Effects of Trait Anxiety on Semantic and Prosodic Processing

Simon Busch-Moreno, Jyrki Tuomainen, David Vinson
Division of Psychology and Language Sciences, University College London
simon.busch.15@ucl.ac.uk

ABSTRACT

Little is known about how individual differences affect speech processing. A number of similarities in hemispheric lateralization patterns suggest close association between acoustic features (prosody), emotional language (threat) and intrinsic affect (anxiety). To address their interaction, we carried out two dichotic listening experiments using semi-naturalistic sentences containing threatening prosody and semantics, and we took participants' worry-levels as a measure of trait anxiety. Results suggest that anxiety does affect speech processing, and this differs depending on task demands and acoustic features.

Keywords: prosody, semantics, affect, anxiety

1. INTRODUCTION

While prosodic information relies on supra-segmental variation of intensity, pitch, voice quality and duration, semantic information relies on morphemes composed by varying combinations of phonological segments [1]. Prosodic and semantic information can develop together in a natural utterance (e.g. emotional sentences), and can convey emotional information simultaneously [2, 3]. To our knowledge, whether intrinsic affect differences between individuals (e.g. variation in trait anxiety) influences the processing of semantics and prosody in a different way remains an unexplored problem.

The present study aims to understand the effect of trait anxiety on these information properties of speech. We use dichotic listening (DL) which provides a robust test of functional hemispheric lateralization [4], tapping into features of both speech (language) and anxiety (affect) processing.

On the language side, evidence suggests the left lateralization of segmental aspects of speech and right lateralization of suprasegmental features of speech [5, 6]. On the affect side, arousal and fear responses associated with escape have been observed to be right lateralized, while responses associated to worry and environmental evaluation follow the opposite pattern [7, 8, 9].

A few dichotic listening experiments have researched the effects of anxiety on emotional speech processing. They either use speech/prosody as an emotion-eliciting stimulus, or use DL mainly as an attentional manipulation technique [10, 11, 12, 13]. As a result, they are limited in the extent to which they reveal the relationship between dynamic variation in emotional language processing (prosody/semantics). Studies with more of a focus upon dynamic properties of emotional language, instead, do not tend to consider individual differences [e.g. 14, 15]. To address these issues, we designed two web- based DL experiments, using semi-naturalistic sentences in order to ensure dynamic language processing beyond the single word level. Participants were asked to discriminate between neutral and threatening sentences (expressing threat via semantics, prosody or both), in the direct-threat condition, identifying whether threatening stimulus occurred on left or right, and indirect-threat, identifying whether neutral stimulus occurred on left or right.

As early over-attention to threat [16] might affect earlier prosody/semantic lateralization patterns [17], and later over-engagement with threat [16] might affect later emotional language evaluation stages [17], we manipulated timing. Experiment-1 required a delayed response: after sentences' offset; Experiment-2 required a fast response: during sentence presentation. This allows us to differentiate responses made at late evaluative stages (delayed response) vs. responses made at earlier attentive stages (fast response).

For Experiment-1 we hypothesize that anxious over-engagement with threat at mid-late evaluative stages [16] should increase left hemisphere (LH) engagement, disturbing possible LH to right hemisphere (RH) information transferring [14, 17]. Hence, we predict that a left ear advantage (LEA), characteristic DL response to prosody/emotional stimuli [14, 18], should decrease as a function of anxiety, especially for semantic threat. This implies slower and less accurate responses for anxious people at their left ear when attending semantic stimuli.

For Experiment-2 we expect that, as responses are forced to be faster (online), prosody effects should be higher as they match early-mid emotional processing stages [17]. Therefore, we hypothesize that higher anxiety should reduce LH involvement [9] due to over-attention to threat effects, which co-occurs at these earlier-mid processing stages [9, 16]. Hence, we predict an enhanced LEA for highly anxious participants, especially for prosodic threat. Thus, faster and more accurate responses for anxious people at their left ear when attending prosodic stimuli.

