The tune drives the text - Schwa in consonant-final loan words in Italian

Martine Grice, Michelina Savino, Alessandro Caffo', Timo Roettger

XXX
XXX

ABSTRACT

In Italian, consonant-final loan words are reportedly produced with or without a final schwa. This paper reveals that variation in the presence of this schwa is dependent on a number of factors, including the metrical structure of the target word and the voicing of the consonant. Crucially, it is also conditioned by intonation: Schwa is more likely to occur – and is acoustically more prominent – when the intonational tune is complex or rising, as opposed to falling. Schwa epenthesis can thus be seen as facilitating the production of functionally relevant tunes.

Keywords: Italian, tune-text association, schwa, compression, epenthesis

1. INTRODUCTION

Autosegmental-metrical approaches to intonation involve the association of tones (the tune) with syllables, words and phrases (the text). This association does not *a priori* privilege one of these two levels. If the text is suboptimal for bearing the tune, adjustments can be made to either the tune (reducing its complexity) or to the text (lengthening the tune-bearing portions of voiced, and therefore tone bearing, material). In this paper we are concerned with adjustments in the text to accommodate the tune, referred to as a local *rate adjustment* [7] or *compression* [1], [8].

In Bari Italian, the variety of Italian investigated in this study, such adjustments have been reported in yes-no questions, which in read speech are typically realised with an accentual rise (L+H*) followed by a fall-rise, represented by a L-H% boundary tone sequence. (Note that in spontaneous questions the accentual rise is generally followed by a fall L-L% [19], [9]). If a phrase-final accented syllable bears the rise-fall-rise, it has been found to be considerably lengthened [11,16] as compared to neutral statements, where there is a simple fall in the lower portion of the speaker's range (H+L* L-L%). These studies investigate words with final stress that end in a vowel, such as bambù (/bamˈbu/ 'bamboo').

The question arises, then, as to what happens when a word with final stress ends in a consonant. Although Italian has a very limited number of consonant-final words in its native vocabulary, the language has incorporated a great number of such words in recent years, including many proper nouns [6]. Crucially, their pronunciation is subject to variation [2,13,17, *inter alia*]. Monosyllabic words such as 'chat' can retain the structure of the donor language (in this case English) with the consonant in final position, [tʃat], or the consonant can be followed by a mid central vocoid (henceforth schwa): [ˈtʃatːə].

Studies on Italian generally analyse this non-lexical word-final schwa as an epenthetic vowel, rather than a phonetic artefact. One strong argument for its phonological status is that it goes hand in hand with a lengthened (geminated) consonant. However, its phonetic properties appear to be prone to inter- and intra-speaker variability, as evidenced in the transcriptions [ə] and [ᵊ] (for some varieties also [e]; cf. [17]). These observations indicate that the insertion of schwa is somewhat gradual in nature, and might be taken as evidence for a less entrenched phonological status.

Returning to Bari Italian, if intonation can condition adjustments to the duration of phrase final lexical vowels, it is also possible that it can play a role in conditioning the epenthesis of schwa. Our main research question is, thus, whether intonation is the driving force behind schwa epenthesis. Since the language has intonation contours other than those in polar questions and neutral statements, we also investigate whether other intonation contours, specifically those in different positions in lists ([24], [25]), can play a role in determining epenthesis, too.

Moreover, consonant final loan words are not always monosyllabic. If they are polysyllabic, the final syllable is rarely stressed (the final consonant being treated as extrametrical). Word-final schwa epenthesis in polysyllabic words has received little attention to date, but has been reported [5]. The question here is whether the metrical structure also plays a role in conditioning the likelihood of epenthesis.

To explore these factors, we elicited intonation contours with differing complexity and direction of pitch change on proper nouns (person names), a common source of loans in the language. Pitch accent placement was elicited on the final and penultimate syllable of the phrase through the choice of monosyllables and disyllabic trochees, respectively, as target words. In the former, the pitch accent and boundary tones crowd together on one syllable, in

the latter they can potentially spread over two syllables.

2. METHOD

2.1. Speech material

Target words consisted of 10 monosyllabic and 6 disyllabic names (*Bill, Moll, Tim, Dan, Dag, Fred, Chris, Jeff, Matt, Dick* for the monosyllables, and *Caleb, Colin, Carol, Edith, Derek, Dennis* for the disyllables).

