The Acoustic Contrasts of Emotional Expressions in New Zealand English

Dona Jayakody, MSc, Emily Lin, PhD., and Valerie Looi, PhD

Department of Communication Disorders, University of Canterbury, Christchurch, New Zealand

Abstract

Listeners can identify emotions based on vocal cues (Banse & Scherer, 1996). This study aims at identifying the acoustical parameters that aid in recognizing different emotions, for a better understanding of the difficulties faced by cochlear implant (CI) and hearing aid (HA) users in real life situations.

Introduction

Speech conveys both linguistic and non-linguistic information. While linguistic aspects of speech convey information about grammar, meaning, and sounds of an utterance, non-linguistic aspects provide information regarding talker identity, gender, age, dialect or accent, and emotional state of the speaker (Peters, 2006). Perceiving emotions is vital for effective communication and also for social interactions (Pereira, 1996; Scherer, 1984). One's ability to perceive vocal expressions plays an important role in interpreting emotions expressed in speech (Banse & Scherer, 1996; Most & Aviner, 2009). As the hearing impaired have difficulties in perceiving the content of speech, emotions expressed by the tone of voice can provide supplementary information in social situations (Most & Aviner, 2009; Pereira, 1996). As part of a treatment efficacy study, an emotion discrimination test was developed to compare the emotion identification abilities of the CI and HA users and normal listeners. This presentation reports the acoustical results of the sentences produced by two actors and two actresses simulating four emotional states.

Methods

Participants and Participant's Task Two male and two female actors, aged between 18-24 years (mean = 21 years, SD = 3), participated in the study. Each participant was asked, in each trial, to say one semantically neutral everyday English sentence to express one of the four target emotional states (Angry, Happy, Sad, and Neutral).

Instrumentation and Procedures Participants were seated 30 cm in front of a Sony ECM-MS907 microphone mounted on a stand at the level of the talker's mouth. Speakers were instructed to imagine a real life situation and to say the sentences with a specified emotion. Three trials were recorded for each target emotion. Recordings were made in a sound treated room and each speech sample was recorded as a PCM recording using a Sony HiMD N50 audio recorder. The sampling rate was 44,100 Hz, with a 16-bit resolution.

Acoustic Analysis The TF32 time-frequency analysis software was used for acoustic measures. The experimental measures include:

I. Whole sentence:
   1. Temporal measure (related to speech rate): number of syllable per second,
   2. Intensity measure: mean RMS value
   3. Pitch measure: mean fundamental frequency (F0), F0 range, standard deviations of F0

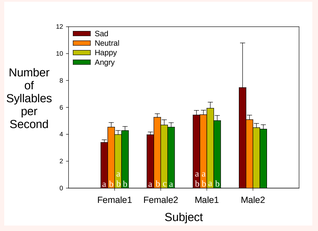
II. Vowel segment (i.e., a 50-ms mid portions of the selected vowel embedded in the sentences)
   1. Perturbation measures (related to voice quality): percent jitter (%Jit), percent shimmer (%Shim), and signal-to-noise ratio (SNR)
   2. LPC (linear predictive coding) spectral measures (related to articulatory precision and vocal tract constriction): frequencies of Formants one and two (F1 &F2)
   3. LTA (long-time average) spectral measures:
      -Related to voice projection power: the amplitude difference between the highest spectral peak in the frequency range between 0 to 2 kHz and that in the 2 to 4 kHz frequency range (i.e., singing power ratio or SPR).

Speech Material Eight semantically neutral utterances were selected from a list of 22 sentences by a panel of linguists. These sentences were selected as they did not require a context to maintain their meanings. The sentences are:

1. She is coming home.
2. I just sold my car.
3. I'm going home.
4. We're watching rugby.
5. They got back together.
6. They are getting married.
7. It's snowing outside.
8. It's raining outside.

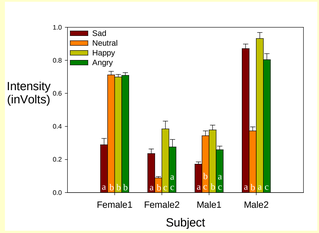
*Vowel segments used to derive perturbation and SPR measures are underlined.

Results



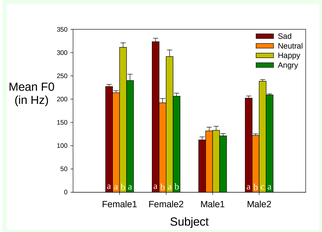
Speech Rate:

- Literature:
  - Happy & Angry: increased (Whiteside, 1999a)
- Findings in this study:
  - Happy > Sad: in Female2
  - Happy & Angry: decreased in Female2
  - Sad: decreased in Females 1 & 2
- One-way repeated measures (RM) analysis of variance (ANOVA) results:
  - Female1: F(3, 21) = 5.587, p = 0.006*
  - Female2: F(3, 21) = 15.641, p < 0.001*
  - Male1: F(3, 21) = 4.211, p = 0.018*
  - Male1: F(3, 21) = 0.711, p = 0.556



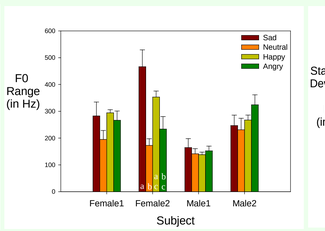
Loudness:

- Literature:
  - Happy & Angry: increased (Whiteside, 1999a, b)
  - Sad: decreased (Whiteside, 1999b)
- Findings in this study:
  - Happy & Angry: increased in Female2 & Male2
  - Sad: decreased in Female1 & Male1
  - Sad: increased in Female 2 & Male2
- One-way RM ANOVA results:
  - Female1: F(3, 21) = 157.532, p < 0.001*
  - Female2: F(3, 21) = 13.484, p < 0.001*
  - Male1: F(3, 21) = 16.175, p < 0.001*
  - Male1: F(3, 21) = 75.674, p < 0.001*



