Mid-year panel report 2016-2017

Stefano Coretta

20/01/2017

Contents

1 Summary of achievements (Year 1 Term 1)

- *Definition of a broad research question*
  - This research will focus on the relative timing between lingual gestures and the configuration of the glottis during speech. Articulatory data will be acquired through ultrasound tongue imaging (UTI) and electroglottography (EGG), which allow, respectively, tracking of tongue movements and calculation of the amount of contact between the vocal folds. These two techniques together will enable a time-aligned account of oral and laryngeal gestures.
- *Training*
  - I received training on the operation of the ultrasound and EGG machines, and on the Articulate Assistant Advanced software (used for the analysis of ultrasound data)
- *Definition of a range of hypotheses to be tested*
  - I have established a series of research hypothesis that are undergoing assessment through an exploratory pilot experiment. Previous studies have shown that vowels followed by voiced consonants are usually longer than vowels followed by voiceless consonants (House & Fairbanks, 1953; Chen, 1970; Hussein, 1994; Durvasula & Luo, 2012). Several explanations for the voicing effect have been proposed in the literature, but none has been conclusive. A promising line of enquiry emerges from the proposal by Chomsky & Halle (1968), who attribute the voicing effect to adjustments of the larynx when moving from a vowel to a consonant. This hypothesis, at that time just the product of a theory, can now be tested, thanks to recent advances in articulatory methods.
- *Application of Ethics Approval for experimental research*
  - I have applied for ethics approval and obtained clearance from the School Ethics Committee.
- *Experimental design and pilot*
  - I ran a pilot study to test the methodology and explore possible variations in relative timing between oral and glottal gestures in different phonological conditions.
  - A speaker of Polish and two speakers of Italian participated in the pilot. The subjects had to read sentences from a computer screen while being recorder. The sentences contained nonce words as target stimuli. The nonce words used in the experiment are shown in Table 1. The frame sentence was *Mówię* ___ *teraz* “Say ___ now” for polish, and *Dico* ___ *lentamente* “I say ___ slowly” for Italian. Audio, ultrasound (US) and electroglottographic (EGG) recordings were

taken simultaneously. Synchronisation of the US and EGG recordings was achieved through cross-correlating the audio signals from those two sources.
  - Due to technical issues, I could not collect EGG data for the Polish participant. I tested the analysis method on the data from one Italian speaker. The preliminary results show that, for the Italian speaker, the configuration of the glottis in vowels is affected by the phonation of the following consonants. This phenomenon is consistent with the idea that the differences in vowels in different consonantal contexts depend on the articulatory implementation of the vowel-to-consonant transition.
- *Fund application*
  - I have applied for a Small Grant by the Endangered Languages Documentation Project (ELDP). If successful, the grant will cover the costs of fieldwork activity in Italy, where I will collect data on Western Lombard. Western Lombard has been reported to show variation in vowel duration depending on the voicing of the following consonant. Interestingly, this language is of great importance to the understanding of this phenomenon, known as the voicing effect, due to an unusual phonological pattern of vowel duration, reported by Prieto i Vives (2000). So far, research made on this matter has been based on older impressionistic descriptions of the phonology of the language (Sanga, 1988; Prieto i Vives, 2000; Torres-Tamarit, 2015). Fieldwork research will help assess the theoretical claims formulated in the literature and test them empirically.

Table 1: Stimuli list

pata	bata	poto	boto	putu	butu
pada	bada	podo	bodo	pudu	budu
paka	baka	poko	boko	puku	buku
paga	baga	pogo	bogo	pugu	bugu

2 Technical report

2.1 Synchronisation of ultrasound and EGG signals

Since the signals from the ultrasound machine and the laryngograph are recorded simultaneously but into two separate laptop, data from both machines need to be synchronised after acquisition. Synchronisation is achieved through the cross-correlation of the audio signals from both sources (Grimaldi et al., 2008). The cross-correlation method creates a new sound file from two audio files. The created new file is a convolution of the original files. The time of the maximum amplitude in the convoluted sound wave is the amount of off-set between the two original files. The off-set is trimmed from the beginning of the longer audio file, with the result that the files will be in sync. A measure taken at any particular time in the ultrasound source can thus be related to a measure taken at that same time in the laryngograph source.

