Italian (Grice et al., 2015), but from these descriptive results there is no indication that vowel epenthesis in TA is a categorical tonal crowding effect, since there is no obvious conditioning due to segmental or metrical context.

4.4 Speaker variation

The incidence of vowel epenthesis appears to vary according to gender in TA, as illustrated in Figure 11 below, with female speakers (f1-f6) displaying on average more epenthesis (76%) than male speakers (m1-m6) (32%). A full breakdown by item and speaker is provided in Table B in the Appendix. In this read speech data, one female speaker (f4) produces an

utterance-final vowel in all of her yes-no questions, whereas there are two male speakers (m2, m6) who never produce any utterance-final vowels. Variation by gender has not been reported in other studies of utterance-final vowel epenthesis.

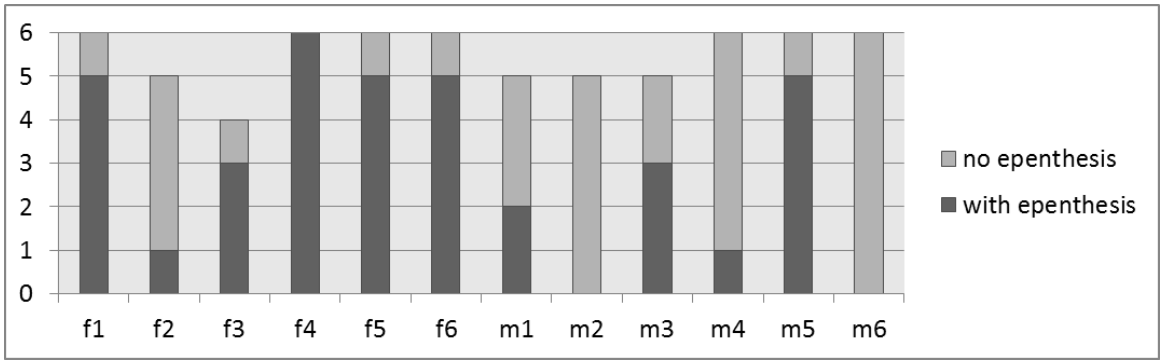

Figure 11: Number of read speech yes-no questions in which vowel epenthesis was observed or not, by speaker. Female speakers = f1-f6; male speakers = m1-m6.

This pattern is mirrored in the set of yes no questions extracted from spontaneous speech, as shown in Figure 12 below. Although there are an uneven number of tokens produced across speakers, due to the spontaneous nature of the elicitation task, nevertheless, the same two male speakers (m2, m6) produce no vowel epenthesis, and there is in general a higher incidence of vowel epenthesis in spontaneous speech produced by female speakers (78%) than by male speakers (38%).

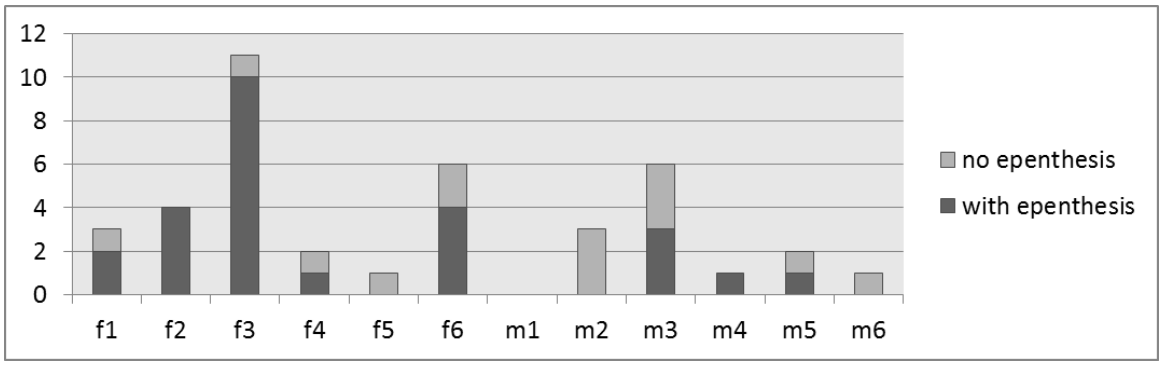

Figure 12: Number of spontaneous speech yes-no questions in which vowel epenthesis was observed or not, by speaker. Female speakers = f1-f6; male speakers = m1-m6.

Overall, in all yes-no questions, female speakers produce vowels in 77% of tokens and male speakers produce vowels in 32% of tokens, as shown in Figure 13.