These target words were elicited in five Prosodic Conditions: (polar) *Questions* and (neutral) *Statements*, and in three distinct positions in lists consisting of six names: *Non-Final, Pre-Final* and *Final*. For each prosodic condition, target words were produced with an appropriate context, as follows (see also Figure 1):

(1) Question:
Ha chiamato [target name]? 'Did [target name] call?'
No, ha chiamato [name]. 'No, [name] called.'
e.g. Ha chiamato Jeff?

(2) Statement (answer):
Chi ha chiamato? 'Who called?'
Ha chiamato [target name]. '[target name] called.'
e.g. Ha chiamato Jeff.

(3) NonFinal, PreFinal and Final:
Ecco la lista dei nomi: 'Here is the list of names:'
[name], [name], [name], [NF target], [PF target], [F target].'
e.g. Dan, Colin, Dennis, Moll, Matt, Fred.

Thus, there were 160 items in total (16 target words x 5 Prosodic Conditions x 2 repetitions).

2.2. Participants and procedure

Ten native Bari Italian speakers participated in the recording session on a voluntary basis. They were all female (aged 22-29) and undergraduate students of Psychology at the University of Bari, without a background in phonetics or prosody.

Speakers were seated in front of a computer screen, wearing a headset microphone (AKG C520) connected to a Marantz PMD 661 digital recorder. Each stimulus was presented on the screen with its context, and speakers were instructed to read first silently and then aloud at a normal pace and in a natural way. They were allowed to repeat the target phrase if they felt it was inappropriate in the context, or in case of disfluency. Speakers were also allowed to take a break any time they needed, and at least once every 20 stimuli.

2.3. Analysis

Target words were manually segmented and annotated with Praat 5.4 [3]. The labelling of potential schwa was not always straightforward. We thus adopted a liberal approach, labelling as a schwa any interval presenting periodic vibrations accompanied by a local increase in the signal energy at the consonantal release, and/or any interval after the consonantal release with formant structure or energy in the F2/F3 region characteristic of vowels.

Data were statistically analysed, using R [15] looking at the question-statement and list data sets separately. Since this is the first look at a controlled and balanced data set produced by a rather homogenous group of speakers, we were able to explore the data set using random forests [4], implemented by the party R package [12,20,21]. For a discussion of these techniques in the context of linguistics and sociolinguistics, see [22]. Random forests is a data mining technique used for classification. It is a so-called "ensemble method" because a multitude of decision trees is constructed (500 in this case). Each tree takes a set of variables and sees which variable best splits the data according to a particular criterion. Each tree is built on a random subset of variables and data. The final classification is based on the overall ensemble of trees.

Random forests allow us to identify which factors are independently relevant for determining the presence vs. absence of schwa. The following factors were included in the analysis: Factors capturing idiosyncratic properties of SPEAKERS and WORDS; factors capturing differences in the identity of the final consonant coded as ± VOICED, and ± SONORANT; a factor capturing metrical characteristics of the target word coded as SYLLABLE NUMBER (including monosyllables and disyllabic trochees); and most importantly, factors capturing prosodic characteristics of the contour coded as PROSODIC CONTEXT, reflecting sentence modality: question and statement in the question-statement data set, and the position in the list: non-final, prefinal, and final in the list data set.

3. RESULTS AND DISCUSSION

3.1. Contours

The intonation contours in both data sets corresponded to our expectations based on previous studies, which were also based on read speech. In the question-statement dataset, questions were produced predominantly with a *rise-fall-rise* (L+H*L-H%) and occasionally with a *rise-fall* (L+H* L-L%), whereas statements had a *low fall* (H+L* L-L%). See Figure 1 (a), (b) for examples.

In the lists, non-final items were produced with a *low rise* (L* L-H%), pre-final items with a *high rise* (H* H-^H%), and final items with a *low fall* (H+L* L-L%, see Figure 1 (c)) as in the statements in the question-statement data set. (Note that although the distinction between H+L* and L* is neutralised in cases where a phrase contains no unstressed syllables before the stressed syllable, by convention we retain the more phonologically oriented H+L* here.)

Figure 1: Representative waveforms and F0 contours for (a) a question, (b) a statement, and (c) a list. Segmental annotations in SAMPA.

