Semantic and Prosodic Threat Processing in Trait Anxiety: Is Repetitive Thinking Influencing Responses?

Simon Busch-Moreno, Jyrki Tuomainen, David Vinson*
Division of Psychology and Language Sciences, University College London

Abstract

The present study attempts to identify how trait anxiety, measured as worry-level, affects the processing of threatening speech Two experiments using dichotic listening tasks were implemented; where participants had to identify sentences that convey threat through three different information channels: prosody-only, semantic-only and both semantic and prosody (congruent threat). We expected different ear advantages (left or right) depending on task demands, information type, and worry level. We used a full Bayesian approach for statistical modelling and analysis. Results indicate that when participants made delayed responses (Experiment 1), reaction times increased with worry level, but under time pressure (Experiment 2) worry level only affected reaction times for semantically neutral, prosodically threatening stimuli. We explain this effect by proposing a fourth stage, associated with goal oriented deliberation, for a three-phasic multistep model of emotional language processing. Higher levels of trait anxiety could induce verbal repetitive thinking (i.e. worry and/or rumination), which might prolong the mentioned deliberation stages, thus slowing-down responses.

Keywords: prosody, semantics, speech, threat, anxiety

*Corresponding author: d.vinson@ucl.ac.uk, room 406, 26 Bedford Way, London, WC1H 0AP.

Introduction

Humans can convey emotion information through different channels, and in the particular case of language the manipulation of tones and/or meanings (i.e. prosody and semantics) are common ways to do so. While prosodic information relies on suprasegmental variation of intensity, pitch, voice quality and duration, semantic information relies on morphemes (minimal meaningful language units) composed by varying combinations of phonological segments (Liu et al., 2013). Prosodic and semantic information can develop together in a complex language emission (e.g. emotional sentences), and can convey emotional information simultaneously (Nygaard et al., 2009; Schirmer and Kotz, 2003). To our knowledge, whether intrinsic affect differences between individuals (e.g. variation in trait anxiety) has differentiable effects on prosody and semantics remains an unexplored problem. Investigating this possible connection can bring to light possible idiosyncratic effects of anxiety on language, moving forward the understanding of individual differences in language processing but also refining understanding of speech information properties.

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 (Hugdahl, 2011), tapping into features of both speech (language) and anxiety (affect) processing. DL can provide a behavioural test of laterality in such a way that information- and affect-related aspects of processing can be disentangled. Normally, responses to DL tasks that do not involve prosody or emotion indicate advantaged response times and/or accuracy for language processing at stimuli presented at the right ear (REA) (Hugdahl, 2011). Differently, DL responses to emotional and/or prosodic stimuli show either diminished REAs or a left ear advantage (LEA) (Godfrey and Grimshaw, 2015; Grimshaw et al, 2003).

The idea of exploring laterality in this way is based on previous theoretical models and supporting evidence indicating that brain hemispheres have different processing functions for both speech's information features and intrinsic affect. On the language side, evidence suggests the left lateralization of segmental aspects of speech and right lateralization of suprasegmental features of speech (Poeppel, 2003; Poeppel et al., 2007; Zatorre, 2001; Zatorre et al., 2002). On the affect side, arousal and fear responses associated with escape have been observed to be right lateralized, while responses associated with worry and environmental evaluation follow the opposite pattern (Heller et al., 1997; Nitschke et al., 1999; Spielberg et al., 2013). As mentioned before, emotions can be simultaneously conveyed through two different information

channels, which can induce different lateralization patters. In addition, these patterns can be affected by intrinsic affect (i.e. anxiety) lateralization patterns. This motivates the question: what is the difference between semantic and prosodic comprehension in a natural emotional expression as processed by anxious people? Before answering this question, we need to find out the points of connection between speech, emotional language and anxiety processing, if any.

Prosody and Semantics: Informational Lateralization

Neuroscientific research has observed that information conveyed through prosody or semantics/syntax is processed via differently lateralized brain routes (Belin et al, 2004). Other findings indicate that this difference might be due mainly to the emotional content of language stimuli (Liebenthal et al., 2005). These differences, however, may not be exclusive. Indeed, if emotional language lateralization is considered as a phasic process, then differences in lateralization might change at any point of the processing time-course (Schirmer and Kotz, 2006). Hence, some of these differences might be related to informational processing and others to the processing of affect/cognition. Therefore, hemisphericity patterns might be due to both emotional and speech processing, but a particular observed left, right or bilateral orientation might be evident depending on the observed time phase. One model addressing this issue is the multistep model of emotional language, which proposes three main processing stages: early stage perceptual processing, mid stage recognition processing, late stage evaluation processing (Kotz and Paulmann, 2011).

Under this model, early stages involve the processing of acoustic properties (purely acoustic information), where greater right hemisphere (RH) engagement would be associated to prosodic processing and left hemisphere engagement (LH) would be associated to phonological processing (Poeppel, 2003; Zatorre, 2001). This leads to the interpretation that LH might process lexical (phonologically composed words) information better, while RH privileges suprasegmental (prosody) information. Mid stages might involve the emotional recognition of stimuli (e.g. integration or previously processed information), implying greater involvement of RH or LH depending on stimulus type and/or conveyed emotion (Schirmer and Kotz, 2006). Late stages would be associated to informational integration and evaluation of emotional stimuli (Kotz and Paulmann, 2011). Another crucial aspect of the model is that it also considers information transferring between hemispheres (Kotz and Paulmann, 2011). Mechanism of callosal relay have been proposed as important aspects for RH to LH (and vice versa) communication of prosodic and syntactic information (Friederici et al, 2007), and also interhemispheric communication of emotional prosody processing (Ross et al, 1997). In

addition, callosal relay mechanisms have been proposed as an explanation for different effects of emotional semantics and prosody processing in a dynamic model of DL (Grimshaw et al, 2003). Hence, the observation of bilateral involvement does not necessarily mean that both hemispheres are processing the same information/task, and the observation of unilateral processing does not necessarily mean that the contralateral hemisphere does not play a role.

With all this in mind, there are some relevant issues that this model does not take into account. First, the process does not need to end at an evaluation stage, as natural responses to emotional stimuli are generally behaviourally oriented (Vuilleumier, 2005) a fourth stage associated to goal-orientation or decision-making might be required to fully understand emotional language processing. Indeed, multistage models of intrinsic affect (i.e. anxiety) have proposed a fourth stage associated to goal-oriented processing (Bar-Haim et al., 2007). Furthermore, anxious earlier or later processing stages have been observed to present different lateralization patterns (Spielberg et al., 2013). These might play a relevant role for emotional language processing, especially if they temporally match some of the stages associated to language informational properties used to convey emotion, such as those proposed by the multistep model of emotional language.

Anxiety and Threat: Affect Lateralization

The lateralization of affect might not only depend upon processing the emotional content of a particular stimulus that can induce a specific emotion (e.g. threat-inducing fear or anxiety), but also upon individual differences between participants that may also cause different lateralization patterns. This is especially indicated by studies that demonstrate such variation not only when processing emotional stimuli but also during resting state (Nietschke et al., 2000; Engels et al., 2007). Variation in lateralization patterns related to anxiety have been discussed from a number of different theoretical perspectives. First, anxiety has been proposed to be elicited by a behavioural inhibition system (BIS), which stops approaching behaviour of the organism in order to allow this organism to scan the environment in search of potential threat (McNaughton and Gray, 2000; Corr and McNaughton, 2012). Second, in the approach-withdrawal model, LH would be more engaged in approach-related emotions, while RH would show more involvement on withdrawal-related emotions (Davidson, 1992). This has been also captured by the valence-arousal model (Heller et al, 1997), where two types of anxiety are distinguished: anxious apprehension (worry-related) and anxious arousal (physiological hyperarousal), processed by LH and RH respectively. Third, a multistage model of anxiety proposes that anxiety develops through four processing stages: pre-attentive threat

evaluation, resource allocation, guided threat evaluation, and a goal engagement system (Bar-Haim el al, 2007). Where a threatening stimulus is pre-attentively perceived first, this elicits a type of cognitive/affective resources allocation (e.g. attend the stimulus or not), which on time induces attentive threat evaluation, and finally this induces goal-oriented/decision-making processing (e.g. approach or withdraw).