2.2 Calculation of dEGG and tracegram analysis

Herbst et al. (2010) describe a new technique, called electroglottographic wavegram, which displays the variations in the EGG and dEGG signals in a single graph. A wavegram contains temporal information on the x and y axis, while changes in the VFCA are rendered as different colour intensities on the z axis.

The extraction of dEGG maxima and minima has been implemented in this study using the PRAAT scripting language. The algorithm consists of the following stages:

1. detection of the glottal periods
2. calculation of the dEGG
3. extraction of absolute dEGG maximum ($dEGG_{max}$) and minimum ($dEGG_{min}$) for each glottal period
4. calculation of $dEGG_{max}$ and $dEGG_{min}$ relative to the glottal period

It is conventional to define a glottal period as the time between two consecutive moments of glottal closure, i.e. two consecutive dEGG maxima. However, since the maxima need to be identified in the first place, an arbitrary definition of glottal period is instead used. Glottal periods correspond to the intervals between two consecutive EGG minima [cf. ...]. First, the EGG signal is band-pass filtered (40Hz-10KHz) and smoothing is applied. A weighted sliding-average smoothing method (triangular smooth) is used, with smooth width $m = 11$. EGG minima are thus extracted from the smoothed EGG signal. The interval between any two consecutive minima constitutes a glottal period.

The dEGG is calculated with the formula $x'_n = x_{n+1} - x_n$, where x_n is the value of the EGG signal at the time n. After calculation, the resulting dEGG is smoothed with the same method as before (triangular smooth, $m = 11$). The algorithm then searches for dEGG maxima and minima within each glottal period (defined as two consecutive EGG minima). Finally, relative $dEGG_{max}$ and $dEGG_{min}$ are calculated as proportions of the respective glottal period. The resulting values are between 0 (beginning of period) and 1 (end of period).

As Herbst et al. (2010) note, the wavegram technique has the limitation of not being suitable for quantitative analysis. A new visualisation technique, based on wavegrams, is introduced here: electroglottographic tracegram. The tracegram method, even if it reduces the displayed dimensions, allows a statistical assessment of the varying $dEGG_{max}$ and $dEGG_{min}$, thus constituting a partial improvement over wavegrams. After the calculation of the relative $dEGG_{max}$ and $dEGG_{min}$, these values are plotted in a graph on the y axis at each time point which corresponds to the beginning of a glottal period. Since the values are restricted between 0 and 1 (being proportions), changes in glottal period (which corresponds to changes in fundamental frequency and hence pitch) are controlled for. The resulting graph, the tracegram, shows the traces of $dEGG_{max}$ and $dEGG_{min}$ as they change in time, in a way similar to the display of pitch contours. Statistical analysis is then performed on both traces separately using a Smoothing Spline ANOVA (SSANOVA).

References

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22(3). 129–159.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York, Evanston, and London: Harper & Row.

Durvasula, Karthik & Qian Luo. 2012. Voicing, aspiration, and vowel duration in Hindi. *Proceedings of Meetings on Acoustics* 18. 1–10.

Grimaldi, Mirko, B. Gili Fivela, Francesco Sigona, Michele Tavella, Paul Fitzpatrick, Laila Craighero, Luciano Fadiga, Giulio Sandini & Giorgio Metta. 2008. New technologies for simultaneous acquisition of speech articulatory data: 3D articulograph, ultrasound and electroglottograph. *Proceedings of LangTech* 1–5.

Herbst, Christian T., W Tecumseh S. Fitch & Jan G. Švec. 2010. Electroglottographic wavegrams: A technique for visualizing vocal fold dynamics noninvasively. *The Journal of the Acoustical Society of America* 128(5). 3070–3078.

House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25(1). 105–113.

Hussein, Lutfi. 1994. *Voicing-dependent vowel duration in Standard Arabic and its acquisition by adult american students*: The Ohio State University dissertation.

Sanga, Glauco. 1988. La lunghezza vocalica nel milanese e la coscienza fonologica dei parlanti. *Romance Philology* 41(3). 290–297.

Torres-Tamarit, Francesc. 2015. Length and voicing in Friulian and Milanese. Or why rule-free derivations are needed. *Natural Language & Linguistic Theory* 33(4). 1351–1386.

Prieto i Vives, Pilar. 2000. Vowel lengthening in Milanese. In Lori Repetti (ed.), *Phonological theory and the dialects of Italy*, 255–272. Amsterdam Philadelphia: John Benjamins Publishing Company.