EMOTIONAL VALENCE AND PSYCHOACOUSTICS OF IRRELEVANT SPEECH

Florian Kattner and Wolfgang Ellermeier
Technische Universität Darmstadt, Germany
kattner@psychologie.tu-darmstadt.de

Abstract

To investigate whether the emotional valence of irrelevant speech has an influence on the disruption of memory performance, two variables were independently manipulated: (1) the semantic content of the background speech which consisted of lists of positive, neutral, or negative words, and (2) the prosody imposed by recording the word lists in either a happy, a neutral, or an angry voice. Fourty participants were exposed to the resulting nine combinations of background speech conditions while performing a serial-recall task. All irrelevant-speech stimuli impaired memory performance when compared to a white-noise control condition. Semantic content did not differentially affect performance, though, while prosody did in that the recordings made in an angry voice produced significantly lower recall scores than the recordings made in a happy or neutral voice. To investigate whether sound features may be identified that carry the crucial prosody information, psychoacoustic analyses of the irrelevant-speech recordings were performed. It turned out that psychoacoustical 'fluctuation strength' - the magnitude of amplitude and frequency modulations present in the signals - predicted the overall pattern of performance quite well.

Immediate serial recall of verbal, or phonologically encoded material from working memory is impaired while listening to irrelevant speech. Though memory performance was shown to be largely unaffected by semantic properties of the irrelevant speech (e.g. LeCompte, Neely, & Wilson, 1997), emotionally laden irrelevant words (particularly negative ones) impair serial recall more than neutral words do (e.g. Buchner, Rothermund, Wentura, & Mehl, 2004). It is further known that magnitude of irrelevant sound effects depends on acoustical properties of the signals such as the amount of frequency changes (Jones & Macken, 1993). Recently, psychoacousticians have claimed that 'fluctuation strength', a sensation based on

We would like to thank Katharina Ernst, Alexandra Margic, Anna Petryszyn, and Christian Wieser for collecting the data, and our colleague Wookeun Song of Brüel & Kjær Denmark for providing us with the software and advice to psychoacoustically analyze the stimuli.

hearing slow frequency and amplitude modulations (Fastl & Zwicker, 2007) might account for the size of irrelevant sound effects across a large number of investigations (Schlittmeier, Weissgerber, Kerber, Fastl, & Hellbrück, 2008). Since those modulations are thought to carry the emotional content in spoken language, it might be interesting to explicitly address the role of emotional intonations in the irrelevant speech effect. The goal of the present study was thus to study the role of the affective valence of irrelevant speech by independently varying (a) the affective meaning of the irrelevant words used (content), and (b) the intonation with which these words were pronounced (prosody).

Method

Materials

Sets of 14 neutral, 14 positive and 14 negative German words were selected from the Berlin Affective Word List (Vo, Jacobs, & Conrad, 2006) and randomly ordered in three different ways per set. One female and one male speaker generated recordings lists by uttering the words at a rate of approximately 1 word/s. In order to modulate the affective prosody of the utterances, the speakers were instructed to articulate each word list once in a neutral voice, once in a happy voice and once in an angry voice, resulting in a total of 54 recordings, 27 (3 randomizations $\times$ 3 types of content $\times$ 3 prosody instructions) from the female speaker and 27 from the male speaker. The recordings were normalized and diotically played back over headphones (Beyerdynamic DT990) at a mean SPL of $65.7 (SD = 2.9)$ dB. White noise having a level of 63 dB SPL was used as a non-speech control condition.

Participants and procedure

19 female and 21 male subjects (mean age: 32 years; range: 19 - 67) were recruited for individual testing in a sound-attenuated listening room. They were randomly assigned to either the male-speaker or the female-speaker recordings. On each trial, the participants had to memorize a random sequence of the digits 1 to 9. Each digit was presented on the screen for 1 s, immediately followed by the next digit. After the last digit, a blank screen was displayed for 5 s. The trial ended with the display of a numeric pad in which the subject had to enter the sequence of digits in order of appearance. During the 14 s from the start of the first digit to the appearance of the response pad, one of 10 sounds (noise, or 3 types of content $\times$ 3 types of prosody) was played back. Subjects were instructed to ignore the sounds. Each sound was presented during 8 trials resulting in a total of 80 trials.

Results

Irrelevant speech effects

Figure 1 shows recall performance in the ten experimental conditions. A 3×3 analysis of variance with content (neutral, positive, and negative words) and affective prosody (neutral, happy, and angry voice) as within-subject factors[1] revealed a significant main effect of prosody

[1] Trials in which white noise was played back were not included in the analysis of variance.

$[F(2, 78) = 4.52; p < .02]$, but no main effect of content $[F(2, 78) = 1.12; p = .33]$ and no interaction [F(4,156)=1.09; p=.36]. On average, serial recall was 53.0% correct when the intonation of background speech was neutral, 52.9% correct when it was pronounced with a happy voice, and 50.0% correct when it was delivered in an angry voice. In the white-noise control condition, serial recall was 62.6% correct which differed significantly from the mean performance in the irrelevant speech conditions $[F(1, 39) = 32.75; p < .001]$.

Psychoacoustic analyses

To verify that the recordings made with different intonations actually produced reliable acoustic differences, they were analyzed (using Brüel & Kjær PULSE 14 software) with respect to objective psychoacoustical indices (loudness, sharpness, roughness, and fluctuation strength; see Fastl & Zwicker, 2007) that may be derived from the signals (compare Table 1). An analysis of variance using the 54 recordings as independent replications of the three prosody conditions (neutral, happy, and angry voice) revealed statistically significant differences in psychoacoustical fluctuation strength between them $[F(2, 36) = 20.83; p < .001]$. By contrast, the emotional content of the word lists did not affect fluctuation strength $[F(2, 36) = 1.46; p = .25]$.