In effect, models of anxiety processing propose BIS as a conflict resolution system (Corr and McNaughton, 2012), where anxiety can be interpreted as a plausible intermediate state between approach and withdrawal, or calm and fear. Here, behaviour inhibition and arousal increase in preparation to approach/withdraw responses when possible or needed (McNaughton and Corr, 2014). In other words, behavioural inhibition for environmental scanning might increase arousal levels (McNaughton and Gray, 2000), which can induce fear-related responses if stimuli within the environment appear threatening enough. Thus, the interplay of lateralization patterns associated to worry and arousal might develop differently through the time-course of stimulus evaluation. Indeed, evidence from a functional magnetic resonance imaging (fMRI) study indicates that emotional language induces different lateralization responses, at different processing stages, for different types of anxiety (Spielberg et al, 2013). Where anxious apprehension is associated to a later and continued involvement of LH structures, interpreted as over-engagement with threat (e.g. rumination), and anxious arousal was associated with a faster and of shorter duration RH response interpreted as over-attention to threat. Important to notice that other models of anxiety also recognize a strong association between worry, rumination and arousal, induced by sustained anxious states (McLaughlin et al, 2007).

Electroencephalography (EEG) research using the event-related potential technique (ERP), which offers high temporal resolution, has observed this over-attention and over-engagement response when threatening faces are used as stimuli (Eldar et al, 2010). Also, over-engagement with threat has been observed when people with generalized anxiety disorder respond to threatening images (MacNamara and Hajcak, 2010). Furthermore, recent research has observed that socially anxious people present a right lateralized over-attention response to threatening words (Wabnitz, 2015). Hence, trait anxiety might directly affect the processing of emotional language. Previous EEG evidence indicates that anxiety has an effect on the recognition of prosody (Pell et al, 2015). However, not much is known about the interaction between emotion (threat) as conveyed through different information channels (prosody, semantics) and intrinsic affect (anxiety). If phasic lateralization patterns are integrated in a multistage model of anxiety (Bar-Haim et al, 2007) and this is compared to a multistep model

of emotional language (Kotz and Paulmann, 2011), then it might be possible predict very specific behavioural responses for anxious and non-anxious people. More precisely, there is a possible overlap between language processing mechanisms and anxiety processing mechanisms, which could become evident by comparing how people with higher trait anxiety processes different types of speech (i.e. prosody and semantics) as compared to less anxious people.

Present Experiment

Few dichotic listening (DL) experiments have researched the effects of anxiety on emotional speech processing (Gadea et al, 2011). They either use speech/prosody as an emotion-eliciting stimulus or use DL mainly as an attentional manipulation technique (Bruder et al., 1999; 2005; Leshem, 2018; Peschard et al., 2016; Sander et al., 2005). As a result, they are limited in the extent to which they reveal the relationship between dynamic variations in emotion language processing (prosody/semantics). Instead, studies focusing on the dynamic properties of emotional language, using DL or not (e.g. measuring laterality through electrophysiological measures), do not tend to consider individual differences (e.g. Godfrey and Grimshaw, 2015; Grimshaw et al., 2003; Kotz and Paulmann, 2007; Paulmann and Kotz, 2012; Techentin et al., 2009; Wabacq and Jerger, 2004). Therefore, on one side of the picture speech stimuli are typically treated as generic threatening stimuli, so possible differences induced by the informational features of speech that may vary over time are overlooked. On the other side, participants are typically regarded as a homogeneous group, so possible differences induced by anxiety-related processing, that may vary over time and may differ across informational features are overlooked.

Another important thing to consider is that in natural speech, emotional prosody might not be constrained to a single word, as is the case in the experimental manipulations of most of the studies we have cited above. However, semantics is always constrained by sentence's structure and lexical meaning. In other words, while a lexical item needs to be identified within a sentence in order for emotional semantics to be recognized, prosody might be expressed from the beginning of a sentence. This makes difficult to generalize from word level, or highly controlled sentences, to real world emotional utterances.

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 a direct-threat condition: identifying

whether threatening a stimulus occurred on the left or right ear, and in an indirect-threat condition: identifying whether a neutral stimulus occurred on the left or right ear. Participant's anxiety level was measured by using a psychometric scale. By so doing we were able take advantage of past studies researching the attentional effects of threatening language on anxiety and of studies researching the dynamics of speech's informational properties within a single study.

Both speech processing and anxiety literature seem to converge on theoretical perspectives incorporating multistep models, so we designed two web-based experiments to tap into different points in processing for which individual variation in anxiety may affect speech. In particular, we aimed to differentiate responses made at late evaluative stages (delayed response) vs. responses made at earlier attentive stages (online response) as early over-attention to threat (Bar-Haim et al., 2007) might affect earlier prosody/semantic lateralization patterns (Kotz and Paulmann, 2011), and later over-engagement with threat (Bar-Haim et al., 2007) might affect later emotional language evaluation stages (Kotz and Paulmann, 2011). Thus, Experiment-1 required participants to wait until after sentences' offset to respond (delayed response), and Experiment-2 required participants to respond during sentence presentation (online response).

For Experiment-1 we hypothesize that anxious over-engagement with threat at mid-late evaluative stages (Bar-Haim et al., 2007) should increase left hemisphere (LH) engagement (Spielberg et al, 2013), disturbing possible LH to right hemisphere (RH) information transferring (Grimshaw et al., 2003; Kotz and Paulmann, 2011). Hence, we predict that a left ear advantage (LEA), usually observed in DL experiments as an effect of prosody/emotional stimuli (Godfrey and Grimshaw, 2015; Grimshaw et al., 2003), 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 responding to semantically threatening but prosodically neutral stimuli (which we named Semantic stimuli).

For Experiment-2 we expect that, as responses are forced to be faster (online), prosody should induce the most noticeable effects, as online responses may overlap with early-mid emotional processing stages (Kotz and Paulmann, 2011). Therefore, we hypothesize that higher anxiety should reduce LH involvement (Spielberg et al., 2013) due to over-attention to threat effects, characteristic of earlier-mid processing stages (Bar-Haim et al., 2007). Hence, we predict an enhanced LEA for highly anxious participants, especially for prosodically threatening but semantically neutral stimuli (which we named Prosody stimuli). Thus, faster

and more accurate responses for anxious people at their left ear when attending prosodic stimuli.

Experiment 1: Delayed Response

Methods

Participants

Participants were recruited using Prolific (prolific.ac). 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. Only participants between 24 and 40 years old were accepted to take part. After exclusion, due to poor accuracy or not finishing the task properly, 44 participants (mean age = 31.7, 27 females) were retained. Participants were remunerated on a £7.5/hour rate. All participants gave their informed consent before participating.

Materials

Four types of sentences were recorded: Prosody (neutral-semantics and threatening-prosody), Semantic (threatening-semantics and neutral-prosody), Congruent (threatening-semantics and threatening-prosody), and Neutral (neutral-semantics and neutral-prosody). Threatening sentences were extracted from movie subtitles by matching them with a list of normed threatening words from the extended Affective Norms for English Words (ANEW) (Warriner et al., 2013). For the present study, any word over 5 points in the arousal scale, and below 5 points in the valence and dominance scales was considered threatening. Every word with less than 5 arousal points and between 4 and 6 (inclusive) valence points was considered neutral. Words' frequencies were extracted from SUBTLEX-UK (van Heuven et al., 2014), only sentences with Zipf log frequencies over 3 were included. Before recording, ten participants rated the threat level of each sentence by using a 0-8 Likert scale presented in Gorilla (gorilla.sc). Sentences' mean ratings were analysed using the Bayesian Estimation Superseeds t-test (BEST) method (Kruschke, 2012). Threatening semantics' ratings (m = 2.92) were considerably higher than neutral semantics' ratings (m = 0.76). See Annex for results.

After this, 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 angry threatening/angry or neutral voice for recording Prosody/Congruent and Semantic/Neutral sentences respectively. Due to a technical problem several sentences were recorded with very low amplitude. Therefore, sentences were normalized and cleaned from noise in Audacity (audacityteam.org). Figure 1 shows oscillograms and spectrograms of four example sentences.

Sentences' prosodic bio-informational dimensions (BIDs) were extracted using ProsodyPro (Xu, 2013) in Praat (praat.org). Three BIDs measures were compared using BEST (see Annex): Hammarber index (maximum energy differences between the 0-2000hz and 2000-5000hz ranges), harmonocity (amplitude signal to noise ratio), and median pitch (F0). Spectral and intensity measures (i.e. Hammarberg index and harmonicity) did not show consistent similarities between types, where Neutral presents the lowest harmonicity and Semantic the highest Hammarberg index, inconsistencies which might be the result of normalization and cleaning process.

