Cognition 115 (2010) 202–206

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Auditory emotional cues enhance visual perception

René Zeelenberg *, Bruno R. Bocanegra

Erasmus University Rotterdam, Department of Psychology, Woudestein, T13-31, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 30 July 2009
Revised 13 November 2009
Accepted 9 December 2009

Keywords:
Emotion
Cross-modal effects
Word recognition
Visual perception

ABSTRACT

Recent studies show that emotional stimuli impair performance to subsequently presented neutral stimuli. Here we show a cross-modal perceptual enhancement caused by emotional cues. Auditory cue words were followed by a visually presented neutral target word. Two-alternative forced-choice identification of the visual target was improved by emotional cues as compared to neutral cues. When the cue was presented visually we replicated the emotion-induced impairment found in other studies. Our results suggest emotional stimuli have a twofold effect on perception. They impair perception by reflexively attracting attention at the expense of competing stimuli. However, emotional stimuli also induce a nonspecific perceptual enhancement that carries over onto other stimuli when competition is reduced, for example, by presenting stimuli in different modalities.

1. Introduction

Ample evidence suggests that emotion affects the processing of visual information. For example, in a rapid serial visual presentation (RSVP) paradigm in which targets are embedded in a stream of distractors, Anderson and Phelps (2001) found that emotional words were more often correctly identified than neutral words. Likewise, Zeelenberg, Wagenmakers, and Rotteveel (2006) obtained evidence for enhanced identification of briefly presented and masked emotional words. Studies with pictures and faces have also obtained evidence for emotion-induced enhanced identification (Alpers & Pauli, 2006).

Other results indicate that emotional stimuli have a negative impact on the perception of subsequently presented neutral stimuli. For instance, several recent studies have shown that the presentation of an emotional stimulus in an RSVP paradigm impairs the identification of a later presented neutral target stimulus (e.g., Mathewson, Arnell, & Mansfield, 2008; Most, Chun, Widders, & Zald, 2005; Most, Smith, Cooter, Levy, & Zald, 2007). These 'emotion-induced blindness' effects have been obtained both when the emotional stimulus was task-relevant (i.e., when the emotional stimulus was a first target in the RSVP stream) or task-irrelevant (i.e., when the emotional stimulus was a distractor). Recently, negative effects of emotional cues on target performance have also been obtained in perceptual identification (Bocanegra & Zeelenberg, 2009) and speeded word recognition tasks (Zeelenberg, Bocanegra, & Pecher, submitted for publication).[1]

Emotion-induced attentional processes may explain why an emotional stimulus impairs performance to a neutral target presented shortly thereafter (Anderson, 2005; Most et al., 2005). If limited attentional resources are allocated to the emotional stimulus, fewer resources will be available for the processing of the neutral target thereby causing an impairment in target identification. Furthermore, if the allocation of processing resources to emotional stimuli happens involuntarily, detrimental carry-over effects may be obtained even when the emotional stimulus is task-irrelevant. Attentional processes may also explain why emotional stimuli themselves are correctly identified

* Corresponding author. Tel.: +31 10 408 9560; fax: +31 10 408 9009.
E-mail address: zeelenberg@fsw.eur.nl (R. Zeelenberg).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.12.004

[1] In the studies cited thus far the cue and target were always presented on a single location throughout the experiment. Discussion of paradigms with spatial uncertainty in which the target and cue are presented on either the same or on different locations (dot probe task) is deferred to Section 5.

more often than neutral stimuli. The reflexive allocation of attention to emotional stimuli will benefit their identification. Within this framework, one might say that the benefits for emotional targets themselves and the deficits for subsequently presented neutral targets are two sides of the same coin. That is, emotional stimuli benefit from attention allocation at the expense of competing stimuli.

There is, however, some evidence suggesting that emotion-induced effects on perception can occur independent of the allocation of attention. For example, masked emotional stimuli presented below the threshold for conscious awareness have been shown to enhance activation of the amygdala (Naccache et al., 2005), a subcortical structure involved in the modulation of perception by emotion (e.g., Anderson & Phelps, 2001; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). Furthermore, emotional modulations of ERPs can arise as early as 60–90 ms after stimulus onset (Stolarova, Keil, & Moratti, 2006), which are unlikely to be caused by selective attentional processes (Di Russo, Martinez, & Hillyard, 2003). Finally, Phelps, Ling, and Carrasco (2006) found that a fearful face cue enhances perception of a subsequent Gabor stimulus irrespective of (spatial) attention.

