Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Does emotional prosody affect word recognition in French?

Caterina Petrone[1], *Arina Antonenko*[1], *Sophie Dufour*[1]

[1] CNRS, Aix-Marseille Université, LPL, UMR 7309, Aix-en-Provence, FR

caterina.petrone@lpl-aix.fr, amilietou@gmail.com sophie.dufour@lpl-aix.fr

Abstract

In the word-processing literature, there is extensive evidence that the semantic context facilitates word recognition. However, much less is known about the potential role of contextual prosodic information. We test whether, in French, semantically neutral sentences (e.g., *She is taking the train*) uttered with emotional prosody (angry, happy) facilitate the recognition of negative emotional words (e.g., *war*). The sentences and the words were semantically unrelated. Two studies were run. The first study employs a unimodal priming paradigm, in which both the sentences and the target words were presented auditorily. Target words were produced with neutral prosody. Data on forty-five listeners showed no priming effects, possibly because of an incongruency between the neutral prosody on the target words and their emotional valence. The second experiment is a preliminary study on fifteen listeners employing an auditory-visual priming paradigm. Here, we found an affective priming effect: Negative words that were more strongly associated with anger were more easily recognized when preceded by angry than by happy prosody. Hence, prosodic context influences word processing depending upon whether prosody and words share the same emotion category, not just the same valence.

Index Terms: prosody, emotion, affective priming, French

1. Introduction

Many studies have shown that the semantic context influences word recognition [1, 2]. Faster reaction times (RT) and higher accuracy are usually found on target words that are related to the preceding semantic context. In addition to the semantic context, the emotional context affects word recognition (see review in [3]). Affective priming effects have been reported with faster responses on target words (*delightful*) that are preceded by a prime word of the same valence (*gift-delightful*) in comparison to target words that are preceded by a prime word of opposite valence (*death-delightful*). This effect seems to be independent of the semantic relatedness between the prime and the target words [4] and it might be explained by, e.g., spreading activation from the prime to affectively congruent targets or by listeners' tendencies to judge the compatibility of the affective components of the prime-target pair [5, 6].

So far, few studies have tested whether prosody provides an emotional context influencing word recognition. [7] conducted a cross-modal priming experiment in German, in which auditory prime sentences (*Yesterday she got her final exam*) uttered with happy or sad prosody were followed by visually presented target words. Target words were either positive (*success*) or negative (*failure*) and they were semantically associated to the previous primes. In the matching condition, the target and the prime had the same valence (*failure* preceded by sad intonation), while in the mismatching condition, the target and the prime had an opposite valence (*failure* preceded by happy prosody). Shorter reaction times were found for words in the matching condition, thus arguing for a role of emotional prosody during word recognition. However, the effects of emotional prosody have proven to be mediated by other extra-linguistic factors, such as the sex of the listeners [7] or task instructions, i.e., whether listeners are explicitly asked to take emotional prosody into account [8]. More recently, [9] also used the cross-modal priming paradigm to explore the effect of emotional prosody on semantically unrelated prime-target pairs in American English. This allowed the authors to examine prosodic influence alone, i.e., in the absence of a semantic context that has likely contributed to the observed effects in [7]. Target words consisted in words with negative (*furious*) or positive (*smile*) valence, and primes consisted in single words with neutral meaning (*pineapple*) uttered with negative (angry), positive (happy), or neutral prosody. Each participant heard only two types of prosody (angry and neutral or happy and neutral), the neutral prosody being the control condition. Affective priming effects emerged more clearly for negative than for positive target words. In particular, for negative words, the magnitude of the priming effects was larger for words having stronger association strength to anger than for negative words having lower association strength to anger.