2. METHODS

2.1 Participants

Participants (20-40 years old) were recruited using Prolific (prolific.ac). For experiment-1, 44 participants were included (mean age = 31.7, 27 females). For Experiment-2, 52 participants were included (mean age = 31, 24 females). Only participants reporting being right-handed, having English as first language, without hearing and neurological/psychiatric disorders, and using only a desktop or laptop to answer the experiment were recruited. All participants signed consent to participate under European Data Protection Act (1998) regulations, and were remunerated for their participation.

2.2 Materials

Four types of sentences were recorded: Prosody-only (neutral-semantics and threatening-prosody), Semantic-only (neutral-semantics and threatening-prosody), Congruent (threatening-semantics and threatening-prosody), and Neutral (neutral-semantics and neutral-prosody). Threatening sentences were extracted from movie scripts by matching them with a list of normed threatening words from the extended Affective Norms for English Words (ANEW) [20]. Sentences were recorded in an acoustically isolated chamber using a RODE NT1-A1 microphone by a male English speaker. The speaker was instructed to speak in what he considered his own threatening/angry or neutral voice for recording Prosody-only/Congruent and Semantic-only/Neutral sentences respectively.

Sentences' prosodic bio-informational dimensions [21] were extracted using ProsodyPro [22] in Praat (praat.org). Median Pitch (F0) comparisons, crucial for defining angry or threatening voices [23], were performed by using Tukey HSD tests in R (R-project.org). These showed no difference for Semantic-only vs. Neutral ($p = 0.31$) and Congruent vs, Prosody-only ($p = 0.93$) comparisons. All other comparisons, involving threatening prosody vs. neutral prosody, showed a significant median F0 difference (all p-values < 0.01). Higher median F0 for threatening stimuli aligns them with hot-anger (rage), which has higher F0 than cold-anger or neutral prosody [23, 24]. Previous DL studies' mean F0 values are consistent with this [14, 25].

2.3 Procedure

Before starting the experiments, participants answered the Penn State Worry Questionnaire (PSWQ) [19] to assess their worry-level. In Experiment-1, participants heard 52 sentences per threatening type (Prosody-only, Semantic-only, Congruent), all dichotically paired with a Neutral sentence of comparable duration. In one half of the study they were instructed to indicate at which ear they heard the threatening sentence by pressing the right or left arrow keys. In the other half of the study they were instructed to respond to the Neutral sentence (indirect-threat condition). This was intended to address attention effects [4, 26]. Starting ear (left or right) and starting condition (direct- or indirect-threat) were counterbalanced. Participants were told to answer only when the sentence finished playing and a bulls-eye (target) image appeared on the screen. A 1400ms inter-stimulus-interval (ISI) was used, the target image stayed on the screen during this period. The same procedure was used for Experiment-2, but participants were instructed to answer only before sentence's end. Experiments were implemented using Gorilla: gorilla.sc.

2.4 Analysis

Data from both experiments was analysed using R (R-project.org). A linear mixed-effects model with reaction time (RT) as a dependent variable; Worry, Type, Ear, and Threat-direction as interaction terms; and subjects and sentences as random effects was selected from a model comparison. A TypeII Satterthwaite's ANOVA was used on this model. For accuracy measures, a generalized mixed-effects binomial linear model was selected from a model comparison, with percentage of correct responses (PC) as dependent variable, same interaction terms as in the previous model, and subjects as random effects (sentences were excluded due to ceiling effects). A TypeII analysis of deviance was used on this model.

3. RESULTS

Experiment-1's RT results, computed from the best-fit model ($R^2 = 0.22$), showed the following relevant effects. An Ear X Worry X Threat-direction interaction ($F_{(1, 8182.9)} = 14.56$, $p < 0.01$), suggesting that ear differences for RTs were mainly restricted to worry-level and threat-direction. Also, a Worry X Type interaction ($F_{(2, 8189.1)} = 9.44$, $p < 0.01$) was observed.