If indeed it is the case that emotion can enhance perception independent of attention allocation, then it may be possible to obtain beneficial effects of emotional cue words on the identification of a subsequent neutral target word. The argument here is that two underlying opposing forces may be at work in emotional cueing paradigms: (a) emotional cues attract attentional resources, harming identification of a subsequently presented neutral stimulus and (b) emotional cues cause a temporary boost in perceptual processing that may improve identification of a subsequently presented neutral stimulus. Whether a deficit or benefit will be obtained depends on the relative influence of these two opposing forces (cf. Zeelenberg, Wagenmakers, & Shiffrin, 2004). Our contention is that under commonly used procedures, such as the RSVP paradigm, performance is dominated by the process that reflexively attracts attention to the emotional stimulus thereby harming performance to target stimuli presented shortly thereafter. However, if we were able to reduce temporal competition for attention between cue and target, we should be able to demonstrate a beneficial effect of emotional cues on a subsequently presented neutral stimulus.

In the present study we aimed to reduce the impact of the emotion-induced 'stealing' of processing resources by presenting the cue and target in different modalities (i.e., auditory and visual). Emotion-induced deficits in RSVP and masked perceptual identification have been obtained in experiments in which the cue and target were presented in the same modality (i.e., the visual modality). Evidence from the attention literature suggests that stimuli presented in the same modality interfere more with each other than stimuli presented in different modalities (e.g., Duncan, Martens, & Ward, 1997; Schupp et al., 2008; Soto-Faraco & Spence, 2002). This suggests that the emotion-induced interference effect may be reduced by presenting cue and target in different modalities. If the interference effect is indeed reduced then the effect of a temporary boost in perceptual processing may become manifest and emotional cues may improve identification of a subsequently presented target.[2]

2. Experiment 1

In Experiment 1, we investigated the influence of an auditory emotional cue word on the identification of a subsequently presented neutral target word in a masked visual identification paradigm. The cross-modal presentation of the cue was expected to reduce attentional competition between cues and targets. As a consequence, we expected that emotional cues would benefit the identification of the neutral targets.

2.1. Methods

2.1.1. Participants

Thirty-two students participated in Experiment 1. The participants in all three experiments of the present study were students at the Erasmus University Rotterdam participating for course credit or a small monetary reward. They gave informed consent and no one participated in more than one of the experiments reported here.

2.1.2. Stimulus materials

The cue words consisted of 104 spoken words: 52 (mostly negative) high-arousal emotional words (e.g., *rape*, *torture*, *kill*, *fuck*) and 52 low-arousal neutral words (e.g., *owl*, *clock*, *shirt*, *hay*). The words were spoken by a female speaker with a neutral prosody and digitally recorded on a computer with a sampling rate of 44,100 Hz. The selected emotional and neutral words were matched in terms of the number of letters (mean = 7.07), mean number of phonemes (6.17 and 6.15), mean number of syllables (2.31 and 2.27), proportion of words from each of three grammatical classes (i.e., verbs, nouns or adjectives), mean log word-frequency per million (0.946 and 0.958), and pronunciation duration (774 ms and 769 ms), all $ps > .40$. An independent group of subjects provided arousal and valence ratings for the cue words. The emotional and neutral words differed in terms of their arousal ratings (8.06 vs. 4.31) and valence ratings (3.11 vs. 5.69), emotional cue words being more arousing and negative than neutral cue words, both $ps < .01$. An additional set of 104 neutral test words (mean log frequency per million = 1.17; Baayen, Piepenbrock, & Gulikers, 1995) were selected to serve as a target and foil in the two-alternative forced-choice perceptual identification procedure.

Four counterbalanced lists were created such that, across participants, each neutral and emotional cue word was used equally often as a cue, and each neutral test word served both as a target and foil.

[2] Note that, for the potential benefit of emotional cues to become manifest, it is not necessary that cross-modal presentation eliminates the interference process entirely. It only requires the interference effect to be reduced to a sufficient extent, such that the emotion-induced benefit outweighs the emotion-induced deficit.

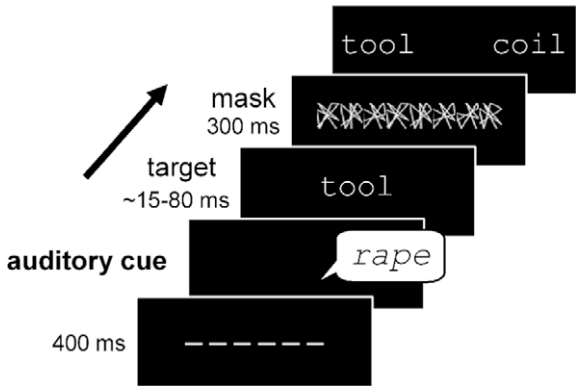

Fig. 1. Illustration of the display sequence in Experiment 1. Visual cue words (rather than auditory ones) were presented in Experiments 2 and 3. See text for details.

