JACEK KUDERA

SEMANTIC AND PROSODIC CONGRUENCE OF EMOTIONAL SPEECH IN DICHOTIC LISTENING

ORIGINAL RESEARCH PAPER
https://doi.org/10.17234/9789533791821.17

This study examines the semantic and prosodic congruence of verbal stimuli in a lexical decision task. It addresses the ear advantage in exposure to emotionally affected speech using a dichotic listening technique. The correct-incorrect answers ratio, coefficient of asymmetry and latency were compared to test the congruency effect (semantic positive vs. negative connotation of a spoken stimulus with the speaker's emotional state) and the ear advantage (speech directed to the left ear with murmur noise presented to the right ear and vice versa). The results show that the congruency effect strongly influences the response time and, as expected, incongruent stimuli require an increased processing effort compared to the congruent variant. Interestingly, this relationship was not confirmed on the basis of response accuracy. Contrary to initial expectations, the congruent condition required increased cognitive effort. These results partially contradict the previous reports in showing that congruence of prosodic and semantic features improves performance and reduces listeners' confusion.

Keywords: dichotic listening, lexical decision, emotional speech, semantic-prosodic congruence, Croatian

INTRODUCTION

It is commonly assumed that linguistic and affective prosody is processed primarily in the right hemisphere of the brain. However, several studies comparing activation patterns during exposure to linguistic and emotional prosody (based on similar acoustic features) suggest a predominant role of the left hemisphere for processing linguistic prosody and semantics and increased right hemisphere activity for processing affective prosody (Hirnstein et al., 2014; Kotz et al., 2006; Kreitewolf et al., 2014; Van der Haegen & Cai, 2019). What remains unknown is how affect and semantics interact when presented in congruent and incongruent ways. To address this question, a lexical decision task was designed using a dichotic listening technique with the following variables: the semantic and prosodic congruence of the stimuli (i.e., the previously assessed positive or negative connotation of a lexical item and the emotional state of a speaker pronouncing it) and the ear (left or right) to which the word was directed. In light of the functional hemispheric specialization (Grandjean, 2021; Grimshaw et al., 2003; Mildner, 1995, 2004) that implies the left ear advantage for processing emotional prosody and right ear advantage for processing semantics, this study attempts to cross-correlate the processing of semantics and emotional prosody by introducing the conditions of emotionally affected speech in a dichotic listening task. First, the semantic connotation of isolated Croatian lexemes was determined by means of a simple questionnaire, then the emotional state of a speaker was tested auditorily, and finally a lexical decision experiment with unilateral perception of spoken stimuli using a dichotic listening technique was employed.

The dichotic listening task was introduced over half a century ago (Broadbent, 1956) in a study on attention mechanisms and later popularized by Doreen Kimura's seminal research on cerebral dominance in speech perception (Kimura, 1961). This technique involves parallel exposure to different auditory stimuli in each ear. Regardless of the rapid development of imaging methods, the dichotic listening task is still a common, non-invasive, behavioral method for investigating perceptual differences between the left and right ear (Hugdahl, 2011). The approach is founded upon the principles of hemispheric functional lateralization, which can be illustrated by the representation of the auditory pathway. In short, an auditory signal received by the right ear arrives in the left hemisphere by a more direct route than a signal received by the left ear, which motivates the expected differences in right versus left ear advantage during exposure to speech presented to one ear and a masking sound directed to the other (Ahlsén, 2006; Mildner, 2008; Pisoni & Remez, 2008; Rouse, 2019; Stemmer & Whitaker, 2008).

Even though linguistic prosody functionally differs from emotional prosody, both are based on the variation of the same parameters, namely fundamental frequency, intensity, and duration. The classical lateralization dichotomy of the cerebral cortex has been investigated separately for emotional and linguistic prosody. This has led to the emergence of several theories, such as right hemisphere dominance for all types of prosody (Bryan, 1989; Kotz et al., 2006) or predominant right hemisphere activation in emotional prosody perception with a broader distribution of activation across the cerebral cortex for linguistic prosody (Cancelliere & Kertesz, 1990; Mildner, 2004). Some studies have suggested a continuum on a linguistic-emotional intonation scale concluding that intonation corresponds to increased left hemisphere activation, whereas affective prosody shifts the activation to the right hemisphere (Mayer et al., 2002; Wildgruber et al., 2004).

Related work

Previous studies have shown that functional lateralization allows us to expect different response patterns after directing a speech signal to one ear only (Ben-David et al., 2016; Bookheimer et al., 1997; Hirnstein et al., 2014; Van der Haegen & Cai, 2019). The introduced semantic versus prosodic dichotomy, which constitutes one of the experimental conditions in this study, is also based on findings that show predominant right hemisphere activation during perception of suprasegmental features and left hemisphere dominance in processing of semantics (Berl et al., 2010; Geschwind & Levitsky, 1968; Grandjean et al., 2006; Szaflarski et al., 2002).

