



Pause ratio effect in emotion ascribing

Eszter Szabó

Department of Cognitive Science, Budapest University of Technology and Economics (BME), Hungary

Abstract

In this study we investigated how pause length and pause ratio influence listeners in ascribing emotional states to the speaker, if nothing but the length of the pauses changes. We did this by taking emotionally neutral speech samples and manipulating the length of pauses. Our results show that the same speech samples were perceived to be happier, more positive, more heated, and concurrently less sad, less disgusted and less scared when pauses were shorter.

Key words: emotion, speech, pause ratio

Introduction

The relationship between emotions and speech is analyzed from many aspects. Speech signal is often analyzed phonetically. One way to investigate the relationship is through inference studies, which look at the cues that produce certain types of emotion inferences in listeners.

The most commonly measured parameters are intensity, tempo or fundamental frequency (review: Johnstone & Scherer 2000 or Scherer 2003). In most of the research only one-to-two word long, or at longest, one-sentence long speech samples are analyzed. Due to the short sample lengths in these studies, pause length and pause ratio are not relevant. Juslin and Scherer (2005) stressed that although pauses and rhythm seem to be important there remains no standardized measure for this.

However, results of some recent studies suggest that these parameters can vary in different emotional states. Deppermann and Lucius-Hoene (2005) analyzed trauma narratives in TV interviews with conversation analytic methods and noted that speech became deeper, more quiet and powerless, and lost its intonation while people were telling sad stories. They also detected more pauses, even within syntactical units. In our previous study (Szabó 2008) sad and happy emotional states were induced by music and autobiographical recall and half minute long speech samples of the recalls were analyzed. It was found that there were significant differences in average duration of silent pauses and in pause ratio (rate of pauses to the whole speech in percentage) between sad and happy states. Jovicic and his collaborates (2004) designed a Serbian emotional speech database. They asked six actors to repeat single words, short and long sentences and a passage while imitating different emotional states. Total speech segment

duration, total duration of pauses between speech segments and speech/pause ratio were measured. It was found that pause duration in passages is a more discriminative parameter between emotions than speech duration. Speech/pause ratio is a good indicator for speech rhythm, and distinguishes sadness, anger, happiness, fear, and neutral emotional states.

In the present research we asked how pause length and pause ratio influence listeners in ascribing emotional states to the speaker if nothing but the length of the pauses change.

Method

Participants

Fifty university students (16 male and 34 female) participated in the experiment. They had a mean age of 20.12, (SD 1.41). All of them were native speakers of Hungarian. Participants received university credits for their participation.

Material

In the experiment we used five emotionally neutral speech samples of the BEA Hungarian speech database (Gósy, 2008) and a previous research (Szabó, 2008). We recorded women describing themselves, their work and an ordinary day in their lives. Speech samples were taken one by one for approximately one minute each (52-66 seconds). The content was emotionally neutral and didn't contain emotional words (like "*it was a very happy part of my life*"), or laughter. The speech samples used were all monologues and were cut out from longer interviews.

The one-minute long speech samples were taken as "original" samples, and each was modified with the program "Audacity" in four different ways. All of the pauses that were longer than 100 msec were elongated by 18 or 50 percent, or were shortened by 21 or 50 percent. Modifications with 21 and 18 percent were based on our previous results (Szabó, 2008), however in a pretest listeners couldn't perceive any differences so we also increased modifications by 50 percent. Finally we had five variants of each speech sample: the original version, pauses elongated by 18 or 50 percent, and pauses shortened by 21 or 50 percent.

Procedure

Participants were asked to judge five speech samples by filling out a questionnaire. After listening to each speech sample, subjects were to rate how angry, sad, disgusted, happy, surprised, scared, positive, and heated the speaker seemed on a scale of 1-to-6.

In all cases we explained the task to the participants in detail and emphasized that subjects should rate the *whole* speech sample and not decide

too early on. They were instructed to assign a rating in the minute long pauses between the speech samples, only after listening to the full sample.