Low-worry follow-up comparisons indicate significant differences for Prosody-only vs. Semantic-only (β = -40.72, SEM = 12.77, $p < 0.01$) and for Semantic-only vs. Congruent (β = 38.47, SEM = 12.44, p = 0.01). High-worry comparisons indicate significant differences for Prosody-only vs. Congruent (β = 86.93, SEM = 15.69, $p < 0.01$) and for Semantic-only vs. Congruent (β = 58.13, SEM = 15.19, $p < 0.01$). The other two comparisons did not show significant differences ($p > 0.05$). These patterns suggest that lower worriers respond faster to conditions containing threatening prosody (Prosody-only and Congruent), but higher worriers respond faster to the Congruent condition only.

Experiment-1's accuracy results, computed from the best fit model ($R^2 = 0.14$), showed the following relevant effects. An Ear X Worry X Threat-direction interaction ($\chi^2 = 13.06$, $p < 0.01$), which suggest that worry effects on ear advantages are driven by Threat-direction differences. Also, a Worry X Type interaction ($\chi^2 = 16.36$, $p < 0.01$) was observed.

Pairwise comparisons for Worry X Type showed significant differences ($p < 0.01$) for all comparisons excepting Prosody-only vs. Semantic-only at low-worry (β = -0.19, SEM = 0.11, p = 0.34). Comparisons for Ear X Type showed a significant difference for the low-worry left-ear vs. right-ear comparison (β = 0.78, SEM = 0.11, $p < 0.01$), but showed no difference for the same comparison at high-worry (β = -0.01, SEM = 0.14, p = 0.88). This suggests that participants were more accurate at recognizing Congruent stimuli, but high-worriers were much less accurate exclusively for Prosody-only (see Figure 1).

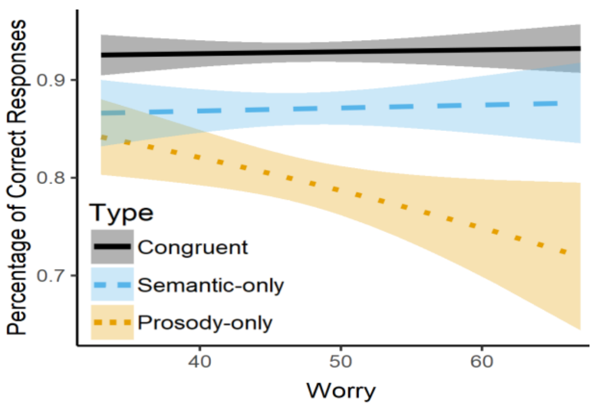

Figure 1: Experiment-1 Worry X Type interaction. Shadows indicate SEMs.

Experiment-2's RT results, computed from the best fit model ($R^2 = 0.37$), showed an Ear X Worry X Type X Threat-direction interaction ($F_{(2, 9691.6)} = 3.43$, p = 0.03). Follow-up comparisons showed significant differences only for Direct-threat: low-worry vs. high-worry at the right ear for Prosody-only (β = 264, SEM = 92.51, $p < 0.01$), and low-worry vs. high-worry for Semantic-only at the left ear (β = 194.37, SEM = 92.44, p = 0.03) and right ear (β = 194.26, SEM = 92.34, p = 0.03). This suggests that higher-worriers were generally faster when answering to Prosody-only stimuli presented at their right ear, and also faster when answering to Semantic-only at both ears (when compared to Prosody-only).

Experiment -2's accuracy analysis, computed from the best fit model ($R^2 = 0.07$), showed several interactions. An Ear X Worry X Type interaction ($\chi^2 = 7.78$, p = 0.02) and an Ear X Type X Threat-direction interaction ($\chi^2 = 25.7$, $p < 0.01$) are the most relevant.