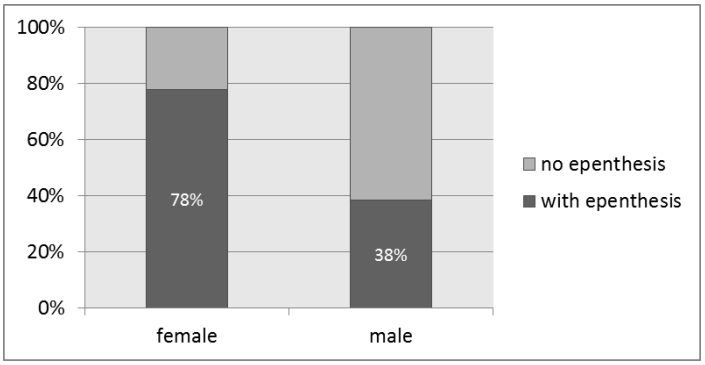

Figure 13: Incidence of vowel epenthesis in all yes-no questions (N=108), by gender.

As well as producing more vowel epenthesis overall, female speakers produced epenthetic vowels which tended to be longer and louder than those produced by male speakers, and with fronter, lower vowel quality (closer to [a] than [u]), as shown in Figure 14.

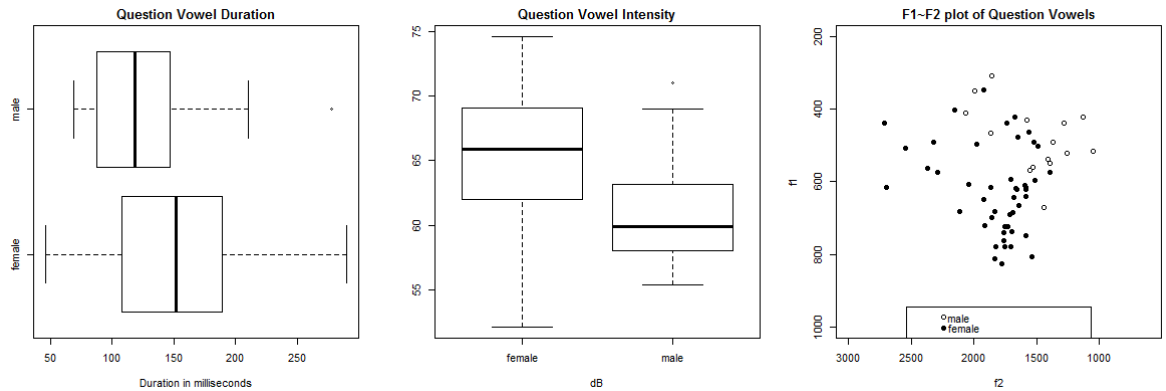

Fig. 14: All epenthetic vowels, split by gender of speaker: boxplots of vowel duration (left) and mean intensity (centre), and f1/f2 at vowel midpoint (right).

Since the quality of the epenthetic vowel after a word-final vowel is highly variable across all speakers, Figure 15 shows F1/F2 for epenthetic vowels after a final consonant only, split by gender of the speaker, and by region of birth of speaker's parents (north or south). From the present limited sample it is not possible to tease apart fully the effects of gender vs. region on epenthetic vowel quality, and we highlight this as a topic ripe for further investigation.

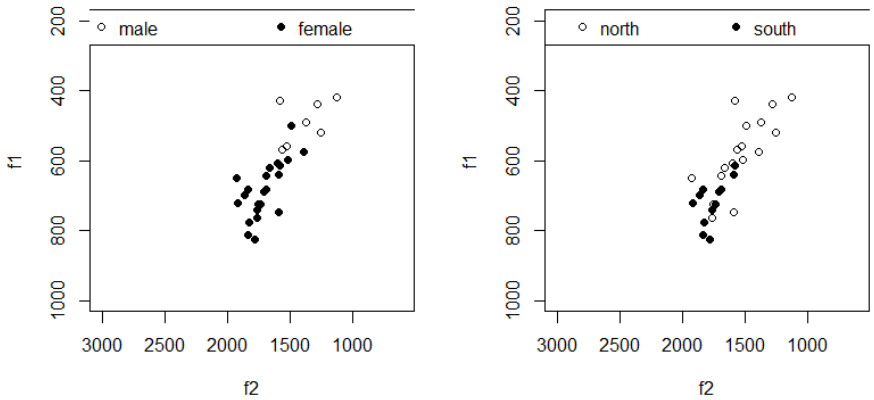

Fig. 15: Measures of f1/f2 at vowel midpoint of epenthetic vowels following a final consonant, split by gender (left) and region of birth of speaker's parents (right).