In this paper, we attempted to provide additional evidence for a role of emotional prosody on French word recognition. While in previous work [7, 8, 9], each listener was exposed to only two types of priming conditions (matching vs. mismatching in [7, 8]; matching vs. neutral in [9]), here participants were tested on three priming conditions (matching, mismatching and neutral). This allowed us to examine the directionality of the priming effect in the matching and mismatching conditions when compared to the neutral condition. Based on previous studies, we expected to find a facilitatory priming effect in the matching condition, i.e., faster RTs in that condition in comparison to the neutral condition. In contrast, we expected either an inhibitory priming effect in the mismatching priming condition, i.e., slower RTs in that condition in comparison to the neutral condition or no effect at all when compared to the neutral condition. Note that to the best of our knowledge, it is the first time that the mismatching condition is compared to the neutral condition, and at a more theoretical level such a comparison would allow us to tease apart between spreading activation or compatibility judgment mechanism between the primes and the targets to account for the affective priming effect. Indeed, an inhibitory priming effect could argue in favor of the intervention of a congruency verification mechanism between the emotional valences of the primes and the targets [3], which lead to an interference effect in case of an incongruency. In

10.21437/SpeechProsody.2022-104

contrast, no difference between the mismatching and the neutral condition would argue in favor of a spreading activation mechanism during prime processing to affectively congruent target words.

The impact of emotional prosodic context was evaluated in the absence of a semantically relation between the target words and the preceding context. On the basis of study by [9], we focused on negative words, for which stronger effects of emotional prosody were observed. In a first experiment, we used a unimodal priming paradigm, in which the primes and the target words were presented auditorily. Target words were produced with neutral prosody, and the primes consisted in neutral sentences uttered with different prosodies (angry, happy, neutral). Because we were interested in spoken word recognition, a first experiment (Experiment 1) was conducted under unimodal presentation, that is with both primes and targets presented auditorily. To anticipate our results, we found no effects of emotional prosody in this specific priming procedure. This may be due to an incongruency between the neutral prosody of the target words and their emotional valence [see in the discussion section], which might have factored out the effect of the emotional/prosodic context. Hence, an additional, preliminary study (Experiment 2) was run on exactly the same material as in Experiment 1, but in which, as in previous studies [7, 8, 9], the targets words were visually rather than auditorily presented.

2. Experiment 1

2.1. Materials and Methods

2.1.1. Stimuli

Forty-five negative target words were selected from *Lexique 3*, a lexical database for French [10]. The score of emotional valence in this database ranged from very negative (-5) to very positive (+5). We included only words with a score higher than |1.5|.

For post-hoc analysis purpose and in order to evaluate the impact of the association strength of the negative words with the emotional category of "anger", an online survey was run, in which 75 French participants rated each word for its degree of representativeness of anger on a 5-point scale (1= "not at all"; 5="absolutely"). Based on this survey, we split reaction times (RTs) for words which were rated either as having very low (i.e., lower or equal to 2) or a very high (higher or equal to 3) association strength score. For instance, the negative word *guerre* "war" had a stronger association strength score to anger (4.14) than the negative word *araignée "*spider" (1.35). The number of words with low vs. high association strength was similar (low = 13 words; high = 12 words).

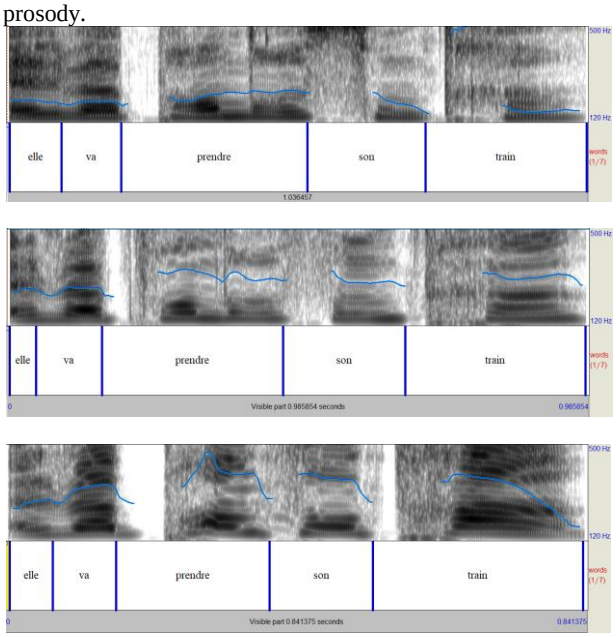
Concerning the primes, five semantically neutral sentences were produced with two emotional prosodies (happy and angry) and with neutral prosody (control condition). The five sentences were similar in length (5 to 8 syllables) and had the same syntactic structure (a subject followed by a verb and a direct or indirect object, e.g., *Elle va prendre son train* « She is going to take her train »). The subject was always a third-person singular pronoun (« she/he »).

