USE OF PITCH FOR PROCESSING EMOTIONS

Masako Hirotani[1, 2]
[1]School of Linguistics and Language Studies, and Institute of Cognitive Science, Carleton University,
1125 Colonel By Drive, Ottawa, Ontario, Canada, K1S 5B6
[2]Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany

1. INTRODUCTION

Previous studies in both linguistics and psych- and neurolinguistics have shown that various prosodic factors contribute to listeners' processing of emotion (e.g., Davidson et al. 2003; Schirmer & Kotz 2006). The present study examined how listeners use one prosodic parameter, pitch, and its role in emotional processing. Specifically, it examined 1) whether fine distinctions in pitch height influence listeners' processing of emotion and 2) to what degree pitch information is utilized when different emotions are processed (i.e., happy vs. sad). Three off-line listing experiments were conducted to address these questions. Overall, the results indicated that pitch expansion has an effect on listeners' perception of emotion, in particular in processing happy emotion.

2. EXPERIMENT 1

Experiment 1 tested whether pitch (pitch expansion vs. compression at three different heights) of a sentence influences listeners' processing of happy and sad emotions.

2.1. Method

Participants. 60 undergraduate students from Carleton University (36 females, mean age 19.83 years old, SD 2.04) participated in the experiment. All participants were right handed, native speakers of Canadian English, with normal or corrected to normal vision, and none of the participants reported any history of neurological or hearing disorders.

Stimuli. A total of 480 English sentences were constructed. They were in a wide variety of syntactic structures but were all similar in length (3.08 s). In order to choose the best 120 sentences with neutral content, the content of the sentences was normed on a seven-point scale (1 being sad, 4 neutral, and 7 happy) by 32 Canadian English undergraduate students (19 females, mean age 22.5 years old, SD 2.78), separate from those that participated in the listening experiments (See Table 1 for example sentences).

Table. 1. Example sentences of neutral content.

1) Alice cleaned the house before she picked up her son.
2) Mick used the coat rack at the restaurant to hang his jacket.

The 120 sentences rated as having "neutral" content in the norming study were recorded at a monoral 16-bit/44.1-kHz sampling rate by a female Canadian English speaker. The experiment had a 2 x 3 design, with the first factor being whether pitch of the sentence was expanded or compressed (pitch range) and the second factor being three different heights in pitch (pitch height). For each sentence, F0 values were extracted every .01 second and converted using two equations, $[x + (x/40)*2]$ for expanded conditions and $[x - \sqrt{x}\,]$ for compressed conditions. To create three different pitch height conditions (i.e., low, mid, and high), 0 Hz, 5 Hz or 10 Hz was added to or subtracted from each of the sentences. Table 2 presents acoustic data for the stimuli.

Table. 2. Acoustic analysis of stimuli.

Condition	Mean F0 (Hz)	Max F0 (Hz)	Mean Intensity (dB)	Max Intensity (dB)
Expanded Low	173.67	328.27	65.96	81.04
Expanded Mid	177.34	331.40	65.97	81.04
Expanded High	180.94	337.74	65.99	81.04
Compressed Low	147.85	287.46	66.21	81.37
Compressed Mid	150.70	289.67	66.17	81.30
Compressed High	154.61	294.87	66.12	81.23

Procedures. Participants were tested individually in half hour sessions. After each sentence had finished playing, the participants indicated which emotion, either happy or sad, the sentence portrayed by choosing the corresponding key on the key pad.

2.2. Results

Figures 1 presents mean percent choice of "happy" for each condition. ANOVAs were conducted using error terms based on participant (F_1) and item variability (F_2). There was a significant effect of pitch range ($F_1(1,59)=72.41$, $p<.0001$, $F_2(1,119)=242.75$, $p<.0001$), which suggests that the pitch expanded conditions were associated with happy emotion more often than the pitch compressed conditions. In addition, a significant interaction of pitch range and pitch height was found ($F_1(2,118)=3.10$, $p<.01$, $F_2(2,238)=4.16$, $p<.05$). This interaction is probably attributed to the difference between the two extreme conditions in pitch height (high vs. low) for the compressed conditions.

Experiment 1 showed that sentences with a large pitch range elicit happy emotion more often than those with a small pitch range. However, as the reader may recall, Experiment 1 used a forced choice task, in which participants were asked to choose between the two options provided (either "happy" or "sad"). This may be unnatural as a task. To overcome this problem, Experiment 2 used seven options.

