# A quantitative analysis of intra- and inter- speaker variation of F0, F1, F2 and vowel duration from *soft* to *loud* speech in Italian.

Massimiliano Canzi
massimiliano.canzi@postgrad.manchester.ac.uk
/mcanzi/baap

## 1. Background

Acoustic features of vowel phonemes are valuable information for both human speaker identification and forensic voice analysis (Remez et al.1997). Possible effects of speech mode change (i.e. from neutral to loud speech) on vowel acoustics have been studied, but a comprehensive analysis of a language's whole monophthong inventory is missing. Most studies in forensic acoustics only deal with few phonemes (e.g. Elliot, 2000). Furthermore, in this study, **soft and loud speech modes are elicited in speakers in no-noise conditions.**

**Research questions:** How does speech mode (soft, neutral, loud) affect speaker identification?

**Hypotheses:** Change is found in the acoustic parameters analysed among the three speech modes of the experiment. General patterns are found throughout the whole vowel monophthong inventory for measures such as F0 (and its SD), F1, F2 and vowel duration

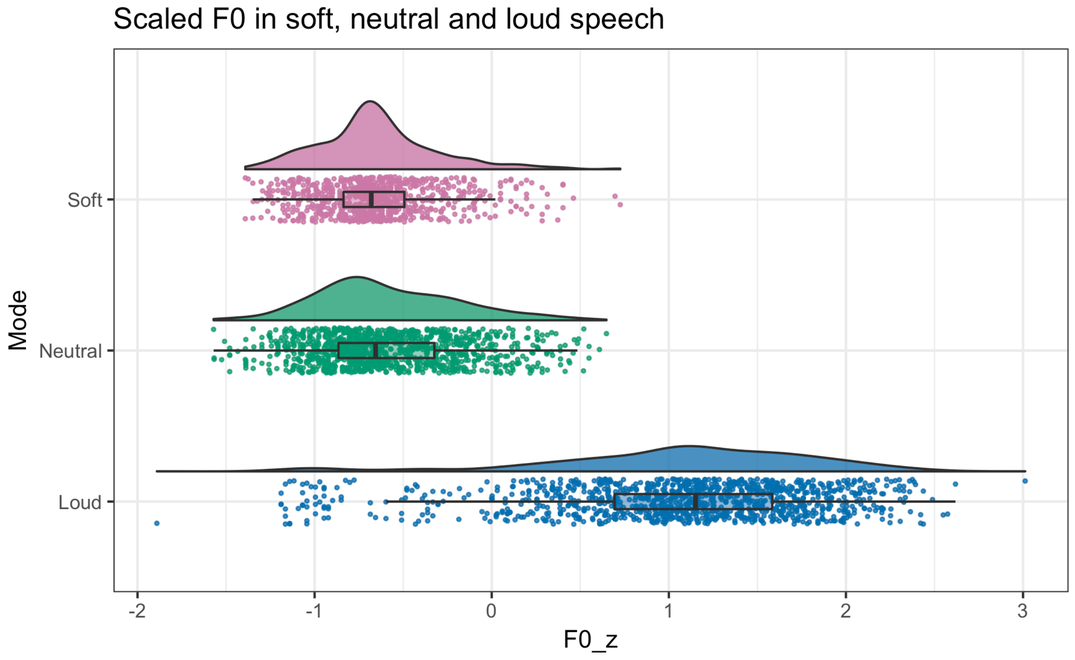
## 2. Methods

8 speakers with L1 Italian (5 males) read a word list of 168 items and a sentence list in each of the three speech modes. Linear mixed regression models - lmer test (Kuznetsova et al 2015) - were used for the analysis of the data.

The three speech 'modes' are described as volitional increase or decrease of *perceived speech loudness*. Speakers were only instructed to read the word list once in each of the three speech modes. No measurements were obtained for acoustic parameters of average SPL in dB.