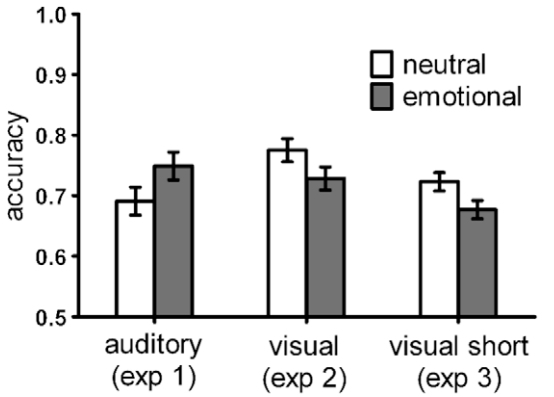

Fig. 2. Percentages of correctly identified visual targets as a function of the emotional status and presentation modality of the cue in Experiments 1, 2 and 3. Error-bars are within-subject standard errors of the mean (Loftus & Masson, 1994).

2.1.3. Apparatus and procedure

The cue words were presented binaurally over headphones (Vivanco type SR 60). The visual stimuli were presented on a gamma-corrected Iiyama 21-inch Vision Master monitor at a 200 Hz refresh-rate and 800 × 600 pixel (Courier New 18 point font, 3–5° visual angle). The presentation procedure for Experiment 1 is illustrated schematically in Fig. 1. Each trial consisted of the presentation of a warning signal, an auditory cue word, a visual target word presented immediately after the offset of the cue (target presentation times were adjusted individually), a backward mask and two choice alternatives (a target word and a foil word). The choice alternatives were presented side-by-side and participants were asked to indicate which one corresponded to the target word by pressing the *z* or *m* key on the keyboard (choice alternatives remained visible until the participant responded). For each trial, the location of the correct choice (left-hand or right-hand word) was determined randomly.

The session started with four practice trials in which targets were flashed for 100 ms (so that the target was clearly visible) to make sure the requirements of the experiment were clear to the participant. Next, 60 calibration trials were presented to estimate the flash-time resulting in 70% correct performance (see Wagenmakers, Zeelenberg, and Raaijmakers (2000) and Zeelenberg et al. (2006), for details). Target presentation times in the main experiment were restricted to a range of 15–100 ms.

The calibration trials were followed by 52 experimental trials. For each participant, 26 visual targets were preceded by an auditory neutral cue word and 26 visual targets were preceded by an auditory emotional cue word. Stimuli were presented in a different random order for each participant.

2.2. Results and discussion

Fig. 2 shows the percentages of correctly identified target words. Visually presented targets words were more often correctly identified when they were preceded by an auditory emotional word than when they were preceded by an auditory neutral word, $t(31) = 2.41$, $p = .022$, $\eta^2 = .16$. Thus, we obtained an emotion-induced enhancement in the perception of visually presented words.

3. Experiment 2

In Experiment 1, we showed that the auditory presentation of an emotional cue word enhanced perception of a subsequently presented visual target word. This finding stands in contrast to previously obtained results in which emotional stimuli that precede the presentation of a target stimulus impair target identification (e.g., Mathewson et al., 2008; Most et al., 2005, 2007). In Experiment 2, we tested whether an impairment would also be obtained with our stimuli and procedures when the cue word was presented visually.

3.1. Methods

Thirty-two students participated. All aspects of the methods were identical to that of Experiment 1, except that cue words were presented visually (cue and target words were presented on the same central location of the screen with a 0 ms ISI). The presentation duration for each word was matched to that of the presentation duration of the spoken word in Experiment 1. For example, pronunciation of the word '*angst*' (*fear*) in Experiment 1 took 630 ms. Therefore, in Experiment 2 visual presentation of the word '*angst*' lasted 630 ms.

3.2. Results and discussion

In contrast to Experiment 1, visually presented target words were less often correctly identified when they were preceded by an emotional word than when they were preceded by a neutral word, $t(31) = 2.37$, $p = .024$, $\eta^2 = .15$ (see Fig. 2). The emotion-induced impairment found in Experiment 2 replicates and extends previous findings, but contrasts with the findings of Experiment 1 in which we found an emotion-induced perceptual enhancement.