2.1.2. Acoustic recordings

Both the targets and the primes were recorded in an anechoic room at Laboratoire Parole et Language. The targets were recorded by a French prosodist, who produced each word with the same neutral intonation (i.e., as broad focus statement with a utterance final L%). The primes were recorded by a professional French actress. For eliciting the three prosodies, each target sentence was preceded by an appropriate pragmatic context.

Prior to the experiment, the sentences were validated in their written and oral forms to ensure that their semantics was neutral and that their prosody was representative of each target prosodic pattern. For the validation of the written forms, 10 participants were asked to read each sentence and to rate its emotional valence on a 7-point Likert scale (1= "negative"; 7="positive"). For the oral forms, 20 different listeners were asked to (i) identify the emotion conveyed by the speaker (ii) to evaluate the plausibility of the sentences, by taking into account both their prosody and their words on a 5-point Likert scale (1="not at all"; 5= "absolutely"). The validation tasks were made via two online surveys. Participants engaged in the validation tasks have not taken part in the main experiment. For validation of the written sentences, only sentences which were judged semantically neutral were retained (i.e., with a score for valence between 3 and 5). For prosodic validation, the accuracy rate was greater than 75% emotion recognition, which is in line with literature on vocal expressions of emotion [11]. The combination of prosody and words was judged as highly plausible (score higher than 3). The intensity level was normalized across stimuli. Spectrograms (with pitch tracked) of three prime sentences are shown in Fig. 1.

Figure 1. Spectrograms and pitch tracks for the sentence *Elle va prendre son train* ("she is going to take her train"), uttered with neutral (top), happy (middle) and angry (bottom) prosody.



2.1.3. Design

Three counterbalanced lists were constructed, so that each target word was preceded by the three types of emotional primes, and participants were presented with each target word only once. Each list included the 45 negative target words. Among them, 15 were preceded by the angry prosody, 15 by the happy prosody and the 15 others by the neutral prosody. As we have constructed five semantically neutral sentences

(e.g. “She is going to take her train”), each sentence within the lists were repeated three times in each prosodic context.
For the purpose of the lexical decision task, we added 45 target pseudo-words to each sub-list, which were created by changing one phoneme of a real word not used in the experimental lists. Pseudo-words were matched in number of phonemes and number of syllables to the target words. Exactly the same sentence primes as for the target words were used for the pseudo words, and as for the target words, 15 were preceded by the angry prosody, 15 by the happy prosody and the 15 others by the neutral prosody.

2.1.4. Procedure

Participants were tested individually in a quiet room at Laboratoire Parole et Langage. They listened to 90 trials, each of them consisting in an auditory prime followed by an auditory target. Primes and targets were played through professional headphones. At 50 ms after the offset of the prime, the visual target (a word or a pseudo-word) appeared in the middle of the screen until participants gave their response. Participants were asked to decide as quickly and accurately as possible whether the target was a word or a non-word. “Word” responses were made with the dominant hand. Each participant began the experiment with a block of 12 practice trials. The experiment lasted about 15 minutes. The experiment was administered via *E-Prime* 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, USA).

2.2. Participants

Forty-five French native speakers were recruited for the experiment. All participants were women as they are supposed to be more sensitive to emotional prosody [7]. None of them participated in the previous validation tasks. Participants were spread across the three sub-lists counterbalancing the happy, angry and neutral auditory primes, with each participant listening to all target words. Listeners signed an informed consent form before the experiment.