Figure 1. Serial recall performance as a function of prosody and content of irrelevant background speech, compared to a non-speech control condition (noise).

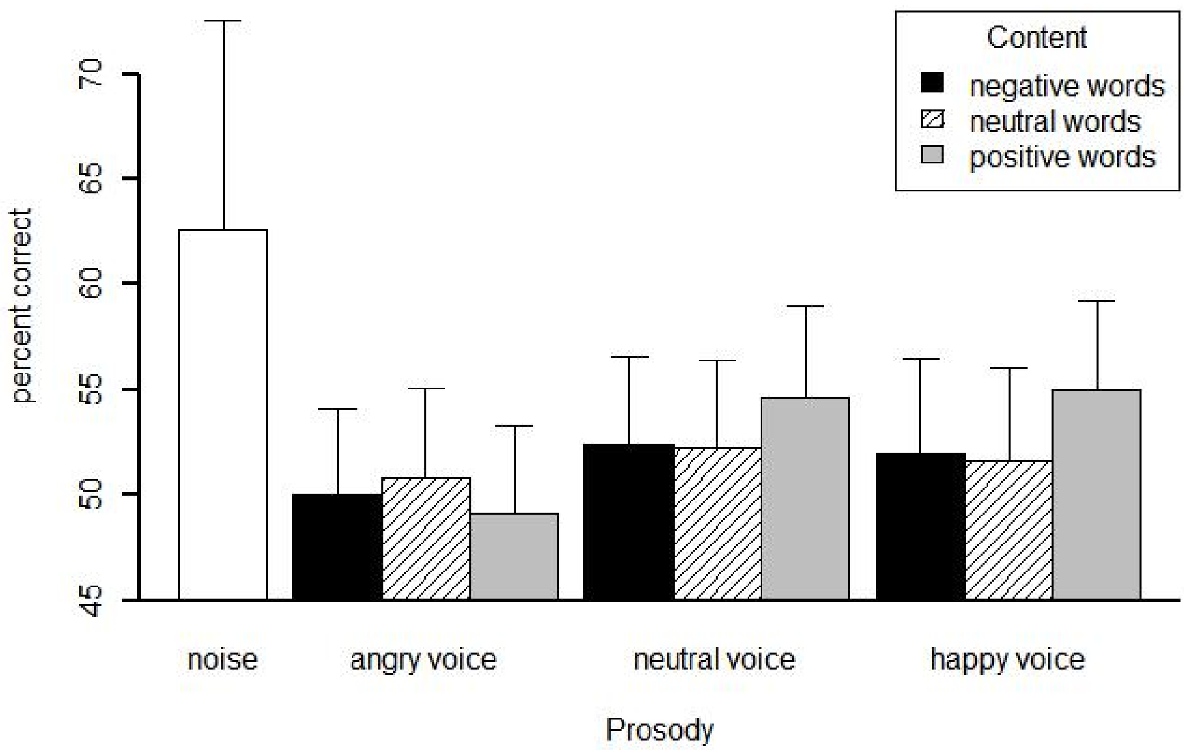

Table 1: Mean psychoacoustic measures of angry, happy and neutral utterances (*SD* in parentheses).

	Voice		
	angry	happy	neutral
Mean loudness [sone]	7.80 (1.51)	8.02 (2.22)	8.35 (1.97)
Fluctuation strength [vacil]	2.70 (0.38)	2.41 (0.33)	2.11 (0.24)
Roughness [asper]	1.76 (0.22)	1.70 (0.29)	2.01 (0.14)
Sharpness [acum(Z)]	1.81 (0.19)	1.64 (0.13)	1.65 (0.24)

More importantly, recall performance was significantly correlated with fluctuation strength ($r = .72; p < .05$). A stepwise multiple linear regression of recall performance on the psychoacoustic measures (fluctuation strength, sharpness, roughness and mean loudness) showed the best fit when using fluctuation strength as the only predictor ($R^2 = .52; p < .05$).

Discussion

The present results show that the intonation of irrelevant speech differentially affects serial recall. Particularly, word lists spoken with an angry voice reliably produced greater working memory impairment than the same lists spoken in a happy or neutral voice. In contrast to Buchner et al. (2004), however, we did not find an effect of the emotional meaning of words. This might be due to a confounding of content and intonation in that earlier study, or to the fact that adjectives might have different psychoacoustic properties than nouns.

Since (a) the speaker's intonation (angry, happy and neutral voice) lead to differences in psychoacoustical fluctuation strength and (b) fluctuation strength was the best signal-based predictor of memory impairment, the results are in line with the hypothesis that the magnitude of the irrelevant speech effect is a function of the amplitude, or - more likely - frequency modulations present in the irrelevant sound (Schlittmeier et al., 2008).

References

Buchner, A., Rothermund, K., Wentura, D., & Mehl, B. (2004). Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory and Cognition*, *32*, 722-731.

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics - facts and models* (3rd ed.). Heidelberg: Springer.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 369-381.

LeCompte, D. C., Neely, C. B., & Wilson, J. R. (1997). Irrelevant speech and irrelevant tones: The relative importance of speech to the irrelevant speech effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 472-483.

Schlittmeier, S. J., Weissgerber, T., Kerber, S., Fastl, H., & Hellbrück, J. (2008, June 29 - July 4). An algorithm modelling the Irrelevant Sound Effect (ISE). In *Acoustics '08.* Paris.

Vo, M.-L., Jacobs, A., & Conrad, M. (2006). Crossvalidating the Berlin affective word list (BAWL). *Behavior Research Methods*, *38*, 606-609.