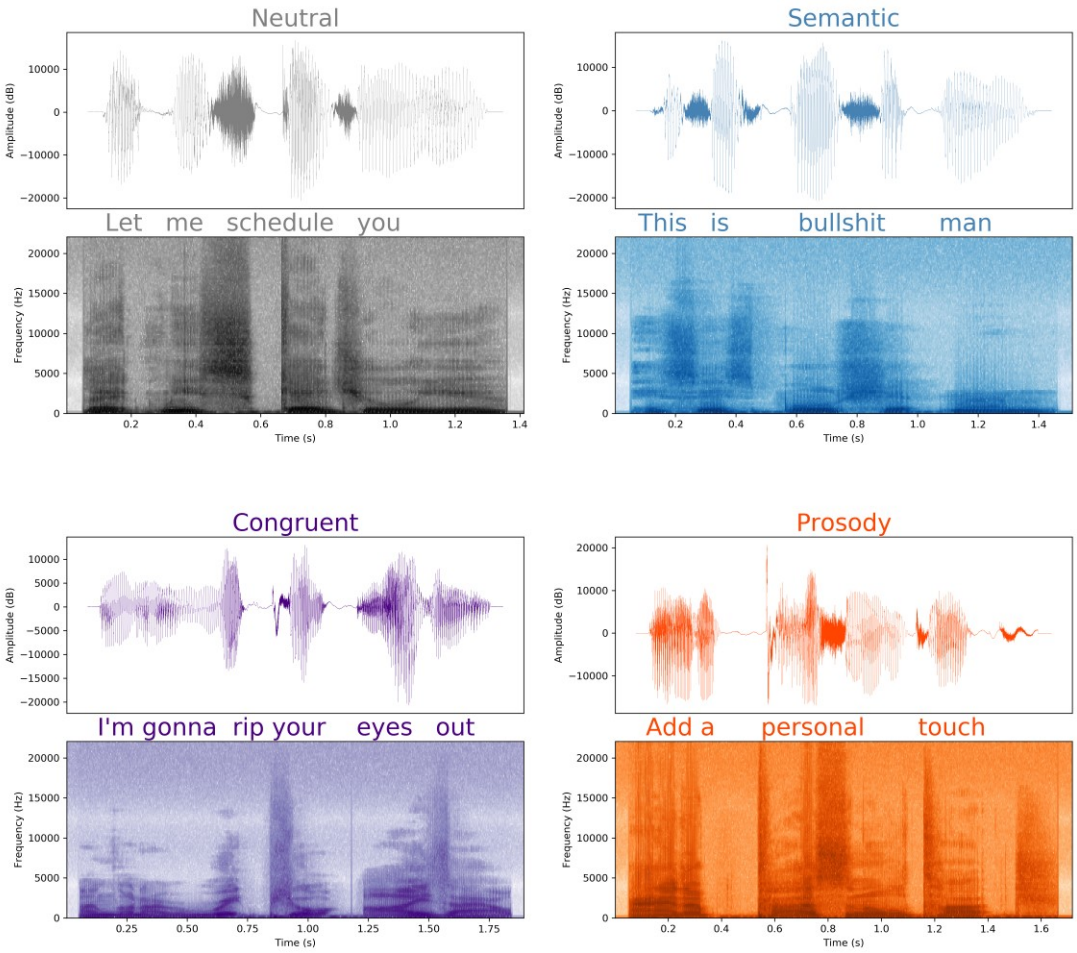

Figure 1. Examples of sentences used as stimuli: neutral prosody and neutral semantics (upper left), neutral prosody and threatening semantics (upper right), threatening prosody and threatening semantics (bottom left), threatening prosody and neutral semantics (bottom right).

However, F0 comparisons, crucial for defining angry or threatening voices (Banse and Scherer, 1996), indicate a higher median F0 for prosodically threatening stimuli, Prosody (m = 137.17Hz) and Congruent (m = 129.04Hz), over prosodically neutral stimuli, Semantic (90.33Hz) and Neutral (m = 97.37Hz). This aligns prosodically threatening stimuli with hot-anger (rage), which has higher F0 than cold-anger or neutral prosody (Banse and Scherer, 1996; Hammerschmidt and Jürgens, 2007). This is also consistent with previous dichotic listening studies' F0 values of angry-prosody stimuli (Godfrey and Grimshaw, 2015; Grimshaw et al., 2009).

To check this, a random subset of 7 prosody-only sentences was compared to a random subset of 7 neutral sentences in an online rating questionnaire in the same manner as semantic threat. Ten participants rated these spoken sentences in Gorilla (gorilla.sc). These ratings were analysed using BEST (see Annex for results). Results showed that threatening prosody (m = 4.06) is rated as more threatening than neutral prosody (m = 1.2). Next, sentences were paired using Audacity, sentences were paired such as their durations were as similar as possible. Silences between words were extended, never surpassing 40ms, to match sentences' latencies as closely as possible. After this, sentences were allocated to one of the stereo channels (left or right) of the recording; each pair was copied with mirrored channels. A silence (~50ms) was placed at the beginning and at end of each pair. This resulted in a total of 480 pairs where 80 sentences of each type (congruent, semantic, prosody) were each paired with a neutral sentence of the same length twice, so every sentence was presented once at each ear.

Procedure

Before starting the experiments, participants answered the Penn State Worry Questionnaire (PSWQ) (Meyer et al., 1990) to assess their worry-level, and the Anxious Arousal sub-scale of the Mood and Anxiety Symptoms Questionnaire (MASQ-AA) (Watson et al., 1995) to assess their arousal level. This follows previous approaches (Nitschke et al., 1999), with the difference that we used PSWQ scores as continuous predictor instead of splitting participants between high and low anxiety groups. Also, we did not use MASQ scores as no participants scored above the median (indicating that all participants presented low arousal levels).

After a practice session, 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 (Aue et al., 2011; Peschard et al., 2016). 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.

Analysis

Reaction time (RT) data were recorded in milliseconds and accuracy was coded as correct=1 and else=0. Participants with hit rates below 70% were excluded. Two Bayesian hierarchical models were built for reaction time (RT) and accuracy. The RT model, shown on Figure 2, was the basic model structure for all analyses, based on Kruschke (2014) and Martin's (2018) guidelines. The model for indirect-threat RT was identical excepting the number of observations (obs. = 5,427). Models for accuracy used a Binomial distribution for the likelihood, where direct-threat = 6,181 obs., and indirect-threat = 5,767 obs. Differences in the number of observations are due to the fact that RT data used only correct responses, and accuracy data was cleaned from erratic responses (i.e. incomplete trials, overlapping responses and extremely delayed response over 3s).

RT models used a robust regression (Kruschke, 2014) in order to account for outliers through a long-tailed Student-t distribution. In this way, RTs that are implausibly fast or implausibly slow do not need to be removed, but can be dealt with statistically. Both accuracy and RT models were sampled using Markov Chain Monte Carlo (MCMC) No U-turn Sampling (NUTS) as provided by PyMC3 (Salvatier et al., 2016). Two chains of 2000 tuning steps and 2000 samples each were used. Plots of results were produced using Arviz (Kumar et al, 2019) and Matplotlib (Hunter, 2007). A region of practical equivalence (ROPE) of 2SDs was established as main criterion for deciding whether posterior distributions provided evidence for effects, where 90% high posterior density intervals (HPDs) which overlapped substantially with the ROPE (over 80%) were considered null. HPDs partially overlapping were considered to provide no evidence (neither null nor effect), HPDs completely outside a 1SD ROPE (but partially inside 2SD ROPE) were considered to provide very weak evidence, and HPDs completely outside the ROPE were considered to provide evidence for the effect.

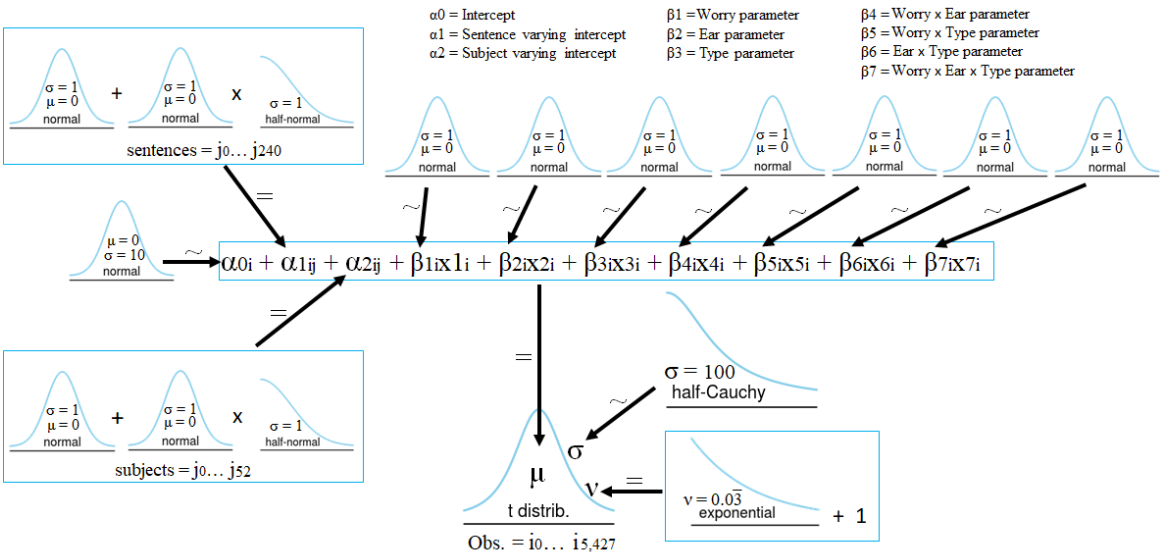

Figure 2. The diagram shows the distributions used for robust regression on amplitude. Arrows indicate the relationship between a parameter and priors/hyperpriors, where tilde (~) indicates a stochastic relationship and equal (=) indicates a deterministic relationship. Observations are equivalent to RTs in milliseconds.