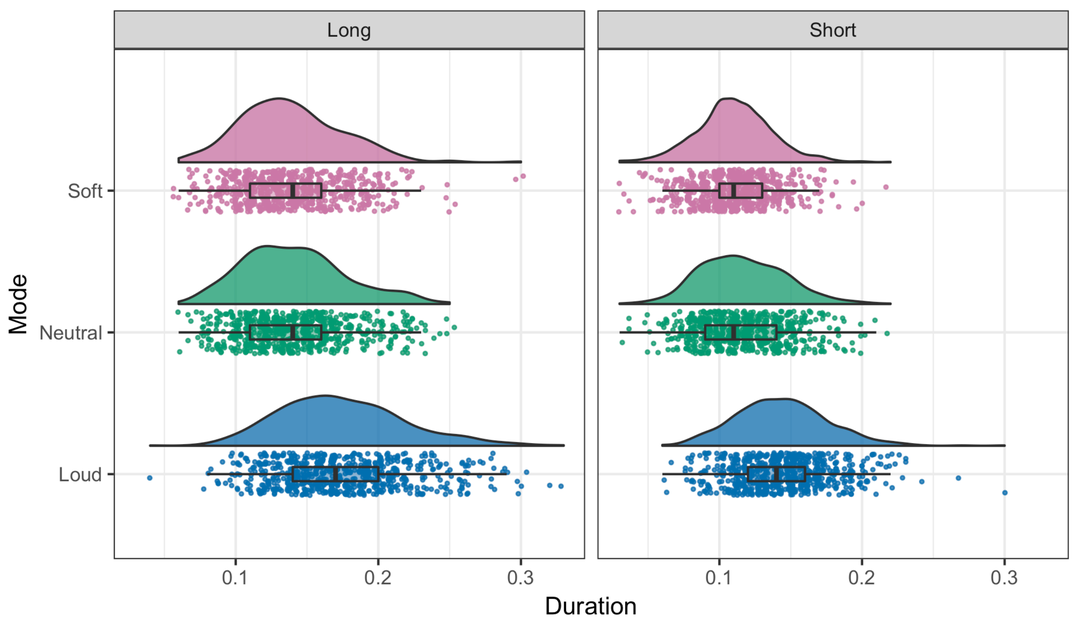
## 3. Results

Scaled F0 in soft, neutral and loud speech



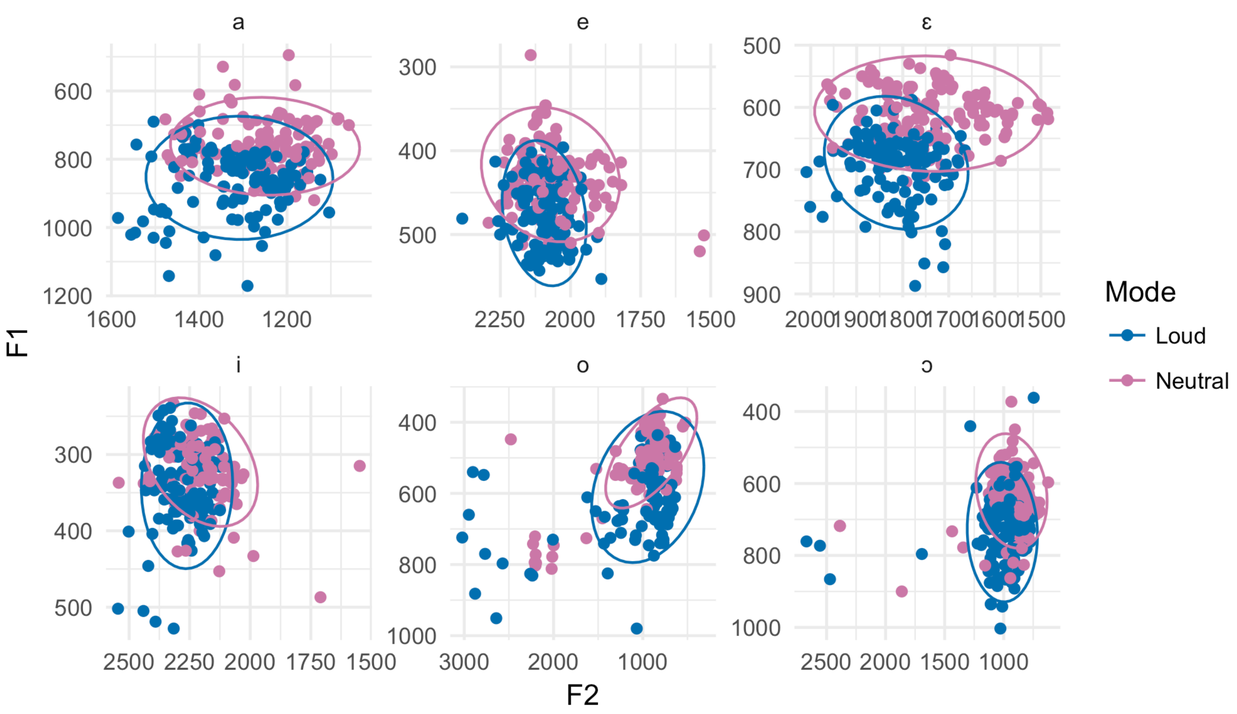

F0 **increases significantly between neutral and loud speech** by 20 Hz (**t(7) = 7.53, p < 0.001**) and so does its standard deviation from soft to loud speech.

Vowel Duration in soft, neutral and loud speech in Seconds