(a) Question

(b) Statement

(c) List: NonFinal, PreFinal, Final
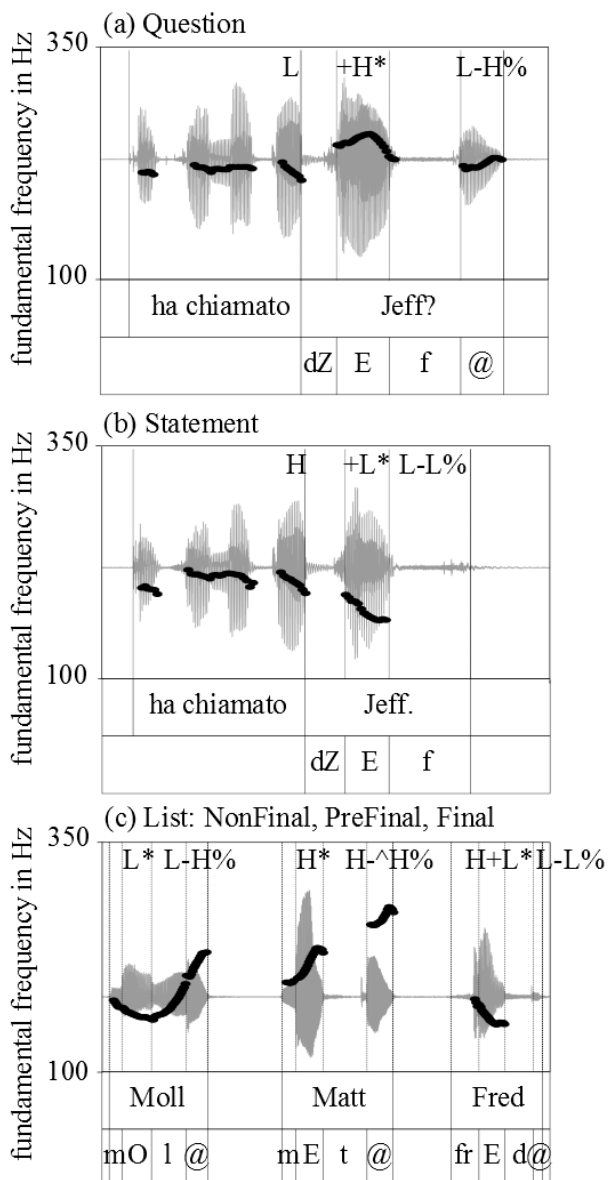
In general, there were many instances of schwa throughout both of the data sets, with schwa being present in 79% of all instances in the question-statement data set and 74% of all instances in the list data set.

3.2. Exploring the data

Fig. 2 shows the "variable importance" calculated on the basis of the random forest analysis, showing the extent to which each factor was important for correct classification. The factors are ranked from top to bottom by importance.

Figure 2: Variable importance measure generated by random forests for the question-statement data set (top) and the list data set (bottom).
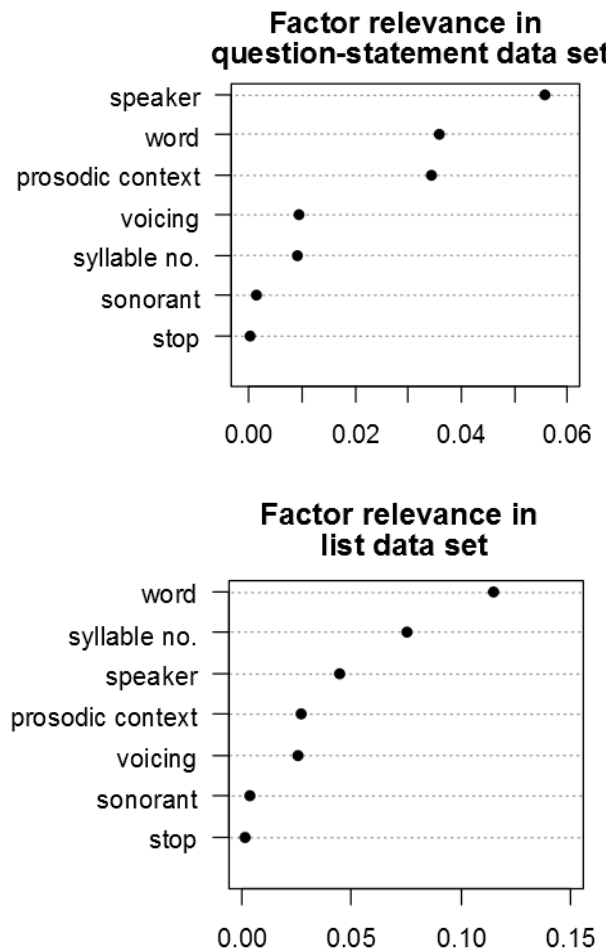
It becomes apparent that a number of different factors are important for predicting whether speakers produce a schwa or not. In both data sets, the idiosyncratic factors SPEAKER and WORD are highly ranked. This ranking reflects the high inter- and intra-speaker variability that appears to be a common characteristic of the production of loan words in the language [17].