Results

Accuracy models for the direct-threat condition showed null results for the main Worry x Ear x Type interaction. This means that Prosody HPD ([-0.02, 0.02]) and Semantic HPD ([-0.04, 0.01]) overlapped substantially with the ROPE ([-0.03, 0.03]). Apparent tendencies toward an Ear effect were revealed to be spurious due to considerable variability, with models providing no clear evidence for any effect; although mean accuracy for left (m=87%, SD=34%) or right (m=72%, SD=45%) ears differed by 15%. Accuracy results for the indirect-threat condition showed similar patterns. Again, Prosody HPD ([-0.03, 0.01]) and Semantic HPD ([-0.04, 0.01]) overlapped substantially with the ROPE ([-0.03, 0.03]). Accuracy means between low worriers left (m=88%, SD=32%) and right (m=88%, SD=33%) ears are almost identical. Otherwise, high worriers left (78%, SD=42%) and right (m=87%, SD=33%) ears seem to exhibit a small, but spurious difference, as the Worry x Ear interaction is confirmed as null (HPD = [-0.03, 0.02], ROPE = [-0.03, 0.03]).

Results for direct-threat RT showed a strong effect of Worry and a weak effect of a Worry x Semantic interaction (HPD = [0.24, 0.73], ROPE = [-0.3, 0.3]), where the HPD slightly overlaps with the 2SDs ROPE. All other effects were either null or did not have enough supporting evidence to warrant conclusions. Results for indirect-threat also indicate a strong effect of Worry, and a mild effect of Worry x Semantic (HPD = [0.34, 0.86], ROPE = [-0.31, 0.31]), where the HPD is barely away from the 2SDs ROPE. Figure 3 summarizes Worry

effects for direct- and indirect-threat. Note that for direct-threat, the effect of worry indicates 236.2ms of estimated relative increase from the lowest to highest PSWQ score, namely from 33 points (m=229.2ms, SD=7.6ms) to 67 points (m=465.4ms, SD=15.4ms). And indirect-threat shows 238.9ms of estimated relative increase from 33 points (m=231.9ms, SD=7.9ms) to 67 points (m=470.8ms, SD=16.0ms). Overall results indicate that effects of Type are very mild or spurious, and the strong main effect of Worry represents a slow down of responses irrespective of threat direction.

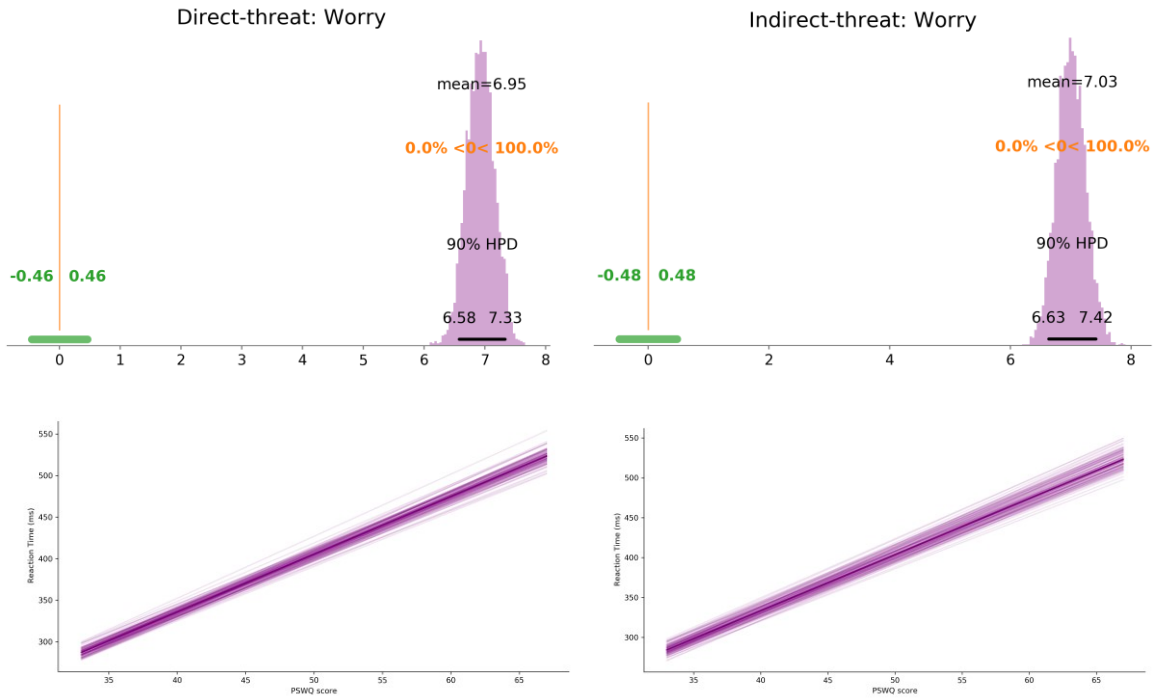

Figure 3. Upper panels show posterior distributions of Worry effect for direct-threat (left) and indirect-threat (right); green bands: 2SDs ROPEs, black bands: 90% high posterior density (HPD) interval. Bottom panels show regression lines from the posterior (mean) for respective conditions; faded lines represent uncertainty of the estimate. Both results indicate that when the PSWQ score increases in 1 point, RT increases ~7ms ± 0.25ms.

Discussion

Present results are not consistent with our hypothesis, instead indicating that when responses are delayed, there are no effects of ear or stimulus type. However, there remains a substantial effect of worry level. These patterns indicate that our extension of dichotic listening models (Grimshaw et al., 2003) does not guarantee laterality effects induced by sentences' information type. However, our prediction of a strong effect of worry level, affecting emotional language processing due to possible over-engagement with threat (Bar-Haim et al., 2007; Spielberg et al., 2013), was strongly supported. Therefore, a plausible explanation is that

delayed responses facilitate over-engagement due to the long latency between sentence presentation and response. This, together with the high variability in sentences' durations and content might have nullified ear and/or type effects. This suggests that any type of threatening language, attended either directly or indirectly, strongly affects higher worriers when a long deliberation (post-evaluation) stage is induced. Therefore, our proposal of adding a fourth stage to a multistep model of emotional language (Kotz and Paulmann, 2011) is only partially supported. In Experiment-2 we aim to tap an earlier stage in the process, using the same materials and task but instead compelling participants to make their responses as quickly as possible: before sentences' end. As this should tap into early-mid stages of emotional language processing, participants with higher anxiety should have reduced LH involvement, reflected in relatively facilitated responses to prosodically threatning stimuli presented at the left ear.

Experiment 2: Fast Response

Methods

Experiment2's methods were the same as Experiment-1's methods, with the following exceptions: 1) 52 participants (mean age = 31, 24 females) were included in the final analysis. Due to the extra difficulty of this experiment, the rejection threshold was reduced and participants with a hit-rate below 60% were excluded. 2) Participants were instructed to answer before the sentence finished and to withhold any response when a stop sign image appeared on the screen after sentences' end. 3) RT models were not reparametrised, such as $\alpha_1 \sim \mathcal{N}(\mu=0, \sigma=10)$ and $\alpha_2 \sim \mathcal{N}(\mu=0, \sigma=10)$; which in this case improved convergence.

Results

Accuracy analyses for direct-threat revealed null effects for all terms, excepting the Worry x Ear x Semantic interaction, which provided no evidence (HPD = [0.001, 0.05], ROPE = [-0.04, 0.04]). Indeed, accuracy means between left (m=75%, SD=43%) and right (m=73%, SD=44%) ears are quite similar. Means between low (m=77%, SD=42%) and high (m=69%, SD=46%) worry levels show a small but spurious difference. Indirect-threat accuracy shows the same pattern as direct-threat. Means for left (m=69%, SD=46%) and right (m=74%, SD=44%) ears show a small and spurious difference, same as low (m=74%, SD=44%) and high (m=67%, SD=47%) worry-level means.