2.3. Results

All analyses were performed on reaction times (RTs) on critical trials, i.e., trials in which lexical decisions concerned real words (2025 trials, i.e., 45 trials x 45 participants). Among those trials, 63 trials (3.1 %) were discarded, either because they were responded incorrectly or because the reaction times were considered as outliers (> 2.5 SD from the mean value). Thus, the final dataset included 1961 trials. Statistical analyses included linear mixed models on (logarithmically transformed) RTs. Log-transformation was applied to achieve normal distribution. The models tested the effect of PRIMING (neutral/matching/mismatching). PARTICIPANTS and WORDS were included as random intercepts. We started the analysis by fitting each model with the two intercepts, and by including by-participant random slopes for PRIME. After backward elimination based on likelihood-ratio tests [12], the final models included only the two intercepts. The reference level was set to the level “neutral” for PRIMING. For pairwise comparisons involving the conditions matching vs. unmatched, the reference level was changed and only p-values smaller than 0.025 were treated as significant (p = 0.05/2).

RTs were very similar across the three PRIMING conditions (matching=950.6 ms; mismatching=951.9 ms, neutral=958.01 ms) with no significant differences between the three priming conditions (p > 0.025). We made an additional analysis, to test whether RTs across priming conditions were modulated by the association strength of the word with the emotion “anger” [see 9]. A linear mixed model testing the effects of PRIMING (neutral/matching/mismatching) and ASSOCIATION STRENGTH (low/high) revealed no significant effects (Figure 1).

3. Experiment 2: Preliminary study

We employed an auditory-visual paradigm to test the effect of the emotional prosodic context in absence of overt prosody on the target words.

3.1. Materials and Methods

We used the same material (primes and targets) as in Exp. 1, with two main differences: (i) target words were visually presented on a screen; (ii) as usual in cross-modal priming studies, the visual target words were presented at the offset of the auditory primes.

3.2. Participants

Fifteen French native speakers were recruited for the experiment. All participants were women. None of them participated in the previous experiment. For each word category, participants were spread across the three sub-lists counterbalancing the happy, angry and neutral auditory primes. Listeners signed an informed consent form before the experiment.