Vowel duration was also found to increase between neutral and loud speech (**t(7) = 8.41, p < 0.0001**). Phonological vowel length, which is labelled as **long** and **short** in the plot above, does not interact with speech mode.

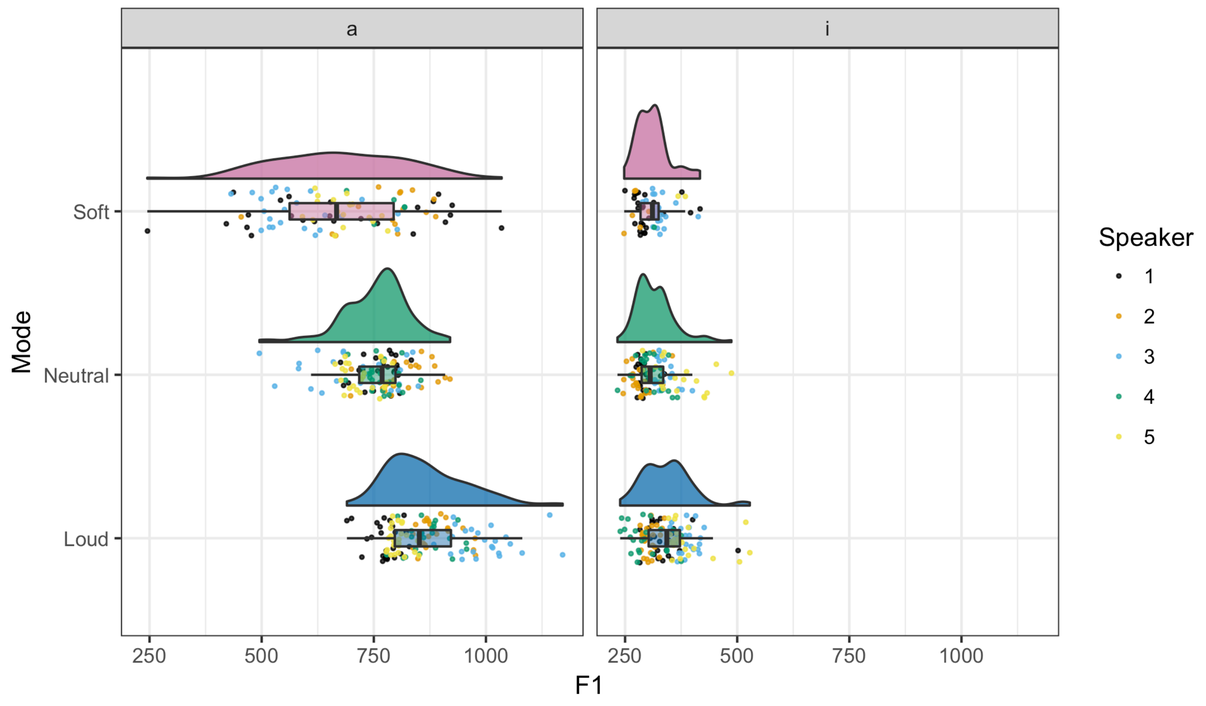
F1-F2 vowel plots for male speakers in neutral and and loud speech