Finally, recall, from Table A (in the Appendix), that male speakers tend to use simple pitch contours, and in particular rises, more than female speakers, as illustrated in Figure 16.

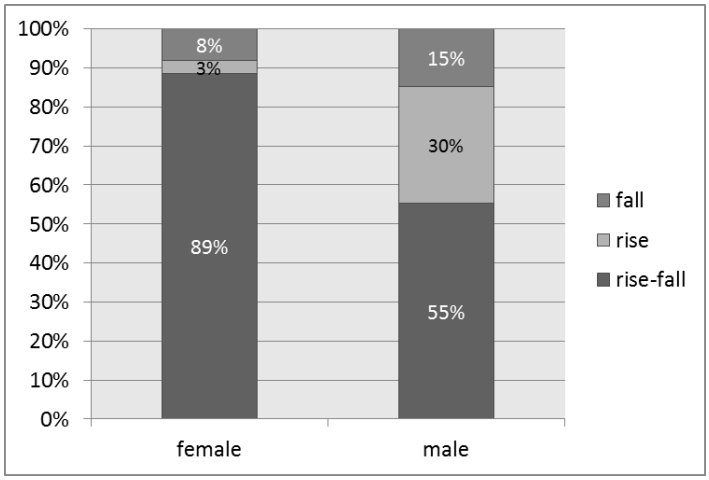

Figure 16: Incidence of different contour types in all yes-no questions (N=108), by gender.

Some of the variation in the data could thus perhaps be explained in terms of a trading relation between the choice to produce an epenthetic vowel and the choice to produce a complex pitch contour L*+H H-L% vs. a simple rise H-H%. Indeed, it may be that male speakers are producing truncated forms of the rise-fall contour, neutralising the contrast between L*+H H-L% and L* H-H%. However, the observed co-variation between choice of contour shape and vowel epenthesis is not total, since male speakers produce complex pitch contours in 55% of tokens (Fig. 16) but produce a vowel in only 32% of tokens (Fig. 13). This is seen also in Figure 6, which includes tokens (on the left hand side) which bear a complex pitch contour but are realised without a vowel, by male as well as female speakers.

Overall then, within the subset of the data which bears a complex pitch contour, there is still some residual variation. A Classification by Regression Trees (CART) analysis was used to determine whether this residual variation was significant (Baayen, 2008 cf. Grice et al. 2015), taking all factors into account (contour shape, final segment, position of stress, number of syllables, speech style and gender). However. a cost complexity pruned CART tree showed only marginal improvement (13% misclassification) over a baseline model (17% misclassification) in which only choice of contour is used as a predictor. The only statistically significant predictor of utterance-final vowel epenthesis in the present sample of yes-no questions in TA is thus the choice of prosodic contour.

4.5 Summary

Utterance-final vowel epenthesis in Tunisian Arabic (TA) yes-no questions was found to be variable across both speakers and items. The primary conditioning factor (present in all observed cases) is a complex pitch contour (analysed here as L*+H H-L%). However, not all utterances bearing a complex pitch contour are realised with a final epenthetic vowel. An important secondary conditioning factor appears, on the surface, to be gender, with female

speakers more likely to produce an epenthetic vowel than male speakers, though this effect is not statistically significant, in the present dataset. The epenthetic vowels produced by females tend also to be longer, louder and more peripheral in vowel quality. An potential trading relation between vowel epenthesis and boundary is observed which opens up the possibility that the gender-driven variation in TA may in fact be between a truncation strategy and an epenthesis strategy. Nevertheless, these two 'strategies' are observed only in yes-no questions, and only in utterances which bear no obvious evidence of tonal crowding (such as polysyllables with non-final stress), in contrast to patterns observed in other languages.

5 Discussion

Analysis of data in Tunisian Arabic (TA) from a corpus of both read and spontaneous speech indicates that utterance-final vowel epenthesis is frequently observed in yes-no questions, but only in exceptional cases in other sentence types. All instances of vowel epenthesis in the present data are found in tokens bearing a complex pitch contour (here, L*+H H-L%), and the presence of this complex pitch contour is a primary conditioning factor. This pattern is thus different from that observed in Alentejo Portuguese, for example, in which vowel epenthesis co-occurs with a variety of different tonal contours.