Results from RT analyses, at both direct-and indirect-threat conditions, indicate null effects for Ear, a small effect of Worry x Prosody, and a very strong effect of Worry. The Worry x Prosody effect indicates a 49.42ms increase from low worry (m=183.22ms, SD= 19.39ms) to high worry (m=232.65ms, SD= 32.47ms) when answering to Prosody, as opposed to a 14.53ms decrease when answering to Semantic. The indirect-threat condition shows a very similar pattern, with a 48.19ms increase between low worry (m=233.03ms, SD= 18.77ms) and high worry (m=184.83ms, SD= 31.29ms) for Prosody, and a negligible 1.06ms decrease for Semantic. Note that low mean RT values are due to the spread of the estimation between answers to different Types. Figure 4 summarizes direct-threat results, indirect-threat results show the same pattern, with a slight certainty decrease.

Effects of Worry are considerably stronger, where direct-threat shows a 753.23ms increase from low worry (m=706.05ms, SD= 19.08ms) to high worry (m=1459.28ms, SD= 31.74ms), and indirect-threat shows a 754.59ms increase from low worry (m= 709.58ms, SD= 18.45ms) to high worry (m= 1464.17ms, SD= 30.52ms), where low worry = 26 PSWQ points, and high = 61 PSWQ points. Figure 5 summarizes these results.

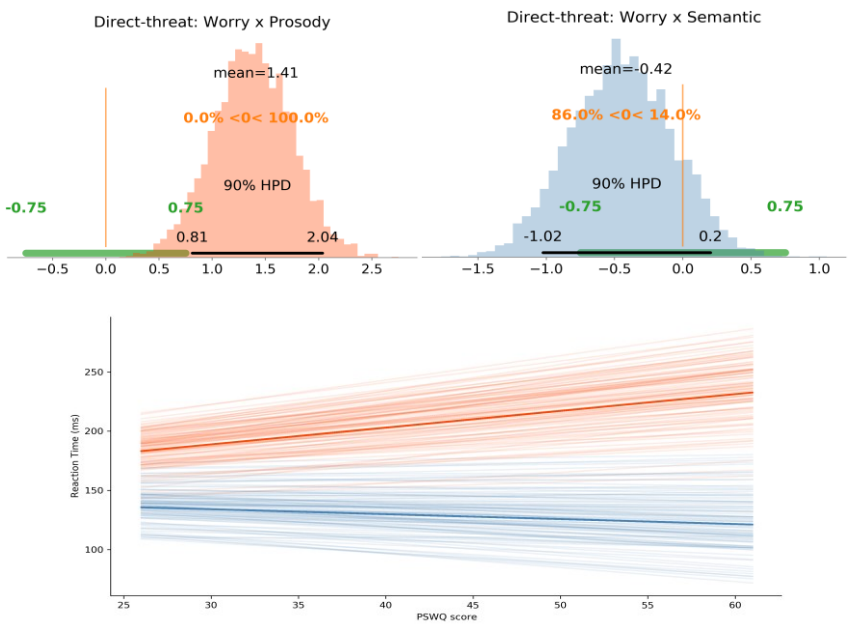

Figure 4. Upper panels show posterior distributions of Worry x Type interaction. Bottom panels show regression lines from the posterior (mean) for Prosody (upper line: orange-red) and Semantic (lower line: light-blue), faded lines represent uncertainty of the estimate. Note that differently from Semantic, Prosody is slightly outside the 2SDs ROPE.

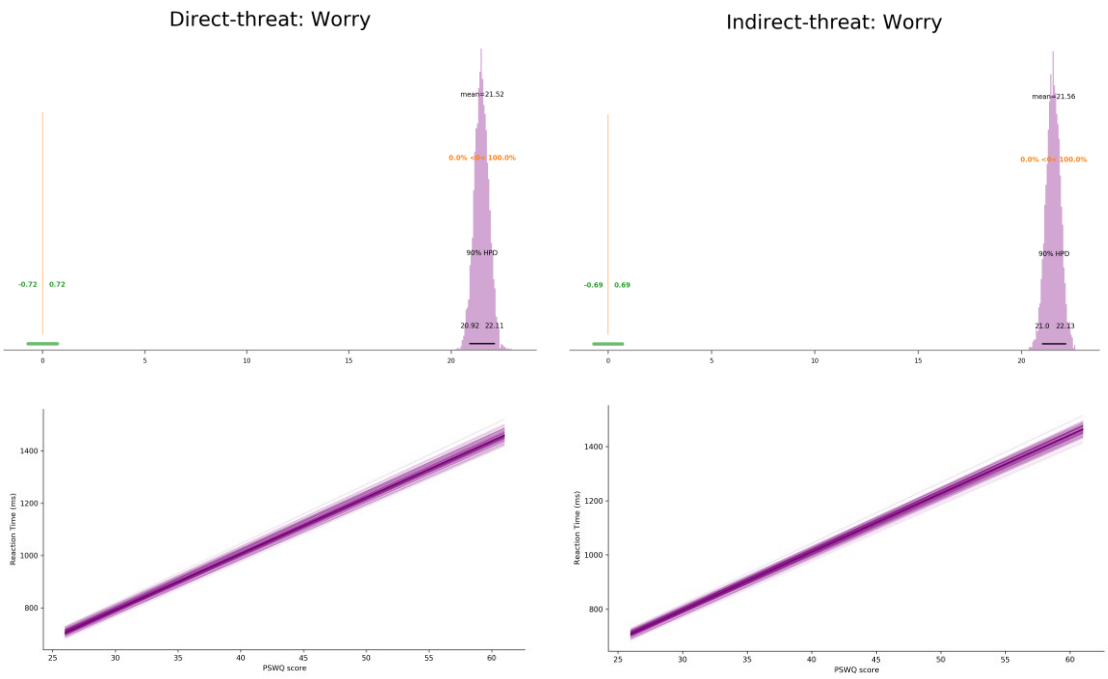

Figure 5. Upper panels show posterior distributions of Worry effect for direct-threat (left) and indirect-threat (right); green bands: 2SDs ROPEs, black bands: 90% high posterior density (HPD) interval. Bottom panels show regression lines from the posterior (mean) for respective conditions; faded lines represent uncertainty. Both results indicate that when the PSWQ score increases in 1 point, RT increases ~21.5ms ± 0.35ms.

Discussion

Present results indicate that there are no ear effects whatsoever, contrary to our expectations. However, forcing fast responses suggest that as worry increases Prosody is harder to recognise; as evidenced by the interaction between Worry and Prosody. This is remarkable, as answering to sentences of the Semantic type requires identifying a threatening word that is seldom present from sentences' onset. A callosal relay mechanism could be proposed as the factor delaying the processing of threatening prosody (Grimshaw et al, 2003), but the lack of ear differences blocks this interpretation. Our second expectation was a pattern resembling previous dichotic listening experiments (e.g. Peschard et al, 2017), where ear preference varies according to type of emotion and threat-direction. Instead, we found that anxiety is the main modulator of responses, no matter threat direction. The only crucial difference was a considerable slow-down of responses for the indirect threat condition, which might simply suggest that participants attempt to identify threat first and answer to the opposite ear later. Overall, the experiment supports the idea that worry, possibly through excessive deliberation before response, interferes with the processing of threatening language.

General Discussion

In order put the present results in context, it is important to recapitulate important aspects that differentiate the current experiments from previous relevant studies: 1) The use of worry-level as a continuous variable. Worry is associated with anxious apprehension (Heller et al, 1997), which implies more chances of participants over-engaging with threat. 2) Stimuli were semi-naturalistic sentences, providing stronger contextual effects. In addition, their longer durations can facilitate engagement with their content. 3) Information channels were manipulated to disentangle effects of semantics and prosody from effects of emotional expression (Kotz and Paulmann, 2007). 4) The use of two tasks measuring responses directed to threatening or neutral stimuli (direct vs. indirect threat, e.g. Sanders et al, 2005) helps to check whether attention effects could be inducing different response patterns. 5) Two experiments were implemented to verify whether answering after sentences' end or during sentence presentations (fast or delayed) can influence laterality patterns by tapping into different moments of a multistep emotional language processing mechanism (Kotz and Paulmann, 2011).

