Using Chess Metrics to Measure the Effect of Emotion on Formants

The phonetic correlates of emotion have been studied since at least Darwin's *Expression of the Emotion in Man and Animals* (1872). However, experimental studies on the subject have suffered from a lack of emotionally-colored data (Gobl & Chasaide, 2003; Murray & Arnott, 1993). On the one hand, eliciting emotions in experimental subjects through subterfuge is difficult and likely not ethical. On the other hand, using recordings from professional actors assumes that all emotional cues are perceptible and mimicable, which is not necessarily the case (Arndt & Janney, 1991). When it comes to phonetic variables, studies about the effect of emotion on formants are scarce (but Goudbeek et al. 2009; Waraama et al., 2018), largely due to the fact that prosodic cues appear to be far more perceptually salient (Carlson et al., 1992; Starkweather, 1960).

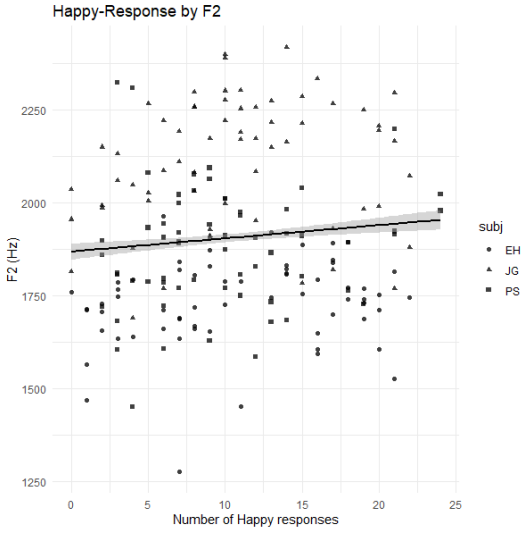
This study focuses chiefly on the effects of emotion on formants and attempts to circumvent issues of obtaining emotionally-colored data by focusing on naturalistic productions (~14 hours) of chess grandmasters (N = 7) playing chess against users of the *chess24.com* website while simultaneously commentating their games. All recordings were in English; speakers varied greatly in age and language background, with 1 native speaker of English, 1 French, 1 German, 1 Russian, 1 Croatian, and 2 Dutch. F_1 and F_2 measurements for the vowel [i] were extracted from a limited set of common chess vocabulary. As a proxy for emotional state, several chess metrics were extracted at each vowel production: **evaluation** (how well the subject is doing in the game), **game time** (amount of thinking time remaining for the speaker), **session time** (amount of time the subject has been playing overall) and **complexity** (the number of pieces remaining).

Results from 2705 vowel tokens reveal that subjects' F_2 was correlated positively with evaluation (see Figure 1). Under the assumption that chess evaluation had a direct effect on emotional valence (i.e. speakers were less happy when they were losing), these findings are in line with existing literature (Goudbeek at al., 2009; Waraama et al., 2018; Ohala, 1984), which associate high F_2 measurements with happiness, a likely contributor to which is lip retraction during smiling (Tartter, 1980).

A follow-up perception study using a subset of the collected data was also undertaken. Listeners (N = 24) were tasked with evaluating the perceived emotion of a given utterance on two dimensions: **valence** (happy ~ unhappy) and **agitation** (tense ~ relaxed). All utterances were lexically uninformative chess terms (*B2*, *C4*, *G6*, etc.). During the analysis, the F_1 and F_2 of the first of the two vowels, which was always an [i], were compared against the perceived emotions.

The perception results confirm that higher F_2 was perceived as an indication that the speaker was more happy (see Figure 2). It was also found that higher F_1 was perceived as an indication that the speaker was more tense (see Figure 3). The latter effect, although hitherto unobserved in perception studies, has been found in a previous production study (Goudbeek et al., 2009).

This project contributes to the currently understudied topic of segmental cues of emotion by confirming previous results with naturalistic data. In addition, the speakers' varied language backgrounds allow inferences about the state of these effects cross-linguistically and/or in L2 speech.



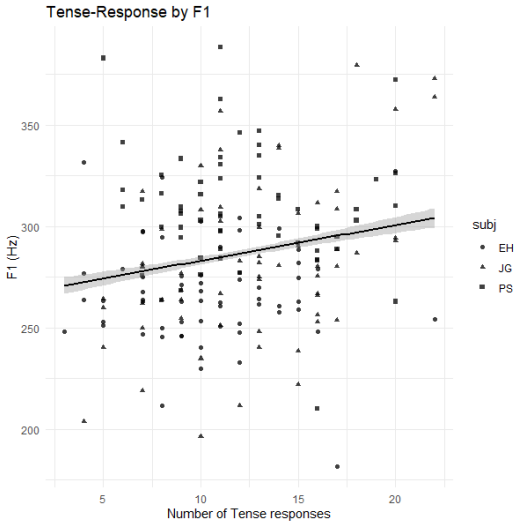


Figure 1 (top left): Violin plot of speakers' F_2 (normalized by speaker and word) by chess evaluation (binned into psychologically intuitive categories: *loss*, *draw*, *win*).

Figure 2 (top right): Speakers' F_2 by items' happiness score (sum of times listeners rated item as 'happy'). p = .023

Figure 3 (right): Speakers' F_1 by items' tenseness score (sum of times listeners rated item as 'tense'). p = .022

References:

Arndt, H., & Janney, R. (1991). Verbal, prosodic, and kinesic emotive contrasts in speech.

Carlson, R., Granstrom B. & Lennart N. (1992). Experiments with emotive speech – acted utterances and synthesized replicas.

Goudbeek M., Goldman J., & Scherer K. (2009). Emotion dimensions and formant positions.

Murray, I., & Arnott, J. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion.

Ohala, J. (1984). An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice.

Starkweather, J. A. (1956). Content-free speech as a source of information about the speaker.

Tartter, V. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech.

Waaramaa, T., Kukkonen, T., Mykkänen, S., & Geneid, A. (2018). Vocal Emotion Identification by Children Using Cochlear Implants, Relations to Voice Quality, and Musical Interests.