Moreover, the identity of the word-final consonant (in the donor language) also has an impact on whether schwa is present. Schwa occurs more often following voiced consonants (85%) than voiceless ones (64%). Thus, despite the fact that schwa is treated as a clearly epenthetic element, it is still to some degree dependent on the laryngeal specifica-

tion of a preceding segment. The distinction between stops and fricatives, and between sonorants and obstruents appear to have no independent effect.

Crucially, the factor SYLLABLE NUMBER is rather highly ranked, especially in the list data set. Overall, monosyllabic words are much more likely to surface with a schwa (91%) than disyllabic (trochaic) words (52%), reflecting the greater attention paid in the literature to schwa in loan words that are monosyllabic.

Table 1: Proportion of observed schwa and its duration and intensity as a function of prosodic context in monosyllabic words.

	proportion (%)	duration (ms)	intensity (dB)
Question *Rise-Fall-Rise*	100.0	120.9	66.1
Statement *Low-Fall*	80.0	84.4	59.6
NonFinal *Low Rise*	99.5	107.1	66.8
PreFinal *High Rise*	97.0	102.8	67.6
Final *Low Fall*	77.5	86.4	60.0

Table 2: Proportion of observed schwa and its duration and intensity as a function of prosodic context in disyllabic trochaic words.

	proportion (%)	duration (ms)	intensity (dB)
Question *Rise-Fall-Rise*	71.2	91.1	64.6
Statement *Low-Fall*	53.3	75.8	58.9
NonFinal *Low Rise*	35.8	61.8	61.7
PreFinal *High Rise*	52.9	61.5	63.5
Final *Low Fall*	45.0	79.9	60.9

Looking at Tables 1 and 2, it is clear that, irrespective of the other effects, PROSODIC CONTEXT is relevant for determining the presence and acoustic prominence of schwa. In the monosyllabic loan words (Table 1), schwa occurs in all questions, but in only 80% of statements. In lists, schwa occurs in almost all non-final and pre-final items, but only in 78% of final items. Moreover, schwas are longer and louder if the tune to be realized is more complex, or if it is rising rather than falling.

In the disyllabic words (Table 2), the overall number of schwas was smaller. Here, schwa occurs in 71% of questions but 53% of statements, and is longer and louder in questions. In the list data set the picture is mixed, with schwa occurring most often in pre-final items but least often in non-final items.

To sum up, besides highly speaker- and word specific patterns, we find that schwa is dependent on both phonetic and prosodic factors. The presence of schwa is to some extent determined by the laryngeal context, with more schwas following voiced consonants. Importantly, the presence and duration of schwa is greatly affected by the number of syllables in the target word and by the tune to be realized on that target word.

In particular if the word is monosyllabic, the text is suboptimal for bearing a pitch movement. The insertion of a schwa in such cases enables the tune to be realized on more voiced material. This adjustment is further affected by the complexity and direction of the tonal movement to be realized. More complex tunes (rise-fall-rise or rise-fall) need more space to be realized than simple tunes (fall), thus schwa is more likely to be inserted in questions than in statements, and if it is inserted, it is longer and louder. Likewise, rising tunes take longer to execute than falling tunes [14, 23], thus schwa is more likely to be needed in the items in lists bearing rising tunes (non-final and pre-final) than those bearing falling ones (final position). The pressure to insert a schwa is less acute in disyllabic words, possibly accounting for the mixed picture in the list data set.

4. CONCLUSION

In this paper we show that in consonant final loan words in Italian the tune can condition adjustments to the text in terms of the insertion and acoustic prominence of schwa. However, we also show that this insertion is probabilistically distributed and dependent on other factors too, such as the metrical structure of the target word and the laryngeal specification of the preceding consonant.

The insertion of a vocalic element to facilitate the realization of functionally relevant tonal movements is in line with findings from genetically unrelated languages, such as Tashlhiyt Tamazight (Berber). In this language, schwa has been found to be highly dependent on prosodic contexts, with schwa being more likely to surface in positions in which tonal movements are located [10,18].

Thus, in Italian, as in Tashlhiyt, intonational tones may not be the only factors involved, but they clearly play a considerable role in determining the restructuring of the textual material with which they are associated. In this sense we can conclude that the tune drives the text.

5. REFERENCES

[1] Bannert, R., Bredvad, A. 1975. Temporal organization of Swedish tonal accent: The effect of vowel duration. *Working papers Phonetics Laboratory 10*, Department of General Linguistics, Lund University.