With this in mind, it is important to carefully interpret the lack of laterality (ear) effects in Experiments 1 and 2, and the mild effect of Prosody in Experiment 2. Null ear effects might be explained by the great variability between items and the high duration (also very variable) of sentences. However, the lack of sensitivity of DL when more naturalistic stimuli/context are provided cannot be discarded as a possible explanation. If DL effects are task dependent (Godfrey and Grimshaw, 2015), increased naturalness on stimuli and context can bring out a myriad of bilateral processing patterns that might make ear advantages disappear on the long run of prolonged auditory stimuli. What made Experiment-2 results different from Experiment-1's was a slow-down in responses to Prosody stimuli as a function of worry-level. Previous evidence suggests a right lateralized pattern for prosody vs. semantic evaluation in an EEG experiment (not considering anxiety), using a congruency (not DL) task with sentences as stimuli (Kotz and Paulmann, 2007; but see: Paulmann and Kotz, 2012). Although this pattern is explained by the strong association between pitch recognition and RH engagement (Kotz and Paulmann, 2007; Zatorre et al., 2002), there are other frequency and spectral features that might be important for recognizing both threatening and neutral sentences (Banse and Scherer, 1996; Hammerschmidt and Jürgens, 2007; Liu et al., 2013; Xu et al., 2013; Zatorre et al., 2002). This could imply that distinguishing prosody and semantics might be a continuous process that can have diversified effects even during sentence presentation.

Indeed, by manipulating angry prosody changes at the beginning and end of sentences, an EEG study has observed that when prosody changes from angry to neutral within sentences, processing is more effortful (Chen et al, 2011). This might indicate that the rich acoustic nature of prosody might be detected quickly but resourcefully. Recent EEG research has observed that anxious people present ERP differences at both early and late processing stages when answering to threatening prosody and non-language vocalizations (Pell et al, 2015). This is consistent with the notion of early over-attention and later over-engagement, and indicates that behavioural responses might change given early or late variations in threat. The present study indicates that when responses are forced to be fast (Experiment 2), the only visible effect is Prosody, namely semantically neutral sentences with threatening prosody only.

Another possible explanation is callosal relay (Atchely et al., 2011; Grimshaw et al, 2003), where increased anxiety would disrupt RH to LH callosal information transferring of threatening prosody. It has been proposed that callosal relay is highly relevant for language informational and emotional processing (Friederici et al, 2007; Kotz and Paulmann, 2011; Steinmann et al., 2017). Hence, interference at one hemisphere (e.g. rumination or worry impacting LH) can have an effect on information transferring to the other. Thus, callosal relay effects could have a relevant impact on how DL tasks are processed, subject to both top-down and bottom-up effects (Westernhausen and Hugdahl, 2008), which is particularly relevant when laterality effects induced by acoustic properties (e.g. prosody) need to be disentangled from those induced by emotional processing (Grimshaw et al, 2003; Leshem, 2018).

Nevertheless, we need to acknowledge that present Prosody effects might be due to present stimuli incongruency, where Prosody stimuli are difficult to distinguish, as acoustic features of threatening prosody were limited to pitch differences. Differently, Semantic and Congruent stimuli might only require lexical recognition. Although it is hard to explain why anxious participants would answer slower to stimuli that are not threatening, we cannot fully discard an alternative explanation: Prosody stimuli being harder to separate from Neutral ones disrupt worriers slowed-down deliberation even more. It might be possible to test this possibility by conducting studies that use more highly controlled materials varying along more spectral and frequency features associated with threatening prosody (e.g. Liu et al., 2013). In addition, the dominant effect of worry we observed, which slow RTs regardless of ear and mostly regardless of stimulus type, need to be further confirmed, as their implications can be deep. The crucial implication, in this case, is that under more naturalistic conditions higher worriers may only differ from lower worriers in that they answer more slowly. In other words, higher worry affects the processing of any type of threatening language in the same way: by

increasing reaction times (although this would not fully explain the effects we observed: in particular that the slowdown in Experiment-2 was not equally observed for all stimulus types)..

A very plausible explanation for this effect is repetitive thinking which can be also associated with rumination and/or worry as a feature of anxious apprehension (Nitschke et al., 1999; Spielberg et al., 2013), or as a marker of an over-reactive behavioural inhibition system (Corr and McNaughton, 2012). The long duration of present stimuli might have been a decisive factor for inducing a strong effect of worry. However, previous research using single words, dichotically presented as direct- and indirect-threat (or anger), and measuring anxiety, did not find differences in RT for left or right ears (Sander et al., 2005; Leshem, 2018; Peshard, 2016), but did find differences in attention focus per ear. Present results indicate that RTs differ in neither of these conditions, which is supported by the remarkably similar posterior distributions for direct- and indirect-threat and the high certainty of these estimates. Recent research (Leshem, 2018) did not find effects of trait anxiety on ear either; present results, going even further, evidence null interactions between ear and worry (trait anxiety). Although the absence of other effects might be induced by stimuli's high variability in length and content, it is also important to emphasize that present analyses are fairly robust. Previous statistical approaches use averaged RTs and accuracy measures, and force such data on normal distributions; thus, increasing chances of type I error. In this sense, current results can be trusted as accurate and are good candidates for replicability.

A more general caveat of our approach lies on the nature of the experiment itself. Behavioural measures such as DL, though able to portrait a very general picture of underlying brain processes, might not be enough. Better spatial and temporal resolution is required to disentangle laterality and early stage effects of threatening language. The latter is particularly relevant, as the time-course of emotional language processing might have crucial differences at much shorter time-scales, as evidenced by previous EEG research (Chen et al, 2011; Kotz and Paulmann, 2007; Paulmann and Kotz, 2012; Pell et al, 2015; Wabnitz et al., 2015; Wambacq and Jerger, 2004). In consequence, present tasks could be replicated by using EEG measures, in particular Experiment-1, where EEG measures such as event-related potentials could provide richer information about processing occurring during sentence listening, before response preparation and response execution. This could also provide lab results as point of comparison with present web-based results. But more importantly, this is crucial for identifying differences in the neural signature of worry and language processing, indispensable for properly understanding time-related models of language and anxiety processing.

In conclusion, present results indicate that extending multistep models of language processing (Schirmer and Kotz, 2006; Kotz and Paulmann, 2011) by including aspects of multistage models of anxiety (Bar-Haim et al, 2007; Corr and McNaughton, 2012) could be a relevant theoretical move. The current multistep model proposes three stages that can be understood as early (perception), mid (recognition), late (evaluation); or as pre-attentive, attentive and evaluative stages. A fourth orientative stage, associated with deliberation, can help to understand aspects of goal-directed processes before response. Complementing this model, however, might be insufficient. Further theoretical development, including quantitative modelling, the inclusion of physiological correlates, and more precise anatomical mappings, might be necessary. Further experimental testing is thus required, in particular by implementing physiological measures such as EEG, and tasks that do not involve DL, using more controlled stimuli and investigating the effects of stimuli below or above the sentence level, such as phrases or narratives.

Declaration of interests

None.

Data statement

All data, analyses' scripts, and additional info can be found at the Open Science Framework (OSF) repository: https://osf.io/z8pgf/?view_only=b5da5ce6c8644bc182231cd9b96be173.

Acknowledgements

This research is supported by CONICYT Becas Chile 72170145 (https://www.conicyt.cl/). Contributions to the study were as follows. **Busch-Moreno**: Conceptualization, Methodology, Software, Analysis, Investigation, Resources, Curation, Writing, Visualization, Administration. **Vinson**: Conceptualization, Methodology, Resources, Revision, Supervision, Administration. **Tuomainen**: Conceptualization, Resources, Revision, Supervision, Administration. Preliminary results of this experiment, using a frequentist framework, were presented at the International Congress of Phonetic Sciences (ICPhS), Melbourne, 2019.

References

Atchley, R. A., Grimshaw, G., Schuster, J., & Gibson, L. (2011). Examining lateralized lexical ambiguity processing using dichotic and cross-modal tasks. *Neuropsychologia*, *49*(5), 1044–1051. https://doi.org/10.1016/j.neuropsychologia.2011.01.010

Aue, T., Cuny, C., Sander, D., & Grandjean, D. (2011). Peripheral responses to attended and unattended angry prosody: A dichotic listening paradigm. *Psychophysiology*, *48*(3), 385–392. https://doi.org/10.1111/j.1469-8986.2010.01064.x

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. https://doi.org/10.1037//0022-3514.70.3.614

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–135. https://doi.org/10.1016/j.tics.2004.01.008

Bruder, G. E., Wexler, B. E., Stewart, J. W., Price, L. H., & Quitkin, F. M. (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. https://doi.org/10.1037/0021-843x.108.2.233

Chen, X., Zhao, L., Jiang, A., & Yang, Y. (2011). Event-related potential correlates of the expectancy violation effect during emotional prosody processing. *Biological Psychology*, *86*(3), 158–167. https://doi.org/10.1016/j.biopsycho.2010.11.004

Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: A two stage (valuation–motivation) approach. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2339–2354. https://doi.org/10.1016/j.neubiorev.2012.09.013