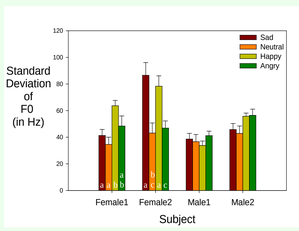
Mean F0:

- Literature:
  - Angry > Neutral > Happy > Sad (Whiteside, 1992 a, b)
- Findings in this study:
  - Happy > Sad: in Female1 & Male2
  - Angry > Neutral: in Male2
  - Happy > Angry & Neutral: in Females 1 & 2, & Male2
  - Sad > Neutral: in Female2 & Male2
- One-way RM ANOVA results:
  - Female1: F(3, 21) = 23.064, p < 0.001*
  - Female2: F(3, 21) = 55.896, p < 0.001*
  - Male1: F(3, 21) = 1.581, p = 0.224
  - Male1: F(3, 21) = 210.188, p < 0.001*



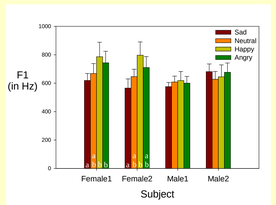
F0 range:

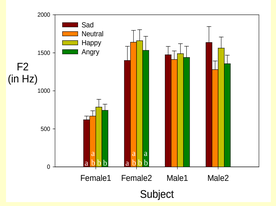
- Literature:
  - Happy & Angry: increased (Whiteside, 1999a)
- Findings in this study:
  - Happy & Angry: increased in Female2
  - Sad: increased in Female 2
  - Sad > Angry: in Female2
- One-way RM ANOVA results:
  - Female1: F(3, 21) = 1.341, p = 0.288
  - Female2: F(3, 21) = 10.6, p < 0.001*
  - Male1: F(3, 21) = 0.319, p = 0.812
  - Male1: F(3, 21) = 1.301, p = 0.3



F0 variability:

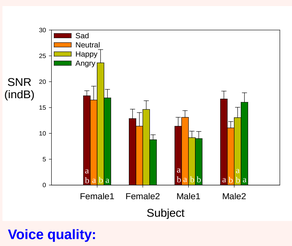
- Literature:
  - Happy & Angry: increased (Whiteside, 1999a)
- Findings in this study:
  - Happy: increased in Females 1 & 2
  - Sad: increased in Female 2
  - Happy > Angry: in Female2
  - Happy > Sad: in Female1
- One-way RM ANOVA results:
  - Female1: F(3, 21) = 5.729, p = 0.005*
  - Female2: F(3, 21) = 8.434, p < 0.001*
  - Male1: F(3, 21) = 0.511, p = 0.679
  - Male1: F(3, 21) = 2.203, p = 0.118





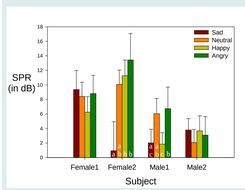
F1 & F2:

- Literature:
  - Angry: F1 increased, F2 decreased (Whiteside, 1999a)
- Findings in this study:
  - F1: Happy > Sad: in Females 1 & 2
  - F1: Angry > Sad: in Female1
  - F2: Happy > Sad: in Females1 & 2
  - F2: Angry > Sad: in Female1



Voice quality:

- Literature:
  - Angry: jitter & shimmer increased (Whiteside, 1999b)
  - Jitter: Sad < Happy & Angry (Whiteside, 1999b)
- Findings in this study:
  - Angry: decreased in Male1
  - No significant emotion effect on jitter & shimmer
  - Happy > Angry: in Female1
  - Sad> Neutral: in Male2



Voice projection power (inversely related to SPR):

- Findings in this study:
  - Sad < Angry: in Femlae2 & Male1
  - Happy < Neutral & Angry: in Male1
  - Sad < Happy: in Female2

Conclusion

Findings in the present study generally agree with the previous observation that emotional state has an effect on speech rate, loudness, pitch, and voice quality. There is also a tendency for F1 to increase with the Happy and Angry conditions, suggesting a more open articulatory setting for these conditions as found in the literature. A few inter-subject differences have also been identified. Further studies with a greater sample size are needed.

References

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70(3)*, 614-636.
Most, T., & Aviner, C. (2009). Auditory, Visual, and Auditory-Visual Perception of Emotions by Individuals With Cochlear Implants, Hearing Aids, and Normal Hearing. *Journal of Deaf Studies and Deaf Education, 14(4)*, 449-464.
Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93*, 1097-1108.
Pereira, C. (1996, 10-12 December). "Angry, Happy, Sad or Plain Neutral? The Identification of Vocal Afect by Hearing Aid Users." Paper presented at the Proceedings of the Sixth Australian International Conference on Speech S cience and Technology., Adelaide.
Peters, K. P. (2006). Emotion Perception in Speech: Discrimination, Identification, and the affects of talker and sentence variability. Unpublished A Capstone Project, Washington University, Washington.
Scherer, K. R. (1984). On the Nature and the Function of Emotion: A Component Process Approach. In P. Ekman (Ed.), *Approaches to Emotion* (pp. 293-317): Hillsdale, NJ: Erlbuam.
Whiteside, S. P. (1999a). Acoustic characteristics of vocal emotions simulated by actors. *Perceptual and Motor Skills, 89*, 1195-1208.
Whitesdie, S. P. (1999b). Note on voice and perturbation measures in simulated vocal emotions. *Perceptual and Motor Skills, 88*, 1219-1222.