[2] Bertinetto, P. M. 1985. A proposito di alcuni recenti contributi alla prosodia dell'italiano. *Annali della Scuola Normale Superiore di Pisa*, 12, 581–643.

[3] Boersma, P., Weenink, D. 2014. *Praat: doing phonetics by computer* (Version 5.4.03).

[4] Breiman, Leo, 2001. Random forests. *Mach. Learn*, 45, 5–32.

[5] Broniś, O. 2014. [amˈburger] vs. [odˈdɔggə]: Word-final Vowel Epenthesis in Italian Loanword Adaptation. *Poster presentation at 22nd Manchester Phonology Meeting*.

[6] D'Achille P. 2010. *L'italiano contemporaneo*. Bologna: Il Mulino

[7] Erikson Y., Alstermark M. 1972. Fundamental Frequency correlates of the grave word accent in Swedish: the effect of vowel duration. *Speech Transmission Laboratory. Quarterly Progress and Status Report* 2–3, KTH, Sweden.

[8] Grabe, E., Post, B., Nolan, F., Farrar, K. 2000. Pitch accent realization in four varieties of British English, *Journal of Phonetics*, 28, 161–185.

[9] Grice M., D'Imperio, M., Savino, M., Avesani, C. 2005. Strategies for intonation labelling across varieties of Italian. In: Sun-Ah Jun (ed.), *Prosodic Typology: the Phonology of Intonation and Phrasing*, New York: OUP, 362–389.

[10] Grice, M., Roettger, T. B., Ridouane, R. (in press). Tonal association in Tashlhiyt Berber. Evidence from polar questions and contrastive statements. *Phonology*.

[11] Grice, M., Savino, M., Refice, M. 1997. The intonation of questions in Bari Italian: do speakers replicate their spontaneous speech when reading? *PHONUS*, 3, Institut fuer Phonetik/Phonologie, Univ. des Saarlandes, Saarbruecken, 1–7.

[12] Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., Van der Laan, M. 2006. Survival Ensembles. *Biostatistics*, 7(3), 355–373.

[13] Krämer, M. 2009. *The Phonology of Italian*, Oxford: OUP.

[14] Ohala, J. J., Ewan, W. G. 1973. Speed of pitch Change, *JASA*, 53, 345(A).

[15] R Core Team. 2014. *R: A language and Environment for Statistical Computing* (Version 3.1.0). The R foundation for statistical computing, Vienna, URL: http://www.R-project.org.

[16] Refice, M., Savino, M., Grice, M. 1997. A contribution to the estimation of naturalness in the intonation of Italian spontaneous speech. In: Proceedings of the 5th EUROSPEECH, vol. 2, 783–786.

[17] Repetti, L. 2012. Consonant-Final Loanwords and Epenthetic Vowels in Italian, *Catalan Journal of Linguistics*, 11, 167–188.

[18] Roettger, T. B., Ridouane, R., Grice, M. 2014. Schwa in Tashlhiyt Berber in voiceless environment. *Poster presentation at 14th Conference of Laboratory Phonology (LabPhon), Tokyo*.

[19] Savino, M. 2012. The Intonation of Polar Questions in Italian: Where is the rise? *JIPA*, 42(1), 23–48.

[20] Strobl, C., Boulesteix, A-L., Kneib, T., Augustin, T., Zeileis, A. 2008. Conditional Variable Importance for Random Forests. *BMC Bioinformatics*, 9(307). URL http://www.biomedcentral.com/1471-2105/9/307.

[21] Strobl, C., Boulesteix, A-L., Zeileis, A., Hothorn, T. 2007. Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinformatics*, 8(25). URL http://www.biomedcentral.com/1471-2105/8/25.

[22] Tagliamonte, S. A., Baayen, R. H. 2012. Models, forests, and trees of York English: was/were variation as a case study for statistical practice, *Language Variation Change*, 24(2), 135–178.

[23] Xu, Y., Sun, X. 2002. Maximum speed of *pitch* change and how it may relate to speech. *JASA*, 111, 1399–1413.

[24] Savino, M. 2001. Non-finality and pre-finality in Bari Italian intonation: a preliminary account, *Proceedings of Eurospeech 2001*, Aalborg, 939-942.

[25] Savino, M. 2004. Intonational cues to discourse stucture in a variety of Italian. In Gilles, P. & Peters, J. (eds), *Regional variation in intonation*. Tuebingen: Niemeyer. 161-187.