Davidson, R. J. (1992). Emotion and Affective Style: Hemispheric Substrates. *Psychological Science*, *3*(1), 39–43. https://doi.org/10.1111/j.1467-9280.1992.tb00254.x

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. https://doi.org/10.1016/j.biopsycho.2010.07.010

Engels, A. S., Heller, W., Mohanty, A., Herrington, J. D., Banich, M. T., Webb, A. G., & Miller, G. A. (2007). Specificity of regional brain activity in anxiety types during emotion processing. *Psychophysiology*, *44*(3), 352–363. https://doi.org/10.1111/j.1469-8986.2007.00518.x

Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (2007). Role of the Corpus Callosum in Speech Comprehension: Interfacing Syntax and Prosody. *Neuron*, *53*(1), 135–145. https://doi.org/10.1016/j.neuron.2006.11.020

Gadea, M., Espert, R., Salvador, A., & Martí-Bonmatí, L. (2011). The sad, the angry, and the asymmetrical brain: Dichotic Listening studies of negative affect and depression. *Brain and Cognition*, *76*(2), 294–299. https://doi.org/10.1016/j.bandc.2011.03.003

Godfrey, H. K., & Grimshaw, G. M. (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. https://doi.org/10.1080/1357650x.2015.1096940

Grimshaw, G. M., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003). The dynamic nature of language lateralization: effects of lexical and prosodic factors. *Neuropsychologia*, *41*(8), 1008–1019. https://doi.org/10.1016/s0028-3932(02)00315-9

Grimshaw, G. M., Séguin, J. A., & Godfrey, H. K. (2009). Once more with feeling: The effects of emotional prosody on hemispheric specialisation for linguistic processing. *Journal of Neurolinguistics*, *22*(4), 313–326. https://doi.org/10.1016/j.jneuroling.2008.10.005

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical Correlates of Affective Prosody. *Journal of Voice*, *21*(5), 531–540. https://doi.org/10.1016/j.jvoice.2006.03.002

Heller, W., Nitschke, J. B., Etienne, M. A., & Miller, G. A. (1997). Patterns of regional brain activity differentiate types of anxiety. *Journal of Abnormal Psychology*, *106*(3), 376–385. https://doi.org/10.1037/0021-843x.106.3.376

Hugdahl, K. (2011). Fifty years of dichotic listening research – Still going and going and…. *Brain and Cognition*, *76*(2), 211–213. https://doi.org/10.1016/j.bandc.2011.03.006

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/mcse.2007.55

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*, *1151*, 107–118. https://doi.org/10.1016/j.brainres.2007.03.015

Kotz, S. A., & Paulmann, S. (2011). Emotion, Language, and the Brain. *Language and Linguistics Compass*, *5*(3), 108–125. https://doi.org/10.1111/j.1749-818x.2010.00267.x

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. https://doi.org/10.1037/a0029146

Kruschke, J. K. (2015). *Doing Bayesian data analysis : a tutorial with R, JAGS, and stan*. Amsterdam Etc: Elsevier, Academic Press, Cop.

Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *Journal of Open Source Software*, *4*(33), 1143. https://doi.org/10.21105/joss.01143

Leshem, R. (2018). Trait Anxiety and Attention: Cognitive Functioning as a Function of Attentional Demands. *Current Psychology*. https://doi.org/10.1007/s12144-018-9884-9

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural Substrates of Phonemic Perception. *Cerebral Cortex*, *15*(10), 1621–1631. https://doi.org/10.1093/cercor/bhi040

Liu, F., Xu, Y., Prom-on, S., & Yu, A. (2013). Morpheme-like prosodic functions: Evidence from acoustic analysis and computational modeling - UCL Discovery. *Ucl.Ac.Uk*. https://doi.org/http://discovery.ucl.ac.uk/1391681/1/Xu_Liu_etAl_JoSS2013.pdf

MacNamara, A., & Hajcak, G. (2010). Distinct electrocortical and behavioral evidence for increased attention to threat in generalized anxiety disorder. *Depression and Anxiety*, *27*(3), 234–243. https://doi.org/10.1002/da.20679

Martin, O. (2018). *Bayesian analysis with Python : introduction to statistical modeling and probabilistic programming using PyMC3 and ArviZ*. Birmingham, Uk: Packt Publishing.

McLaughlin, K. A., Borkovec, T. D., & Sibrava, N. J. (2007). The Effects of Worry and Rumination on Affect States and Cognitive Activity. *Behavior Therapy*, *38*(1), 23–38. https://doi.org/10.1016/j.beth.2006.03.003

McNaughton, N., & Corr, P. J. (2014). Approach, avoidance, and their conflict: the problem of anchoring. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00124

McNaughton, N., & Gray, J. A. (2000). Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety. *Journal of Affective Disorders*, *61*(3), 161–176. https://doi.org/10.1016/s0165-0327(00)00344-x

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *28*(6), 487–495. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2076086

Nitschke, J. B., Heller, W., Palmieri, P. A., & Miller, G. A. (1999). Contrasting patterns of brain activity in anxious apprehension and anxious arousal. *Psychophysiology*, *36*(5), 628–637. https://doi.org/10.1111/1469-8986.3650628

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. *Cognitive Science*, *33*(1), 127–146. https://doi.org/10.1111/j.1551-6709.2008.01007.x

Paulmann, S., Jessen, S., & Kotz, S. A. (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. https://doi.org/10.1016/j.neuropsychologia.2012.03.014

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological Psychology*, *111*, 14–25. https://doi.org/10.1016/j.biopsycho.2015.08.008

Peschard, V., Gilboa-Schechtman, E., & Philippot, P. (2017). Selective attention to emotional prosody in social anxiety: a dichotic listening study. *Cognition & Emotion*, *31*(8), 1749–1756. https://doi.org/10.1080/02699931.2016.1261012

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time.' *Speech Communication*, *41*(1), 245–255. https://doi.org/10.1016/s0167-6393(02)00107-3

Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1071–1086. https://doi.org/10.1098/rstb.2007.2160

Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of Affective Prosody in Brain and the Callosal Integration of Hemispheric Language Functions. *Brain and Language*, *56*(1), 27–54. https://doi.org/10.1006/brln.1997.1731

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, *2*, e55. https://doi.org/10.7717/peerj-cs.55

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage*, *28*(4), 848–858. https://doi.org/10.1016/j.neuroimage.2005.06.023

Schirmer, A., & Kotz, S. A. (2003). ERP Evidence for a Sex-Specific Stroop Effect in Emotional Speech. *Journal of Cognitive Neuroscience*, *15*(8), 1135–1148. https://doi.org/10.1162/089892903322598102

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Spielberg, J. M., De Leon, A. A., Bredemeier, K., Heller, W., Engels, A. S., Warren, S. L., … Miller, G. A. (2013). Anxiety type modulates immediate versus delayed engagement of attention-related brain regions. *Brain and Behavior*, *3*(5), 532–551. https://doi.org/10.1002/brb3.157

Steinmann, S., Meier, J., Nolte, G., Engel, A. K., Leicht, G., & Mulert, C. (2017). The Callosal Relay Model of Interhemispheric Communication: New Evidence from Effective Connectivity Analysis. *Brain Topography*, *31*(2), 218–226. https://doi.org/10.1007/s10548-017-0583-x

Techentin, C., Voyer, D., & Klein, R. M. (2009). Between- and within-ear congruency and laterality effects in an auditory semantic/emotional prosody conflict task. *Brain and Cognition*, *70*(2), 201–208. https://doi.org/10.1016/j.bandc.2009.02.003

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Wabnitz, P., Martens, U., & Neuner, F. (2015). Written threat: Electrophysiological evidence for an attention bias to affective words in social anxiety disorder. *Cognition and Emotion*, *30*(3), 516–538. https://doi.org/10.1080/02699931.2015.1019837

Wambacq, I. J., & Jerger, J. F. (2004). Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. *Cognitive Brain Research*, *20*(3), 427–437. https://doi.org/10.1016/j.cogbrainres.2004.03.015

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., & et al. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, *104*(1), 3–14. https://doi.org/10.1037//0021-843x.104.1.3

Westerhausen, R., & Hugdahl, K. (2008). The corpus callosum in dichotic listening studies of hemispheric asymmetry: A review of clinical and experimental evidence. *Neuroscience & Biobehavioral Reviews*, *32*(5), 1044–1054. https://doi.org/10.1016/j.neubiorev.2008.04.005

Xu, Y. (2013). ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis - UCL Discovery. *Ucl.Ac.Uk*. https://doi.org/http://discovery.ucl.ac.uk/1406070/1/Xu_TRASP2013.pdf

Xu, Y. (2019). Emotional expressions as communicative signals | Yi Xu, Andrew Kelly and Cameron Smillie. *Benjamins.Com*. https://doi.org/https://benjamins.com/catalog/ill.13.02xu

Zatorre, R. J. (2001). Neural Specializations for Tonal Processing. *Annals of the New York Academy of Sciences*, *930*(1), 193–210. https://doi.org/10.1111/j.1749-6632.2001.tb05734.x

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46. https://doi.org/10.1016/s1364-6613(00)01816-7

Annex

Results contained in this annex correspond to analyses of materials, where differences between means were tested as a methodological check-up to guarantee that stimuli at both the semantic and prosodic levels belonged to recognizable categories. In other words, a group of people provided ratings on how threatening the 'content' of sentences is, and another group on how threatening sentences 'sound'. In addition, we tested differences between three acoustic properties of sentences: Hammarberg index, Harmonicity and Median Pitch (F0). Only F0 was consistently different in threat vs. neutral comparisons, and consistently similar in threat vs. threat and neutral vs. neutral comparisons. Nonetheless, stimuli were rated as recognizably threatening in both their 'content' and 'sound'. Results below are displayed as plots of posterior distributions.

