

Vowel duration, voicing duration, and vowel height: Acoustic and articulatory data from Italian

Stefano Coretta

29/11/2018

1 Methods

1.1 Participants

Participants for this study were recruited in the province of Verbano-Cusio-Ossola (VCO), Piedmont, Northern Italy. Inclusion in the study was dependent on the participant being a native speaker of Italian, of adult age (18 yo or older), with no reported hearing or speaking disorders, and normal or corrected to normal vision. The target sample size was set up to be 30, to be reached within a maximum of three months from the start of data collection. A total of 19 participants were recorded within the established timeframe. Fifteen were female and the mean age was 40 (range 18–59). Seventeen participants lived all or most of their lives in the VCO province. IT03 is from a neighbouring province (Varese), and the variety of Italian he speaks is indistinguishable from that of the VCO speakers. IT07 is from Sardinia, but she spent the 10 years ahead of recording in the VCO province. Her variety of Italian is different from the one of the other speakers, but excluding her data from the analyses produced negligible or no differences in the results, so they have been kept. Ethical clearance for this study was obtained from the University of Manchester (REF 2016-0099-76). The speakers were not paid for their participation.

1.2 Equipment

Synchronised audio and electroglottographic signals were obtained with a Glottal Enterprises EG2-PCX2 electroglottograph and a RØDE Lavalier microphone, at a sample rate of 44100 Hz (16-bit; downsampled to 22050 Hz for analysis). The acquisition of the signals was controlled with Audacity (2.2.2) running on a MacBook Pro (13-inch, Mid 2014, ...) with macOS The EGG electrodes were placed around the participant's neck with a velcro strap, at the height of the glottis, one on each side of the thyroid cartilage. Accuracy of the electrode vertical placement was assessed by checking the EGG unit placement monitor light (green indicates correct placement). The low frequency (LF) limit selector was set on 20 Hz. The microphone was clipped on the participants' clothes, about 20 cm from their mouth.

1.3 Materials

To test the influence of vowel height on vowel and voicing duration, a set of target words with a stressed vowel flanked by two voiceless stops. The words have a CVCV structure, with lexical stress on the first syllable, and where $C_1 = /p, t, k/$, $V_1 = /a, e, \text{ɔ}, i, u/$, $C_2 = /p, t, k/$, and

$V_2 = /o/$. All possible combinations were used, except for *peto* and *caco* (which are bad words). The full list of target words is given in ... The traditional description of the Italian vocalic system includes both mid-low $/\epsilon, \text{ɔ}/$ and a mid-high vowels $/o, \text{ɔ}/$. While these vowels are usually ascribed to different phonemes (Krämer 2009), their functional load and actual phonetic realisation are less straightforward (Renwick & Ladd 2016). Although some minimal pairs exists both for the front and the back contrast, stressed open syllables in Italian tend to have the mid-high front vowel $/e/$ and the mid-low back vowel $/ɔ/$. Note also that penultimate stressed open syllables in Italian have a phonetically long vowel, for example $/poko/ = [\text{p}ɔ\text{:ko}]$ (Renwick & Ladd 2016).

These words were embedded in 4 frame sentences, adapted from Renwick & Ladd (2016): *Scrivete X sul foglio* ‘Write (you pl.) X on the sheet’, *Ha detto X sei volte* ‘S/he said X six times’, *Ripete X da sempre* ‘S/he’s being repeating X for ever’, *Sentivo X di nuovo* ‘I heard X again’. In all the sentences, the target word is preceded and followed by three syllables, where the second syllable is stressed. Using different frames reduces possible repetition effects, while allowing the acquisition of 4 word tokens per speaker. All word/frame combinations were used, generating a total of $43 \text{ words} \times 4 \text{ frames} = 172$ observations per speaker.

1.4 Procedure

Participants familiarised themselves with the materials prior to recording and practiced reading them aloud. The stimuli sentences were presented in a randomised order on the laptop screen used for recording, through the software PsychoPy v1.90.3 Peirce (2009). The participants were given a change to take a break half the way through the task, which lasted 8–10 minutes. At the end of the speaking task, the participants filled a short sociolinguistic questionnaire. The whole experiment session lasted 30 minutes.

1.5 Data processing and measurements

The audio recordings were downsampled to 22050 Hz (16-bit) for acoustical analysis. Noise removal with the Praat built-in function Remove noise (Boersma & Weenink 2018) was applied to the audio recordings of IT01, IT02, IT03, IT04, IT07, IT10, and IT12. An automatic time-aligned transcription was performed with the SPeech Phonetisation Alignment and Syllabification software (SPPAS, Bigi 2015), and subsequently manually fixed according to the criteria in Machač & Skarnitzl (2009). In particular, vowel onset and offset were placed in correspondence of the appearance and disappearance of higher formant structure in the spectrogram. Only the following relevant boundaries were corrected: sentence onset and offset, target word onset and offset, target word segment boundaries. The stop releases of C1 and C2 were automatically detected with a Praat implementation of the algorithm described in Ananthapadmanabha et al. (2014) and then manually corrected if necessary.

The EGG signal was pass-band filtered (40 Hz–10 KHz) to remove hardware noise, and smoothed using a weighted moving average with smooth width = 11 (Macerata et al. 2016). The Praat built-in function To TextGrid (vuv) was used to detect voiced/voiceless (Voiced/UnVoiced) intervals. The onset and offset of the voicing of V1 was then corrected manually. The boundaries were placed at the signal minimum before/after the first/last visible glottal pulse.

The following interval durations were measured with a custom script:

- **V1 duration:** from the onset to the offset of higher formant structure of V1.
- **Duration of voiced interval of V1** (V1 voicing duration): from the voice onset to the voice offset of the voiced interval of V1.
- **Voice Onset Time:** from the release of C1 to V2 voice onset.
- **Release to Vowel Offset:** from release of C1 to V1 offset.
- **Speech rate:** calculated as syllables per second (number of syllables / sentence duration, Plug & Smith 2018).

1.6 Statistical analysis

1.7 Open Science statement

References

- Ananthapadmanabha, T. V., A. P. Prathosh & A. G. Ramakrishnan. 2014. Detection of the closure-burst transitions of stops and affricates in continuous speech using the plosion index. *The Journal of the Acoustical Society of America* 135(1). 460–471. doi:10.1121/1.4836055.
- Bigi, Brigitte. 2015. SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician* 111–112. 54–69.
- Boersma, Paul & David Weenink. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.40. <http://www.praat.org/>.
- Krämer, Martin. 2009. *The phonology of Italian*. Oxford: Oxford University Press.
- Macerata, Alberto, Andrea Nacci, Mariangela Manti, Matteo Cianchetti, Jacopo Matteucci, Salvatore Osvaldo Romeo, Bruno Fattori, Stefano Berrettini, Cecilia Laschi & Francesco Ursino. 2016. Evaluation of the electroglottographic signal variability by amplitude-speed combined analysis. *Biomedical Signal Processing and Control* 37. 61–68.
- Machač, Pavel & Radek Skarnitzl. 2009. *Principles of phonetic segmentation*. Epocha.
- Peirce, Jonathan W. 2009. Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics* 2(10).
- Plug, Leendert & Rachel Smith. 2018. Segments, syllables and speech tempo perception. In *Proceedings of the 9th international conference on speech prosody 2018*, 279–283. doi:10.21437/SpeechProsody.2018-57.
- Renwick, Margaret & Robert D. Ladd. 2016. Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7(1). 1–29. doi:10.5334/labphon.17.