Follow-up comparisons for Ear X Worry X Type show a significant difference for high-worry for stimuli presented at the right ear for Prosody-only (β = -0.9, SEM = 0.16, $p < 0.01$) and Semantic-only (β = -1.2, SEM = 0.16, $p < 0.01$) when compared to Congruent stimuli. At the left ear all differences were significant ($p < 0.01$), excepting low-worry for Prosody-only vs. Congruent (β = -0.5, SEM = 0.2, p = 0.11) and high-worry for Prosody-only vs. Semantic-only (β = 0.03, SEM = 0.14, p = 0.9). When comparing right vs left ear, the only significant difference was for low-worry and Semantic-only (β = -0.62, SEM = 0.18, $p < 0.01$).

This, as illustrated in Figure 2, indicates that low-worriers were more accurate for all stimuli types at their right ear, while at their left ear they were more accurate for Prosody-only and Congruent stimuli. When comparing left vs. right ears, only low-worriers were more accurate for their right ear, and only for Semantic-only stimuli. Caution is required due to model's low effect size.

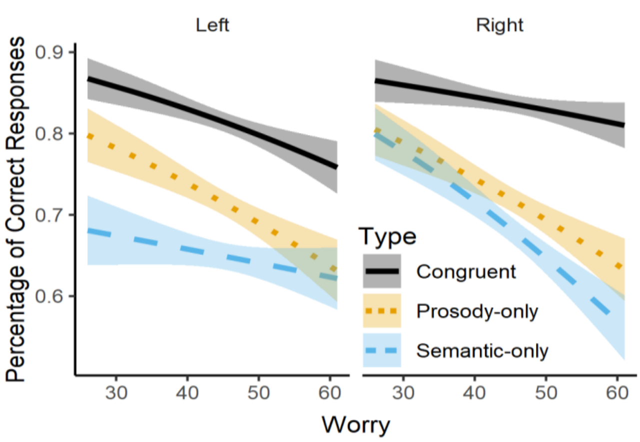

Figure 2: Experiment-2 Worry X Ear X Type interaction. Shadows indicate SEMs.

4. DISCUSSION

Overall, present study's hypotheses were not thoroughly confirmed due to some of the fine-grained patterns of differences we observed. Unexpectedly, for Experiment-1, lower RTs were associated with higher-worry, and lower accuracy with Prosody-only stimuli overall. However, the Ear X Worry X Threat-direction interaction was partially consistent with our predictions, in that direct-threat induced an accuracy LEA for both high and low worriers, which was reduced for high worriers in indirect-threat.

For Experiment-2, also unexpectedly, high-worriers reacted faster than low-worriers for Semantic-only stimuli in the direct-threat condition. High-worriers also answered less accurately than low-worriers for Prosody-only stimuli. Although we had predicted an enhanced LEA for high-worriers in the prosody-only condition, instead we found that only low-worriers showed a REA for Semantic-only stimuli. Despite these unexpected patterns, the Threat-direction X Ear X Type interaction is consistent with our predictions, in particular a prosody LEA for direct-threat regardless of anxiety level.

Although these patterns are somewhat inconsistent with our general predictions, they still suggest that online vs. delayed responses differ for prosody and semantics. In addition, they indicate that worry-level (trait anxiety) can affect processing of the informational features of speech, consistent with multistep models of emotional language processing [17], and multistage models of anxiety [16].

Nevertheless, explaining these unexpected patterns requires some reinterpretation of how anxious laterality patterns interact with language laterality patterns. First, present results indicate that in Experiment-1, where participants must delay responding, laterality effects are mainly associated to threat-direction [11, 12], rather than directly associated to the informational features of speech.