All listeners participated in the experiment alone, with the speech samples played by a computer via a headset.

Arrangement

In the experiment we had five experimental groups, each with 10 participants. Each participant judged each speaker only once, and all of the speech samples he or she listened to were modified with a different pause ratio (see Table 1.). Participants in group A listened to speech *a* with 50 % abbreviated pauses, speech *b* with 21 % abbreviated pauses, speech *c* in original version etc..

Participants belonging to the same group listened to speech samples in an arrangement called “Balanced Latin Square” (Williams, 1949), where effect of the order of targets is balanced.

Table 1. Speech sample (a–e) and pausing (0,5–1,5) in the different experimental groups.

Group A	Group B	Group C	Group D	Group E
a 0.5	a 1.5	a 1.18	a 1	a 0.79
b 0.79	b 0.5	b 1.5	b 1.18	b 1
c 1	c 0.79	c 0.5	c 1.5	c 1.18
d 1.18	d 1	d 0.79	d 0.5	d 1.5
e 1.5	e 1.18	e 1	e 0.79	e 0.5

Results

We used Spearman’s correlation to measure whether pause length has an effect on listeners’ judgments of how sad, happy, positive etc. they found the speaker. We found that as the length of pauses increased, participants rated the speaker as more sad ($r=0.204$, $p<0.01$), more disgusted ($r=0.174$, $p<0.01$), more scared ($r=0.131$, $p<0.05$). With longer pauses, the speaker was perceived as less happy ($r= -0.179$, $p<0.01$), less positive ($r= -0.153$, $p<0.05$), and emotionally less heated ($r= -0.126$, $p<0.05$).

Discussion

We demonstrated that the length of silent pauses influences listeners in attributing emotional states to the speaker. Results show that with shorter pauses, the same speech samples were perceived to be happier, more

positive, and more heated and concurrently less sad, less disgusted and less scared. Although correlations are significant, it is important to stress that the rate of correlation is not exceptionally high. This is not surprising because the content of all the monologues was emotionally neutral.

Since all of the speech samples were manipulated with all of the pause ratios, and listeners listened to all of the speakers only once, they did not have the opportunity to observe differences, or compare pause ratios between the samples. It might therefore be possible that people have an unconscious conception about the way a speaker might speak in a neutral emotional state, and by perceiving differences in pause ratios, listeners can ascribe emotional states to the speaker. If a pause pattern perception like this does exist, it may be learned and used differently in childhood.

Also, if we know that pauses are relevant for emotion attribution, there are then important implications for speech synthesis.

References

- Deppermann, A., Lucius-Hoene, G. 2005. Trauma erzählen – kommunikative, sprachliche und stimmliche Verfahren der Darstellung traumatischer Erlebnisse. *Psychotherapie und Spezialwissenschaft. Zeitschrift für Qualitative Forschung und klinische Praxis*, 1. 35–73.
- Gósy, M. 2008. Magyar spontán beszéd adatbázis – BEA. *Beszéd kutatás* 2008.116–128.
- Johnstone, T., Scherer, K. R. 2000. Vocal communication of emotion. In Lewis, M., Haviland-Jones, J. (eds.) 2000, *Handbook of Emotions, Second Edition*, 220–235, New York: Guilford Press.
- Juslin, P.N., Scherer, K.R. 2005. Vocal expression of affect. In JHarrigan, J.A., Rosenthal, R. and Scherer, K.R. (eds.) 2005. *The new handbook of methods in nonverbal behavior research*, 65-135, New York: Oxford University Press.
- Scherer, K.R. 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40, 227–256.
- Szabó, E. 2008. A szomorú és a vidám érzelmi állapot megjelenése a beszédben. *Magyar Pszichológiai Szemle*, 63, 4, 651–668.
- Williams, E.J. 1949. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research, Ser. A* 2, 149–168.