One potential complication in comparing Experiments 1 and 2 is that, even though presentation times for cues were

carefully matched, the meaning of cues may have become activated at different points in time with auditory and visual presentation. With visual presentation of the cue word all stimulus information is presented simultaneously whereas with auditory presentation of the cue stimulus information gradually becomes available over time. Therefore, one might argue that the difference in results between Experiments 1 and 2 could be due to differences in the point in time in which semantic information (and hence emotional content) becomes available, rather than the presentation of the cue and target in different modalities. Note, however, that in a well-known cross-modal priming study, Zwitserlood (1989) showed that the meaning of a spoken word becomes available quickly after the onset of its presentation (i.e., before the entire word has been presented). Thus, just as with visual presentation (e.g., Duňabeitia, Carreiras, & Perea, 2008; Pecher, Zeelenberg, & Wagenmakers, 2005) semantic information becomes available before processing of the input has been completed.

Nevertheless it seems likely that the meaning of cue words became available somewhat later with auditory presentation than with visual presentation. To ease any concerns that this may have caused the different results obtained in Experiments 1 and 2, we ran an additional experiment in which the cue was presented visually, but with a much shorter SOA (200 ms rather than the average 771 ms SOA in Experiment 2). If the benefit for emotional cues in auditory presentation was somehow due to the meaning of the cue becoming available only shortly before presentation of the target word, then we would expect a benefit with visual presentation as well. If, however, the benefit with auditory presentation was due to the cue being presented in a different modality than the target we should again find a deficit for visually presented cues.

4. Experiment 3

4.1. Methods

Thirty-two students participated. The experiment was identical to Experiment 2 except that all cue words were presented for 200 ms.

4.2. Results

In Experiment 3, we again found that visually presented target words were less often correctly identified when they were preceded by a visually presented emotional word than when they were preceded by a visually presented neutral word, $t(31) = 2.83$, $p = .008$, $\eta^2 = .21$ (see Fig. 2).

5. General discussion

The experiments of the present study show that the effect of an emotional cue word on visual perception reverses as a function of the presentation modality of the cue. With visual presentation of both the cue and the target emotional cue words impaired subsequent target identification, replicating and extending previous findings (e.g., Mathewson et al., 2008; Most et al., 2005, 2007). Most importantly, however, with auditory presentation of the cue and visual presentation of the target, emotional cue words improved subsequent target identification. This demonstrates for the first time that an emotional auditory stimulus can enhance the threshold identification of a neutral visual stimulus.

Studies using a dot probe paradigm (e.g., Mogg & Bradley, 1999; Pourtois, Grandjean, Sander, & Vuilleumier, 2004) have shown that emotional cues speed up responding to a subsequent neutral stimulus but in these studies the location of the target varied from trial to trial (usually either to the left or to the right of fixation). When the cue and target are presented on the same location (valid condition) emotional cues enhance performance; when the cue and target are presented on different locations (invalid condition) emotional cues impair performance. Emotional modulations in the dot probe task likely reflect the orienting of attention to the location of the emotional stimulus (as well as a failure to disengage from that location; Fox, Russo, Bowels, & Dutton, 2001).

Dot probe paradigms differ from the present paradigm on several dimensions. Perhaps most important, in the present experiments the target was always presented at the same location (i.e., there was no spatial uncertainty as to the location of the target). Studies in which the cue and target are presented on the same location throughout the experiment have consistently obtained perceptual deficits for neutral targets preceded by emotional cues (e.g., Mathewson et al., 2008; Most et al., 2005, 2007) a finding we replicated in Experiments 2 and 3 of the present study. These findings contrast with those of the dot probe paradigm where the presentation of an emotional cue on the same location as the target enhances performance. Also of interest is a recent cross-modal dot probe study. In this study, Brosch, Grandjean, Sander, and Scherer (2008) presented pairs of meaningless utterances (pseudowords) over headphones in such a manner that the sounds seemed to originate from a specific location in space. One member of the pair was pronounced with neutral prosody and the other member of the pair was pronounced with angry prosody. Subsequently, a dot probe was presented on a computer screen and participants indicated on which side (left or right) the dot appeared. Subjects responded faster on congruent trials (i.e., when the angry prosody and dot were on the same side) than on incongruent trials. Thus, in contrast to the results of the present study, the dot probe paradigm yields similar results irrespective of whether cue and target are presented in different modalities or in the same modality.

The present results clearly demonstrate that the beneficial effects of emotional stimuli are not tied to the emotional stimulus itself; rather we obtained evidence for a more general boost in the visual processing of words irrespective of the allocation of spatial attention (see Phelps et al. (2006), for a related finding with Gabor patterns). It is often thought that the amygdala is involved in the emotional modulation of perception (e.g., Anderson & Phelps, 2001). Given that the amygdala projects to all levels of the ventral visual stream emotion may affect early as well as later stages of visual processing (Vuilleumier, 2005).