However, it is possible that delayed evaluative responses by high-anxious participants, with increased activity at LH [13], tend to interfere more with prosody. Considering that all participants were right handed, responses to threatening prosody might require RH to LH transferring from mid to late stages before response, a mechanism that could explain simultaneously slower RTs and lower accuracy. This is consistent with the callosal relay mechanism proposed by dynamic models of dichotic listening [8, 27]. It is also consistent with LH to RH transfer as proposed by multistep models [17]; which could explain the better performance for Semantic-only stimuli. This may also be associated with an additional decision-making/goal-engagement stage discussed in the anxiety literature [9, 16], not discussed in language processing models, but compatible as a stage following or derived from evaluative processing.

For Experiment-2, in which participants were under time pressure to respond (online responses), high-anxious participants were generally faster but less accurate for all stimuli but Congruent sentences. This pattern of results is consistent with a double mechanism of apprehension and arousal for trait anxiety associated to high worry [9, 28]. From a perspective of multistep models this would imply over-attention to threat at early stages [16, 17], inducing a right lateralized pattern [9], and over-engagement with threat at later stages [16, 17], inducing a shift to a left lateralized pattern [9].

This would explain why high-worriers tend to be less accurate for both Prosody-only and Semantic-only stimuli. As they might be prone to over-attend early and easily recognisable Prosody-only stimuli, inducing higher rate of false alarms; or to over-engage with harder to recognize Semantic-only stimuli, inducing a higher rate of misses. However, better temporal resolution is required to confirm this interpretation, including electroencephalographic (EEG) measures of experiments with or without dichotic listening.

Regarding low-worriers, however, they did show a diminished LEA for Semantic-only stimuli. This might simply confirm a dynamic processing of dichotic listening involving callosal transfer [14], as Semantic-only stimuli might require LH assessment before emotional evaluation [17]. This effect might not be present in high-worriers, as they might engage faster into an emotional evaluation stage [29], which is supported by their faster responses.

Previous evidence suggests a right lateralized pattern for prosody vs. semantic evaluation in an EEG experiment (not considering anxiety), using a congruency but non-dichotic task [15, but see: 30]. Although this pattern is explained by the strong association between pitch recognition and RH engagement [15, 6], there are other frequency and spectral features [1, 6, 21, 23, 24], that might be important for recognizing both threatening and neutral sentences. Further research is required to know whether these mainly engage RH processing in a way that can be influenced by trait anxiety.

Taken together present experiments show that high-worry (trait anxiety) can affect the processing of semantic and prosodic threatening stimuli. In addition, evaluation time seems to play a relevant role in threatening speech recognition. However, results indicate that current predictions need to be reconsidered and further research, including the use of electrophysiological measures, is required to fully understand this phenomenon.

5. REFERENCES

[1] Liu, F., Xu., Y., Prom-on, S., Yu, A.; (2013) Morpheme-like prosodic functions: Evidence from acoustic analysis and computational modelling. *Journal of Speech Sciences,* 3 (1) pp. 85-140.

[2] Nygaard, L., Herold, D., Namy, L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. *Cognitive Science*, *33*(1), 127-146.

[3] Schirmer, A., Kotz, S. (2003). ERP Evidence for a Sex-Specific Stroop Effect in Emotional Speech. *Journal Of Cog. Neuro.*, *15*(8), 1135-1148.

[4] Hugdahl, K. (2011). Fifty years of dichotic listening research – Still going and going and…. *Brain And Cognition*, *76*(2), 211-213.

[5] Hickok, G., Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.

[6] Zatorre, R., Belin, P., Penhune, V. (2002). Structure and function of auditory cortex: music and speech. *Trends In Cognitive Sciences*, *6*(1), 37-46.

[7] Heller, W., Nitschke, J., Etienne, M., Miller, G. (1997). Patterns of regional brain activity differentiate types of anxiety. *Journal Of Abnormal Psychology*, *106*(3), 376-385.

[8] Nitschke, J., Heller, W., Palmieri, P., Miller, G. (1999). Contrasting patterns of brain activity in anxious apprehension and anxious arousal. *Psychophysiology*, *36*(5), 628-637.