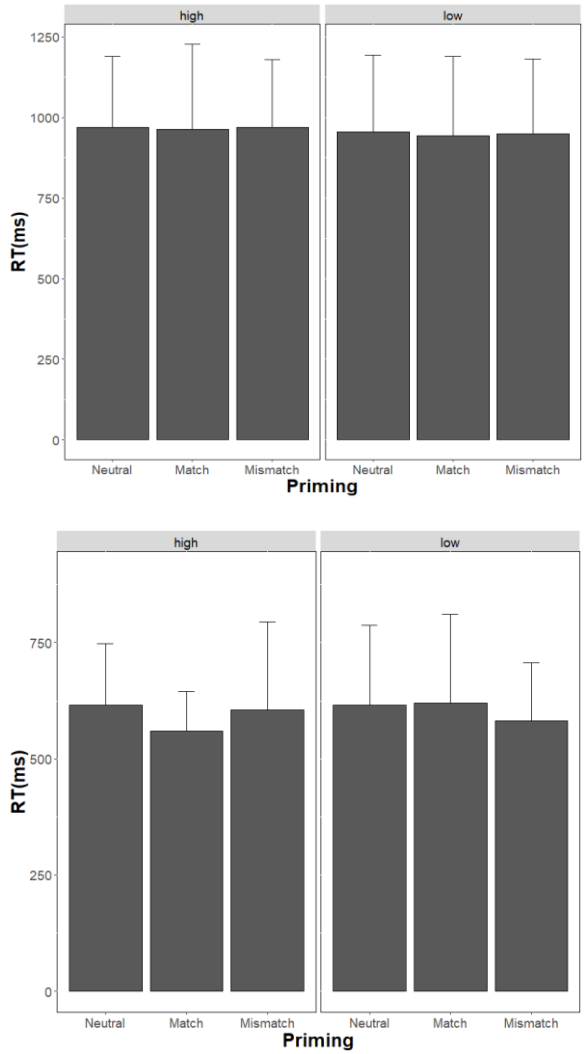
3.3. Results

All analyses were performed on reaction times (RT) on critical trials (675 trials, i.e., 45 trials x 15 participants). Among those trials, 28 (4.1%) were discarded from the analyses, either because they were responded incorrectly or because reaction times were outliers. RTs were very similar in the matching (597.3 ms) and mismatching (594.4 ms) and neutral condition (606.4 ms). The statistical models showed that there were no significant difference between matching and mismatching conditions [negative: β = 0.005, SE = 0.02, t = 0.24, p =0.8]. Similarly, the neutral condition was not significantly different both from the matching and the mismatching conditions (p > 0.025).

Like for Exp. 1, an additional analysis was run to evaluate the impact of association strength. For target words which were more strongly associated to anger (Figure 2), RTs were longer in the mismatching (604.8 ms) than in the matching condition (560.05 ms). RTs in the neutral condition had similar values as the mismatching condition (615.2 ms). The difference between mismatching and matching conditions was significant [β = 0.05, SE = 0.02, t = 2.27, p =0.023], and the difference between the neutral and the matching conditions approached the significance [β = -0.0047, SE = 0.002, t = -1.83, p =0.06]. The difference between the neutral and the mismatching conditions was not significant [β = 0.001, SE = 0.002, t = 0.41, p =0.67]. For words with low association strength, the effects of PRIMING, ASSOCIATION STRENGTH and their interactions were not significant.

Figure 2. Barplots for RTs (ms) for the lexical decision task, for Experiment 1 (top) and Experiment 2 (bottom).

Data are split by priming conditions and association strength.



4. Discussion

Our results suggest that prosodic context may modulate word recognition, but only under specific circumstances. In particular, differences in reaction times between the matching and the mismatching conditions appeared only in a cross-modal priming situation, and for words which were more strongly associated with anger.

No effects of prosodic context were found in Experiment 1 when the targets were auditorily presented. A possible explanation is that the pronunciation of words with emotional connotation in a neutral voice may seem somewhat surprising for participants. The surprise generated by the neutral prosody for emotional words could have masked the impact of the preceding prosodic context on spoken word recognition. Evidence for this interpretation comes from the literature, showing that incongruencies between prosodic properties and semantic content have a negative impact on spoken word recognition In particular, speakers recognize or repeat emotional words slower when they are produced with neutral or incongruent prosody than with congruent prosody [14].

Because of this possible bias in Experiment 1, a preliminary study (Experiment 2) was run with visual rather than auditory target words. Similarly to [9], we found some evidence that affective priming is modulated by the association strength with an emotion, i.e., it is more likely to emerge for words that are the more representative of an emotional category. It is thus appeared that RTs to negative target words conveying “anger” were faster when these words are preceded by sentences uttered with angry prosody than when they were preceded by sentences uttered with happy prosody. In contrast, RTs to negative target words with low association strength to “anger” were unaffected by the prosody of sentence primes. Although our preliminary findings must be interpreted with cautions due to the smaller group of participants, and because they are based on post-hoc analysis only, they replicate results by [9] on American English in respect to the comparison between a match and a mismatch in the emotion conveyed by the primes and the targets. Perhaps more importantly, our preliminary finding extend those of [7,8] to semantically neutral sentence primes, thus suggesting that prosody in itself is sufficient to provides an emotional context influencing word recognition. One might also argue that our difficulty to find a robust effect across our two experiments is due to the fact that contrary to [15] we used intermixed lists of stimuli in which listeners were exposed to the three types of prosody. Note however, that affective prosodic priming effect was also observed in [9] with intermixed lists of stimuli. Given this discrepancy, further study should be done to provide a more in depth examination of the impact of list composition on the magnitude of the affective priming effect. Although our finding need to be replicated with a more powered experiment, they can bring some information that help us to tease apart between an account of affective priming effects in terms of spreading activation or in terms of compatibility judgment between the primes and the targets [3, 5, 16]. The lack of a significant difference between the mismatching and the neutral priming conditions argue against a congruency checking mechanism between the emotion conveyed by the prime and target stimuli, since in this case an interference effect should have been observed. As a result, the 55 ms difference, near the significance (p=.06), between the neutral and matching priming condition for words with high association strength with anger is likely due to spreading activation mechanism. Through this mechanism, the anger conveyed by the prosodic context generates activation to words associated with this emotional category, thus facilitating the subsequent processing of that word in comparison to both the neutral and mismatching priming conditions.