F1 increases of **25 Hz** from soft to neutral (**t(7) = 3.45, p < 0.004**) and of **41 Hz** from neutral to loud speech (**t(7) = 7.77, p < 0.001**).
However, the effect size is different depending on the vowel and speaker.

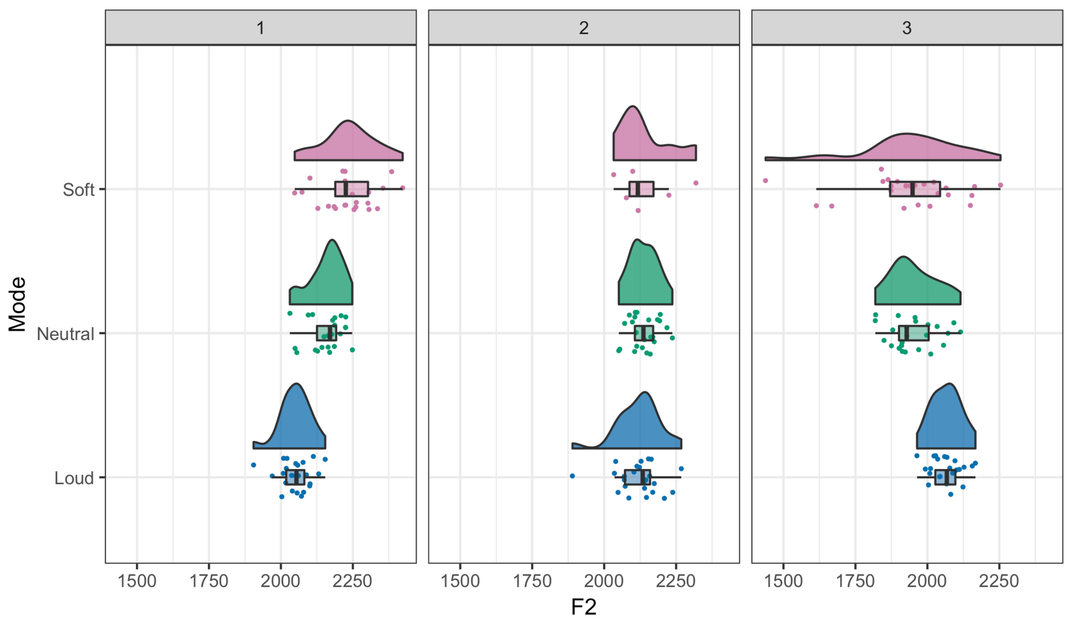
The plot below shows the differences between two vowels: /a/ and /i/. While the values for /i/ are less spread out for all speakers, the increase is smaller than that of /a/. where the values for all speakers are less concentrated.

F1 for /a/ and /i/ by male speakers in soft, neutral and loud speech




While F2 changes significantly between neutral and loud speech (**t(7) = 3.46, p < 0.01**) the direction of the change depends on single vowels and speakers, as shown in the example below.

F2 for /e/ by male speakers 1, 2 and 3 in soft, neutral and loud speech




## 4. Conclusion

Overall, changes are more pronounced from neutral to loud speech. Variance is generally higher in loud rather than in soft and neutral speech. Increased variance and unpredictable change is not favourable for speaker identification, especially when effects - although present - and their directions are hard to quantify and predict.

## 5. References

Elliot, J. et al. (2000). Comparing the acoustic properties of normal and shouted speech: a study in forensic phonetics. In *Proc. SST-2000: 8th Int. Conf. Speech Sci. & Tech*, pages 154–159.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2015). Package: lmer test. *R package version*, 2.
Remez, R. E., Fellowes, J. M., and Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance*, 23(3):651.