Furthermore, functional dominance is subject to change as a function of the listener's emotional state (Papousek et al., 2011, 2012). In addition to behavioral tests, several neuroimaging and clinical studies have reported hemispheric specialization with regard to the processing of affective prosody and semantics. These studies have demonstrated that the processing of suprasegmental properties of speech signal is asymmetrical, with distinct activation patterns observed in the two cerebral hemispheres (Boemio et al., 2005; Borod et al., 2002; Erhan et al., 1998; Frühholz & Grandjean, 2013; Poeppel, 2003; Scott & Wise, 2004; Wiethoff et al., 2008; Wildgruber et al., 2005; Zatorre & Belin, 2001). Other reports provide evidence for a higher degree of distribution of neuronal activity across the hemispheres during the processing of linguistic prosody (Alba-Ferrara et al., 2011; Hervé et al., 2012; Mitchell & Ross, 2008). Different interpretations of this phenomenon have been inspired by the adaptation of signal analysis methodology. For instance, Zatorre et al. (2002) proposed that both hemispheres exhibit complementary skills and that the right hemisphere acts as a spectral analyzer with a long time window, which results in poor time resolution. In contrast, the left hemisphere acts as a better parser for temporal resolution with rather poor spectral discrimination. These findings have contributed to the callosal transfer and cue lateralization hypotheses (Grimshaw et al., 2003), which both suggest a predominant activation of the right hemisphere in perception of suprasegmental features of speech. Moreover, some studies have reported that the exposure to intonation contours and emotional prosody can result in the increased activation of similar brain structures (Raithel & Hielscher-Fastabend, 2004). Therefore, this study attempts to address the question of the ear advantage in the perception of affected speech by introducing the condition of semantic and prosodic congruence of unilaterally directed spoken stimuli. Thus, the semantic-prosodic congruence was defined as the emotion the speaker portrays when producing the stimuli. In the experiment, the incongruent condition includes the utterances of semantically positive words pronounced with sadness and semantically negative words uttered with happiness; whereas the congruent condition includes the set of semantically positive lexemes articulated with happiness.

AIMS AND PREMISES

The aim of this experiment was to test the effect of semantic-prosodic congruence and ear advantage on the latency and correctness of responses in a lexical decision task conducted using dichotic listening technique. The experiment addressed the question of whether emotionally affected speech (exhibiting a happiness-sadness polarity) affects latency and the ratio of correct to incorrect responses in a lexical decision task. The study also sought to test the ear advantage in exposure to semantically and prosodically congruent and incongruent speech.

Accordingly, three hypotheses were formulated. The first hypothesis concerns the relation between the response accuracy and the ear of exposure. It is assumed that due to functional lateralization, the dichotic projection of a stimulus (speech projected to one ear and a murmur noise, SNR = 0 dB, to the other) may influence the perception of emotionally affected items. It is therefore anticipated that more accurate responses will be observed following the presentation of stimuli directed to the right ear and within a congruent set of stimuli. This can be attributed to the right-ear advantage for language.

The second hypothesis concerns the ear advantage and the semantic-prosodic congruence of speech. Due to functional specialization and privilege of the right hemisphere in processing suprasegmental features of auditory stimulus, the unilateral perception of incongruent stimuli can evoke the left ear advantage operationalized as a laterality index.

The third hypothesis refers to the latency of the lexical decision influenced by the ear advantage and the congruency effect. It is hypothesized that longer latency may result from directing speech into the left ear and occurs after incongruent stimuli rather than the congruent ones. This is due to an increased cognitive effort required in processing semantic and prosodic mismatch, which is presumed to be more challenging than processing congruent stimuli.

METHODS

Having considered possible implications of the experimental design and the measurable effect of prosody processing (Mildner, 2013; Kotz et al., 2006), a simple behavioral, non-invasive test was designed. The dichotic perception of stimuli was employed to test the ear advantage in a lexical decision paradigm. The ear advantage index was computed with correctness of responses (calculated as coefficient of asymmetry: CAS) and latency (the response time measured until the participants' reaction). Before the lexical decision task was conducted, the isolated lexical items were evaluated in two pretests: the first one aimed to assess the semantics of isolated words; while the second pretest aimed to assess the emotional state of the speaker pronouncing the stimuli. Both factors were then used as test variables. For the sake of simplicity, the term 'prosody' is arbitrarily used throughout this paper to convey a speaker's emotional state as assessed in the second auditory pretest.