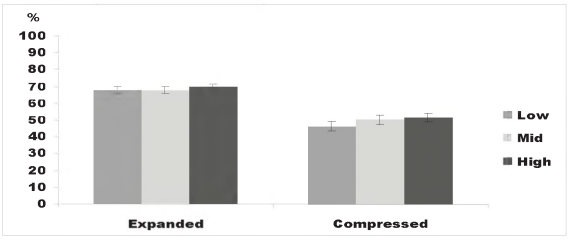

Figure. 1. Mean percent choice for "happy".

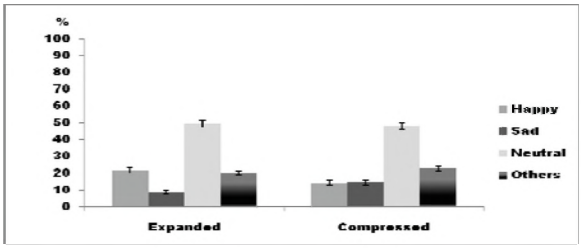

Figure. 2. Mean percent choice for emotion categories.

3. EXPERIMENT 2

3.1. Method

Participants. 60 participants (41 females, mean age 19.86 years old, SD 2.27) took part in the experiment. Participants were recruited from the same participant pool, using the same criteria, as in Experiment 1.

Stimuli and Procedures. The stimuli and procedures for the experiment were the same as those for Experiment 1 except that in this experiment, participants were provided seven emotion categories, i.e., six basic emotions (Ekman 1992) and neutral, from which they were asked to choose the emotion best associated with the sentence they heard.

3.2. Results

The results of experiment are presented in Figures 2. Data were averaged over three different conditions of pitch height (low, mid, and high) for each type of pitch range conditions (pitch expansion and compression). In addition, results for "surprise", "disgust", "fear", and "anger" were clustered into the category "others".

As shown in Figure 1, "neutral" was chosen most often for both pitch expansion and compressed conditions (about 50% of all responses). This is probably due to the fact that the stimuli were created based on the sentences with neutral content (see the stimuli section for Experiment 1). More importantly, "happiness" was chosen more often than "sadness" for the pitch expansion conditions (21.67% happiness vs. 7.75% sadness), whereas such a data pattern was not found for the pitch compression conditions. For the pitch compression conditions, "happiness" and "sadness" were chosen equally often (14.20 % happiness and 14.65% sadness). The statistical analyses (omitted due to space limitations) support the description of data presented above.

The results of Experiment 2 are consistent with those of Experiment 1, suggesting that a large pitch range is associated with happy emotion more frequently than sad.

4. EXPERIMENT 3

As a follow-up to the previous experiments, Experiment 3 tested listeners' sensitivity to subtle differences in pitch. This was to examine a relation between the listeners' choice of emotion and pitch perception.

60 undergraduate students from Carleton University (31 females, mean age 19.90 years old, SD 2.02) participated in the experiment. The participants were asked to listen to the sentence stimuli used for Experiments 1 and 2 and evaluate pitch height by using a 1-7 scale (1 being lowest and 7 highest). The results are presented in Figure 3.

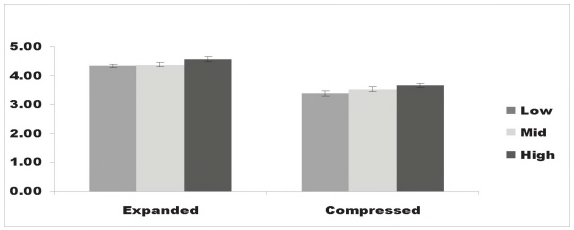

Figure. 3. Mean rating of pitch height.

The data showed a significant effect of pitch range ($F_1(1,59)=87.85$, $p<.0001$, $F_2(1,119)=903.55$, $p<.0001$) as well as pitch height ($F_1(2,118)=24.28$, $p<.0001$, $F_2(2,238)=23.62$, $p<.0001$). These results indicate that listeners were capable of perceiving the small differences in pitch associated with the present stimuli.

5. CONCLUSION

This study investigated the role of pitch in emotion processing. While listeners are sensitive to subtle pitch differences, a pitch range (pitch expansion vs. compression) plays a crucial role in associating sentences of neutral content with specific emotion categories (e.g., happy, sad).

REFERENCES

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169-200.

Davidson, R.J., Scherer, K.R., & Goldsmith, H.H. (2003). Vocal expression of emotion. *Handbook of affective sciences* (pp. 433-453). US: Oxford University Press.

Schirmer, A., & Kotz, S.A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*(1), 24-30.

ACKNOWLEDGEMENTS

The present study was supported by the research grant, Social Sciences and Humanities Research Council of Canada, Standard Research Grant # 410-2008-0987.