5. Conclusions

Contrary to studies in other domains [3], this work suggests that the congruency in emotional valence between the prosodic primes and the target words is not sufficient to trigger affective priming. Affective prosodic priming is in fact possible only when the prime and the word share the same emotion category. It appeared thus that listeners automatically extract the emotion conveyed by spoken words/sentences uttered with emotional prosody, which in turn generate activation on words that are associated with the extracted emotion. Consequently, just as semantic context facilitates the processing of semantically associated words, emotional context facilitates the processing of emotionally associated words.

6. References

[1] H. Neely (1991), "Semantic priming effects in visual word recognition: A selective review of current findings and theories." In D. Besner and G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*, pp. 264–336. Lawrence Erlbaum Associates, Inc..

[2] D. A. Swinney (1979). "Lexical access during sentence comprehension: (Re)consideration of context effects," *Journal of Verbal Learning & Verbal Behavior*, 18(6), 645–659.

[3] L. Ferrand, F. Ric, and M. Augustinova (2006). "Quand "amour" amorce "soleil" (ou pourquoi l'amorcage affectif n'est pas un (simple) cas d'amorçage sémantique?)," *L'Année Psychologique*, 106 (1), pp. 79–104.

[4] D. Hermans, D. Smeesters, J. De Houwer, and P. Eelen (2002), "Affective priming for associatively unrelated primes and targets," *Psychologica Belgica.* 42. 10.5334/pb.994.

[5] R. H. Fazio, D. M. Sanbonmatsu, M. C. Powell, and F. R. Kardes (1986), "On the automatic activation of attitudes," *Journal of Personality and Social Psychology*, 50(2), pp. 229–238

[6] Wentura, D. (2000). "Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgment tendencies," *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 456–469.

[7] A. Schirmer, S. A. Kotz, A. D. Friederici (2002), "Sex differentiates the role of emotional prosody during word processing," *Cognitive Brain Research*, 14(2), pp. 228–233.

[8] A. Schirmer, S.A. Kotz, A.D. Friederici (2005), "On the role of attention for the processing of emotions in speech: sex differences revisited," *Cognitive Brain Research*, 24(3), pp. 442-52.

[9] S. K. Kim and M. Sumner (2017). "Beyond lexical meaning: The effect of emotional prosody on spoken word recognition," *The Journal of the Acoustical Society of America* 142, EL49--EL55.

[10] B. New, C. Pallier, M. Brysbaert, M. and L. Ferrand (2004). "*Lexique 2* : A new French lexical database," *Behavior Research Methods, Instruments, & Computers* 36**,** 516–524.

[11] K. R. Scherer (2003), "Vocal communication of emotion: A review of research paradigms," *Speech Communication*, 40(1-2), 227–256.

[12] J. C. Pinheiro, D. M. Bates (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer. Chicago

[13] L. C. Nygaard, and E. R. Lunders, (2002), "Resolution of lexical ambiguity by emotional tone of voice", *Memory and Cognition*, 30, 583–593.

[14] L. C. Nygaard, and J. S. Queen (2008). "Communicating emotion: Linking affective prosody and word meaning," *J. Exp. Psychol. Hum. Percept. Perf.* 34(4), 1017–1030.

[15] Wurm, L. H., Vakoch, D. A., Strasser, M. R., Calin-Jageman, R., and Ross, S. E. (2001). "Speech perception and vocal expression of emotion," Cognit. Emotion 15(6), 831–852.

[16] J. De Houwer, D. Hermans, K. Rothermund, & D. Wentura (2002). "Affective priming of semantic categorisation responses," *Cognition and Emotion*, 16(5), 643–666.