

Intuitive Data Representation Techniques for Representing Paralinguistic Speech Data

- Akash Chaudhary

Motivation

- Non-native speakers have problem speaking English in a stress timed pattern.
- They tend to speak it in the way their native language is spoken.
- Hence, they lack speech intentionality.
- Native speakers find their speech unintelligible.
- This reduces scope for job opportunities and effective communication capabilities, thereby limiting the the pace of individual growth.
- Hence, there exists a need to design a text visualization that is intuitive for non-native speakers to understand speech intentions.

Proposed design features of text visualization

- Represents paralinguistic raw audio information (pitch, intensity, silences, syllable extension) over and above the word level information.
- Represents all of the above parameters together.
- Represents context.
- Does not decrease the legibility of reading.
- Is intuitive to read and perceive.

Context

- 4 types of context
- “(1) **Phonetic/linguistic context**, that is, what a speaker produced before or after the unit we want to analyse
- (2) **Multimodal context**, that is, which body posture, gestures, and facial gestures the speaker produces concomitantly, synchronously or before and after the unit we want to analyse
- (3) **Immediate situational context**, that is, the overall setting (communication partners, type of communication, room characteristics, etc.)
- (4) **General context in time and space**, that is, generally speaking, in which historic/geographic situation the communication partners are.”

Ref - Schuller, Björn, and Anton Batliner. *Computational paralinguistics: emotion, affect and personality in speech and language processing*.

John Wiley & Sons, 2013.

Context (Contd.)

- Context can be narrow, concentrating on the speaker herself and her personal situation and what she has experienced in recent hours, it can be wide, including macro-sociological and political constellations, and it can simply be narrowed down to membership of class, etc.
- Modeled the first context in visualizations.

Discussion

- Context can be changed to new and interesting cases.
- Relative change can be changed to absolute quantitative change.
- Normal written language follows Grice's 4 cooperative maxims to communicate -
- ***The maxim of quantity***, where one tries to be as informative as one possibly can, and gives as much information as is needed, and no more.
- ***The maxim of quality***, where one tries to be truthful, and does not give information that is false or that is not supported by evidence.
- ***The maxim of relation***, where one tries to be relevant, and says things that are pertinent to the discussion.
- ***The maxim of manner***, when one tries to be as clear, as brief, and as orderly as one can in what one says, and where one avoids obscurity and ambiguity.
- If we change the maxim of manner, we get figures of speech, like sarcasm and irony. It will be to visualize these figures of speech.