There are tokens in the present dataset bearing a complex pitch contour which are realised without vowel epenthesis however (N=18), but none of the constraining factors observed in SEP and Bari Italian - such as sonority of the final segment or the number of syllables in the word, nor an equivalent factor of position of stress in the word - are robust predictors of the presence or absence of vowel epenthesis in TA.

It is puzzling that utterance-final vowel epenthesis is observed in TA in contexts which appear to provide sufficient segmental material on which to realise a complex pitch contour. Under a standard AM framework phonological account, it is hard to characterise this as a pure case of 'text-tune adjustment'. There are some indications of a trading relation between truncation (somewhat preferred by male speakers) and vowel epenthesis (preferred by female speakers), which, if confirmed in further data in future, might still suggest that an element of tonal crowding is nevertheless at work.

A potential alternative explanation as to the origins of this TA epenthetic vowel pattern would be historical migration patterns into North Africa from the Arabian Peninsula, resembling similarly discontinuous patterns of syntactic variation across Arabic dialects (Lucas & Lash, 2010). Holes (2016) reports the existence of a clitic suffix [-ə] which attaches to a word or sentence to create a yes-no question, in the Baḥarna dialect of Bahrain (BBA), and in sedentary dialects in Oman. In BBA, the /-ə/ clitic is realised as [-hə] or [-jə] after a vowel or glide, and there are parallel cases in the present dataset, in that the utterance-final word [ˈjamani] is realised by some speakers as [ˈjamanijə].[2] Holes notes that rising intonation is used in BBA in tag questions, which mirror the vowel clitic in meaning, and confirms that the /-ə/ clitic is also accompanied by a characteristic rise-fall tune (Holes, p.c.). In TA, we have assumed thus far that it is the complex boundary which conditions vowel epenthesis, rather than the reverse, but it is tempting to consider the possibility of treating the TA

[2] For one speaker (f4), there is accompanying shift of stress to the penult yielding [jamaˈnijə], showing that for this speaker the appended vowel is integrated into the phonological word, for purposes of stress assignment.

epenthetic vowel as a question particle, by analogy with the BBA case. Another potential source of such a particle would be from contact with Tamazight, since an interrogative clitic [a] polar question particle is reported for Zwara Tamazight (Gussenhoven, 2015), a variety of Tamazight spoken in western Libya close to the border with Tunisia. However, in Zwara Tamazight, this particle is expressly reported to be accompanied only and always by a final falling intonational contour, so the match with the TA pattern is only partial.

A further puzzle is that, if we treat these TA epenthetic vowels as question particles, arising in a historical contact scenario, adoption of vowel epenthesis is quite clearly not a case of linguistic innovation, since the BBA phenomenon is reported by Holes only in the speech of an older generation of speakers, and we can assume that any Tamazight substrate influence pre-dates the arrival of Arabic to Tunisia. If vowel epenthesis is, then, a conservative feature of TA we might expect it to be used more by male speakers than female, rather than the reverse as observed here. Indeed, we note that the one male speaker (m5) who produces rather more epenthetic vowels than other male speakers, also displays consistent palatalization in his speech, which is a known innovative feature in Egyptian Arabic, initially adopted by female speakers (Haeri, 1996) but which we now also observe in the speech of young male EA speakers, for example in IVAr corpus data, (cf. also Youssef, 2016). A similar *sociolinguistic* distribution of schwa epenthesis has been reported for Parisian French (Hansen, 1997) – that is, used more frequently by young female speakers – and we might thus entertain the possibility that French is the source of the pattern observed in our data here, since French is spoken bilingually alongside TA by most educated Tunisians, and indeed by all of the participants in our study. However, although utterance final vowel epenthesis is accompanied by a rise-fall contour more frequently than other contours, in Hansen's (1997) Parisian data, the pattern is not confined to yes-no questions, but is instead used more often to express emphasis or indignation, or in a non-final position in the speaker's turn; and, in addition, the incidence of schwa epenthesis varied according to the type of final segment. This contrasts with the TA pattern observed in our data whereby an epenthetic vowel is appended only in utterances bearing the rise-fall contour, and almost exclusively in yes-no questions, and showed no categorical variation dependent on the utterance-final segment.