Thus, our benefits may have arisen at multiple stages of visual processing.

We argue that previous studies (and Experiments 2 and 3 of the present study) obtained emotion-induced deficits because the latent emotion-induced perceptual enhancement was obscured by the preferential allocation of visual temporal visual attention to the emotional cue at the expense of the target stimulus. By presenting the cue and target in different modalities we were able to reduce the detrimental effect of the cue, thereby allowing the beneficial effect of emotional cueing to become manifest. Recently, our lab also found evidence for an emotion-induced benefit when a long (1075 ms) SOA was used and also with a short SOA when emotional cues were masked so that subjects were not aware of the emotional cue (Bocanegra & Zeelenberg, 2009), further bolstering the idea that emotional cues can enhance perception of subsequently presented target stimuli when temporal competition between the cue and the target is reduced. It is also interesting to point out that the use of visually dissimilar stimuli in the Phelps et al. (2006) study may have reduced competition (cf. Einhäuser, Koch, & Makeig, 2007) and contributed to their finding of an emotion-induced benefit. Thus, we argue that emotional cueing effects are the net result of two opposing forces. Whether an emotion-induced impairment or enhancement will be observed depends on the relative influence of these two forces.

References

Alpers, G. W., & Pauli, P. (2006). Emotional pictures predominate in binocular rivalry. *Cognition & Emotion, 20*, 596–607.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General, 134*, 258–281.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411*, 305–309.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*. Philadelphia: Linguistic Data Consortium.

Bocanegra, B. R., & Zeelenberg, R. (2009). Dissociating emotion-induced blindness and hypervision. *Emotion, 9*, 865–873.

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2008). Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody. *Cognition, 106*, 1497–1503.

Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex, 13*, 486–499.

Duñabeitia, J. A., Carreiras, M., & Perea, M. (2008). Are coffee and toffee served in a cup? Ortho-phonologically mediated associative priming. *Quarterly Journal of Experimental Psychology, 61*, 1861–1872.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387*, 808–810.

Einhäuser, W., Koch, C., & Makeig, S. (2007). The duration of the attentional blink in natural scenes depends on stimulus category. *Vision Research, 47*, 597–607.

Fox, E., Russo, R., Bowels, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130*, 681–700.

Loftus, G. R., & Masson, M. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

Mathewson, K. J., Arnell, K. M., & Mansfield, C. A. (2008). Capturing and holding attention: The impact of emotional words in rapid serial visual presentation. *Memory & Cognition, 36*, 182–200.

Mogg, K., & Bradley, B. P. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition & Emotion, 13*, 713–740.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review, 12*, 654–661.

Most, S. B., Smith, S. D., Cooter, A. B., Levy, B. N., & Zald, D. H. (2007). The naked truth: Positive, arousing distractors impair rapid target perception. *Cognition & Emotion, 21*, 964–981.

Naccache, L., Gaillard, R., Adam, C., Hasboun, D., Clémenceau, S., Baulac, M., et al. (2005). A direct intracranial record of emotions evoked by subliminal words. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 7713–7717.

Pecher, D., Zeelenberg, R., & Wagenmakers, E. J. (2005). Enemies and friends in the neighborhood: Orthographic similarity effects in semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 121–128.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science, 17*, 292–299.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex, 14*, 619–633.

Schupp, H. T., Stockburger, J., Bublatzky, F., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2008). The selective processing of emotional visual stimuli while detecting auditory targets: An ERP analysis. *Brain Research, 1230*, 168–176.

Soto-Faraco, S., & Spence, C. (2002). Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology, 55A*, 23–40.

Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: Evidence for sensory plasticity in early affective perception. *Cerebral Cortex, 16*, 876–887.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Science, 9*, 585–594.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience, 7*, 1271–1278.

Wagenmakers, E. J., Zeelenberg, R., & Raaijmakers, J. G. W. (2000). Testing the counter model for perceptual identification: Effects of repetition priming and word frequency. *Psychonomic Bulletin & Review, 7*, 662–667.

Zeelenberg, R., Bocanegra, B. R., & Pecher, D. (submitted for publication). Emotion-induced impairments in speeded word recognition.

Zeelenberg, R., Wagenmakers, E. J., & Rotteveel, M. (2006). The impact of emotion on perception: Bias or enhanced processing? *Psychological Science, 17*, 287–291.

Zeelenberg, R., Wagenmakers, E. J., & Shiffrin, R. M. (2004). Nonword repetition priming in lexical decision reverses as a function of study task and speed stress. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 270–277.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32*, 25–64.