[9] Spielberg, J., De Leon, A., Bredemeier, K., Heller, W., Engels, A., Warren, S. et al. (2013). Anxiety type modulates immediate versus delayed engagement of attention-related brain regions. *Brain And Behavior*, 3(5), 532-551.

[10] Bruder, G., Wexler, B., Stewart, J., Price, L., Quitkin, F. (1999). Perceptual asymmetry differences between major depression with or without a comorbid anxiety disorder: A dichotic listening study. *Journal Of Abnormal Psychology*, *108*(2), 233-239.

[11] Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M., Scherer, K., Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage*, 28(4), pp.848-858.

[12] Peschard, V., Gilboa-Schechtman, E., Philippot, P. (2016). Selective attention to emotional prosody in social anxiety: a dichotic listening study. *Cognition and Emotion*, 31(8), pp.1749-1756.

[13] Leshem, R. (2018). Trait Anxiety and Attention: Cognitive Functioning as a Function of Attentional Demands. *Current Psychology*.

[14] Grimshaw, G., Kwasny, K., Covell, E., Johnson, R. (2003). The dynamic nature of language lateralization: effects of lexical and prosodic factors. *Neuropsychologia*, *41*(8), 1008-1019.

[15] Kotz, S., Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*, 1151, pp.107-118.

[16] Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M., van IJzendoorn, M. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, 133(1), pp.1-24.

[17] Kotz, S., Paulmann, S. (2011). Emotion, Language, and the Brain. *Language And Linguistics Compass*, *5*(3), 108-125.

[18] Godfrey, H., Grimshaw, G. (2015). Emotional language is all right: Emotional prosody reduces hemispheric asymmetry for linguistic processing. *Laterality: Asymmetries Of Body, Brain And Cognition*, *21*(4-6), 568-584.

[19] Meyer, T., Miller, M., Metzger, R., Borkovec, T. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research And Therapy*, *28*(6), 487-495.

[20] Warriner, A., Kuperman, V., Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.

[21] Xu, Y., Kelly, A., Smillie, C. (2013). Emotional expressions as communicative signals. In S. Hancil, D. Hirst, *Prosody and Iconicity* (pp. 33-60). John Benjamins Publishing Co.

[22] Xu, Y. (2013). ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis (pp. 7-10). Aix-en-Provence, France: TRASP.

[23] Banse, R., Scherer, K. (1996). Acoustic profiles in vocal emotion expression. *Journal Of Personality And Social Psychology*, *70*(3), 614-636.

[24] Hammerschmidt, K., Jürgens, U. (2007). Acoustical Correlates of Affective Prosody. *Journal Of Voice*, *21*(5), 531-540.

[25] Grimshaw, G., Séguin, J., Godfrey, H. (2009). Once more with feeling: The effects of emotional prosody on hemispheric specialisation for linguistic processing. *Journal Of Neuroling.*, *22*(4), 313-326.

[26] Aue, T., Cuny, C., Sander, D., Grandjean, D. (2011). Peripheral responses to attended and unattended angry prosody: A dichotic listening paradigm. *Psychophysiology*, 48(3), pp.385-392.

[27] Atchley, R., Grimshaw, G., Schuster, J., Gibson, L. (2011). Examining lateralized lexical ambiguity processing using dichotic and cross-modal tasks. *Neuropsychologia*, *49*(5), 1044-1051.

[28] McNaughton, N., Gray, J. (2000). Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety. *Journal Of Affective Disorders*, *61*(3), 161-176.

[29] Eldar, S., Yankelevitch, R., Lamy, D., Bar-Haim, Y. (2010). Enhanced neural reactivity and selective attention to threat in anxiety. *Biological Psychology*, *85*(2), 252-257.

[30] Paulmann, S., Jessen, S., Kotz, S. (2012). It's special the way you say it: An ERP investigation on the temporal dynamics of two types of prosody. *Neuropsychologia*, *50*(7), 1609-1620. doi: 10.1016/j.neuropsychologia.2012.03.014.