6 Conclusion

This paper highlights, for the first time, the existence of a phonological pattern of prosodically conditioned vowel epenthesis in TA yes-no questions. Our exploration here of the conditioning factors governing the incidence of utterance-final vowel epenthesis in TA suggests, firstly, that this is not a case of 'text-tune' adjustment at all, but rather a 'question vowel' particle of some type. This in turn suggests, secondly, that we should set aside the possibility that the pattern arises due to contact from languages across the Mediterranean (Portuguese or Italian), and instead focus on older contact scenarios via historical migrations and/or substrate influence from varieties of Tamazight. We hope that investigation of the intonation patterns of neighbouring dialects of Arabic along the coast of North Africa might yield further insights as to the spread of the phenomenon and its origins. Similarly, a perceptual study might determine whether yes-no questions produced with a 'truncated' boundary rise are perceptually equivalent to those produced with a complex pitch contour, and what role the presence or absence of an appended vowel plays in the interpretation of

utterances. Lastly, but by no means least, our findings suggest that vowel epenthesis in yes-no questions would be a useful variable for inclusion alongside others in sociolinguistic analysis of TA, to determine what other sociolinguistic factors may be relevant, how widespread the pattern is in Tunisia and whether the incidence of the pattern is changing.

Acknowledgements
The IVAr database project was funded by a grant to the author from the UK Economic and Social Research Council (ES/I010106/1). The IVAr project is indebted to the TA speakers who participated in data collection and gave generously of their time and language expertise. I am grateful to: Rana Almbark who carried out data collection in Tunis for the IVAr project; to Nadia Bouchhioua who hosted our visit to Tunis; to Zaineb Chihi who, as a native speaker of TA, ran all of the data elicitation sessions; and to Sihem Saadi who transcribed the data. Pitch trace figures were produced using a script kindly shared by Francesco Cangemi. Thanks are due to the editors, and to three anonymous reviewers, for their very helpful suggestions.

Appendix

	ynq1	ynq2	ynq3	ynq4	ynq5	ynq6
f1	complex	complex	complex	complex	complex	complex
f2	complex	complex	complex	complex		complex
f3	complex	complex	complex		complex	complex
f4	complex	complex	complex	complex	complex	complex
f5	complex	complex	complex	complex	complex	*rise*
f6	complex	complex	complex	complex	**fall**	complex
m1	complex	complex	complex	complex		complex
m2	*rise*		*rise*	*rise*	*rise*	*rise*
m3	complex	complex	complex	*rise*	complex	*rise*
m4	complex	complex	complex	*rise*	**fall**	*rise*
m5	complex	complex	complex	complex	complex	complex
m6	complex	**fall**	complex	complex	**fall**	complex

Table A: Prosodic contour on last lexical item in read speech *ynqs*, by target utterance and speaker.
Dark shaded cells: with vowel epenthesis; light shaded cells: no vowel epenthesis; empty cells: missing token.
Key: complex = L*+H H-L%; *rise* = L*+H H-H%; **fall** = H* L-L%.

	f1	f2	f3	f4	f5	f6	m1	m2	m3	m4	m5	m6	*Total*
ynq1	✓	×	×	✓	✓	✓	✓	×	×	×	✓	×	6
ynq2	✓	×	✓	✓	✓	✓	×	--	✓	×	✓	×	7
ynq3	×	×	✓	✓	✓	✓	×	×	✓	✓	✓	×	7
ynq4	✓	×	--	✓	✓	✓	×	×	×	×	×	×	4
ynq5	✓	--	✓	✓	✓	×	--	×	✓	×	✓	×	6
ynq6	✓	✓	✓	✓	×	✓	✓	×	×	×	✓	×	7
Total	5	1	4	6	5	5	2	0	3	1	5	0	

Table B: Incidence of vowel epenthesis in read speech *ynqs*, by target utterance and speaker.
Key: ✓ vowel observed; × no vowel observed; -- denotes a missing token due to disfluency.

References

Aloulou, M. (2003). *The intonation patterns of Tunisian Arabic and English: a comparative study.* MA dissertation, Institut Supérieur des Langues de Tunis, University of Carthage, Tunis.

Anderson, A., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S. et al. (1991). The HCRC Map Task Corpus. *Language and Speech, 34*, 351-366.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R* Cambridge University Press.

Beckman, M., & Elam, G. A. (1997). *Guidelines for TOBI Labelling (version 3.0 1997).* The Ohio State University Research Foundation.

Beckman, M., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic Typology: The phonology of intonation and phrasing* (pp. 9-54). Oxford: Oxford University Press.

Boersma, P., & Weenink, D. (2015). *Praat: doing phonetics by computer (Version 5.4.09) [http://www.praat.org]*.

Bouchhioua, N. (2008). *The acoustic correlates of stress and accent in TunisianArabic: a comparative study with English.* (PhD dissertation, University of Carthage).