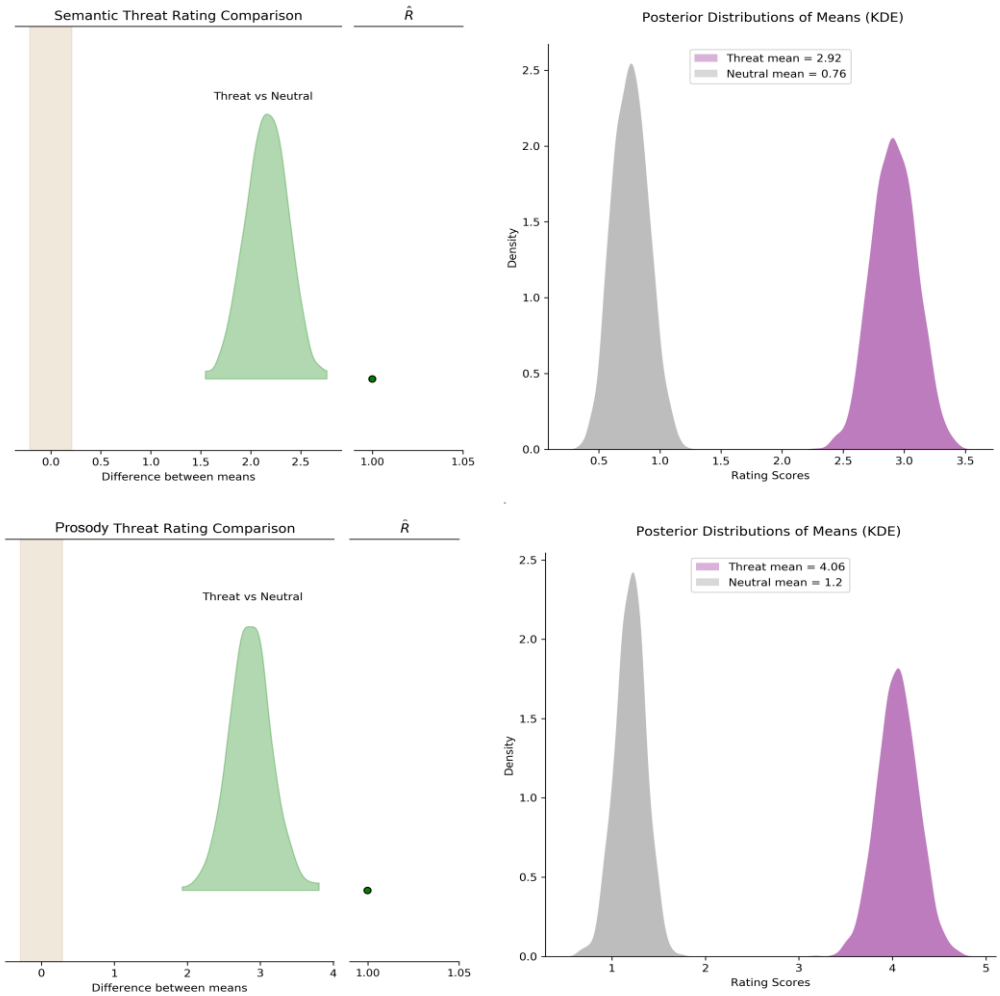

Figure 1. Left panels show a ridge plots of 90% HPDs in green, beige bars indicate 2SDs ROPEs. Right panels show kernel densities of posterior distributions of mean ratings. Upper panels: ratings of semantic threat. Lower panels: ratings of prosodic threat. $\hat{R}$ indicates the Gellman-Rubin convergence statistic.

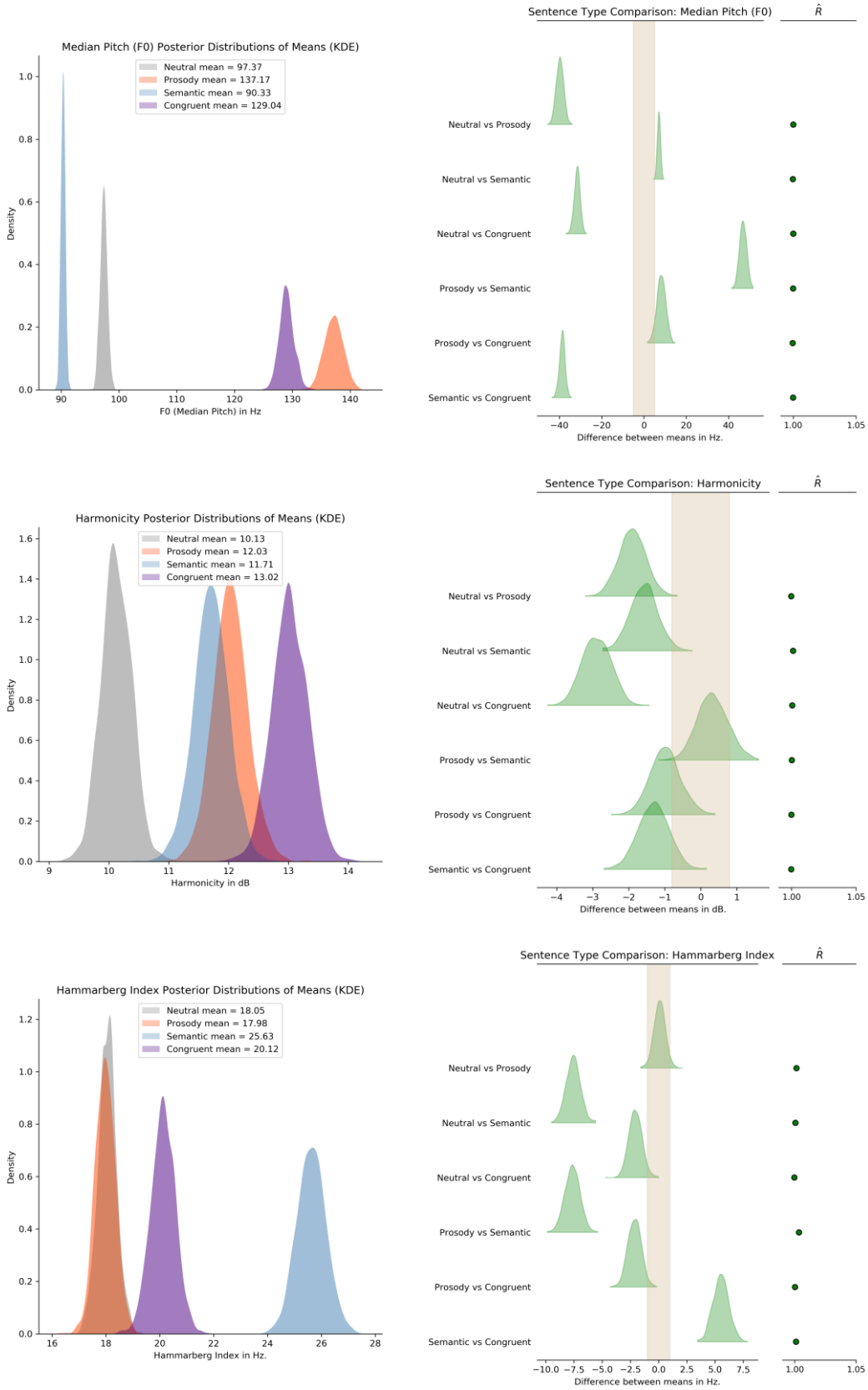

Figure 2. Plots shows results of BEST estimation on three prosody measures. Left: posterior distributions of each category mean. Right: 90% HPD distributions of difference between means in green, sand-coloured middle column shows the 2SDs ROPE. R̂ indicates the Gellman-Rubin convergence statistic.