Cruz, M. (2013). *Prosodic variation in European Portuguese: phrasing, intonation and rhythm in central-southern varieties.* PhD dissertation, Universidade de Lisboa.

Eckert, P., & McConnell-Ginet, S. (2003). *Language and gender*. Cambridge: Cambridge University Press.

Frota, S. (2002). Nuclear falls and rises in European Portuguese: a phonological analysis of declarative and question intonation. *Probus, 14*(1), 113-146.

Frota, S., Cruz, M., Fernandes-Svartman, F., Collischonn, G., Fonseca, A., Serra, C. et al. (2015). Intonational variation in Portuguese: European and Brazilian varieties. In P. Prieto & S. Frota (Eds.), *Intonation in Romance* (pp. 235-283). Oxford: OUP.

Frota, S., & Prieto, P. (2015). *Intonation in Romance*. Oxford: OUP.

Ghazali, S., Hamidi, R., & Knis, K. (2007). Intonation and rhythmic patterns across the Arabic dialects continuum. In E. Benmamoun (Ed.), *Perspectives on Arabic Linguistics: Papers from the annual symposium on Arabic Linguistics Volume XIX: Urbana, Illinois April 2005* (pp. 97-122). Amsterdam: John Benjamins.

Ghazali, S. (1973). *Tunisian Arabic and French interference with English: word stress and the phoneme.* MA dissertation, University of Texas at Austin.

Grabe, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics, 26*(2), 129-143.

Grabe, E. (2004). Intonational variation in urban dialects of English spoken in the British Isles. In P. Gilles & J. Peters (Eds.), *Regional Variation in Intonation* (pp. 9-31). Linguistische Arbeiten. Tuebingen: Niemeyer.

Grice, M., Savino, M., Caffo, A., & Roettger, T. B. (2015). The tune drives the text: schwa in consonant-final loan words in Italian. In The Scottish Consortium for ICPhS (Ed.), *Proceedings of the 20th ICPhS, Glasgow* (pp. Paper 0381).

Gussenhoven, C. (2015). Zwara Tamazight (Berber, Afroasiatic). *Ms submitted to Journal of the International Phonetic Association*.

Haeri, N. (1996). *The sociolinguistic market of Cairo: gender, class and education.* London: Keegan Paul International.

Hansen, A. B. (1997). Le nouveau [ə] prépausal dans le français parlé à Paris. In *Polyphonie pour Ivan Fonagy*. L'Harmattan.

Hellmuth, S. (2013). Phonology. In J. Owens (Ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 45-70). Oxford: Oxford University Press.

Hellmuth, S., & Almbark, R. (2017). *Intonational Variation in Arabic database*. Retrieved from: http://ivar.york.ac.uk/

Hellmuth, S., & Almbark, R. (in preparation). *Intonation in spoken Arabic dialects.*

Holes, C. (2016). *Dialect, Culture, and Society in Eastern Arabia: Phonology, Morphology, Syntax and Style* Brill.

Hutchby, I., & Wooffitt, R. (2008). *Conversation Analysis* (2nd ed.) Polity Press.

Knis, K. (2004). *?atharu ?allahajaat ?al?arabiyya fi tanghiim ?al-fuSHaa [The effects of colloquial dialects on the intonation of Formal Arabic]* MA dissertation, Institut Supérieur des Langues de Tunis, University of Carthage, Tunis.

Kulk, F., Odé, C., & Woidich, M. (2003). The intonation of colloquial Damascene Arabic: a pilot study. *IFA Proceedings, 25*, 15-20.

Labov, W. (2001). *Principles of linguistic change Volume 2: Social factors*. Oxford: Blackwell.

Ladd, D. R. (2008). *Intonational phonology* (Second ed.). Cambridge: Cambridge University Press.

Lucas, C., & Lash, E. (2010). Contact as catalyst: the case for Coptic influence in the development of Arabic negation. *Journal of Linguistics, 46*(02), 379-413.

Ng, E.-C. (2013). Paragoge as an indicator of language contact. Poster presented at LSA 2013 (Linguistic Society of America), Boston, 3-6 Jan 2013..

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from ISBN 3-900051-07-0: http://www.R-project.org

Saadi, S. (2014). Regional Effects on the Intonation Patterns of Three Different Groups of Tunisian EFL Learners (MA dissertation, University de la Manouba, Tunis.).

Youssef, I. (2016). Palatalization in educated Cairene Arabic. *Nordlyd: Tromsø University Working Papers on Language & Linguistics, 42*(1), 21-31.