Laterality index

The coefficient of asymmetry (CAS) is a common quantitative laterality index used to estimate the privilege of one ear based on the number of correct responses given after the dichotic projection of stimuli (Bryden & Sprott, 1981). It is calculated according to the following equation: $CAS = (R - L)/(R + L)$, where R is the number of the correct responses given after directing the stimulus to the right ear and murmur to the left ear; whereas L marks the number of correct responses given after directing speech to the left ear and noise to the right ear. The CAS coefficient should therefore be interpreted as left ear advantage for $CAS < 0$, right ear advantage for $CAS > 0$, and no ear advantage for $CAS = 0$.

Pretest 1

The objective of the first pretest was to assess the semantic polarity (positive versus negative) of a set of isolated Croatian words. A total of 12 undergraduate students, native speakers of Croatian, participated in the first pretest. This group was excluded from participation in the subsequent phase. In a paper-and-pencil test, the subjects were given a word list consisting of 130 lexical items. The participants were instructed to mark each word as having a positive, negative, or neutral connotation. In the second pretest, only the lexemes that were marked as either 100 percent positive or negative were used.

Pretest 2

The second pretest was designed to assess a speaker's emotional state and allowed for the introduction of the congruency factor into the experimental design. To control for stimulus length, only the three-syllable tokens were selected from the set of qualified items from the first pretest. Then the pseudowords created according to the phonotactic rules of Croatian (Barić et al., 1997) were added to the existing lexemes. In total, 80 items were recorded by a trained male Croatian native speaker. The professional speaker was instructed to read the items twice conveying happiness and sadness respectively. The list was given to the speaker prior to the recording session for familiarization and to ensure effortless pronunciation. The stimuli were recorded in an acoustically controlled environment to an uncompressed format at 48 kHz sampling rate and 16-bit depth. The recorded samples were then used in the second pretest. In total, 15 subjects participated in the emotion judgment test. Their task was to rate the emotion of the speaker who pronounced the words in either a happy or a sad manner. In contrast to the first pretest, only the emotional state of a speaker was evaluated at this stage. The participants who took part in the second pretest were also excluded from further experimental procedure. Consequently, only the samples with 100 percent consistency in marking were selected for the final test to ensure clear disambiguation of the introduced variable.

Dichotic listening in lexical decision task

A lexical decision task in a dichotic listening technique was the third stage of the experimental procedure (Dobrić, 2012; Mildner, 1995). First, a brief interview was conducted with the participants. The questions included history of past injuries, recent illnesses, and potentially disqualifying factors such as diagnosed hearing disorders. A Weber test was then performed with a 440 Hz tuning fork to ensure intact air and bone conduction and equal sound perception in both ears. After the Weber test, the extended version of Croatian handedness and footedness questionnaire was given to participants. The survey consisted of 22 questions about daily activities performed predominantly with the right/ left hand or leg. The questionnaire also included a section on right-handedness, left-handedness, or ambidexterity in close family members (Mildner, 2000). Subjects were then introduced to a lexical decision task that took place in the sound-attenuated booth. Instructions were read aloud and projected on a screen in their native language (Mildner & Golubić 2003). The participants were asked to listen to the stimuli and decide whether the tokens they heard were real Croatian words. They were instructed to respond by pressing the word or non-word button for each item as accurately and quickly as possible. To familiarize the participants with the procedure and the experimental interface, the test was preceded by a short training session with

12 tokens in each condition. The data obtained from the training part were excluded from the analysis. After the trial stage, participants were given an opportunity to clarify the procedure and ask questions. They were then asked to sign the informed consent form. They were also instructed that they could leave the booth at any time in case of fatigue, stress or any test-related discomfort. The studio session lasted for approximately 30 minutes. The lexical decision tests consisted of 20 meaningful words and 11 logatomes. Each subject was exposed to 372 randomized stimuli (31 tokens x 2 projection sides x 2 emotions x 3 repetitions) and 360 responses were analyzed from each session. The dichotic perception was achieved by projecting a murmur noise (SNR = 0 dB) to one ear and stimuli to the other in a randomized fashion via AKG-K77 Perception Studio headset. The binaurally projected fixation sound was played after each trial. The E-Prime software was used for stimulus presentation, randomization and latency tracking.

Participants

In total, 39 participants (35 males and 4 females), aged 22-27, students of phonetics at the University of Zagreb took part in the final studio experiment. They received a credit point for their participation in the experiments as a part of their study program. None of the participants had undergone ear or throat surgery, nor had they been diagnosed with alexithymia (inability to recognize emotions), which would have been a disqualifying factor for the audio test. None of the subjects reported a recent history of ear, nose, or throat disease, and all participants were qualified for the studio experiment based on the Weber test results.

Data analysis

The basic descriptive statistics were calculated with the Kolmogorov-Smirnov test. Then one-way and two-way repeated measure ANOVAs were conducted to answer the research questions. The classical threshold of statistical significance was set at $\alpha = 0.05$ and the values within the range $0.05 < p < 0.1$ were interpreted as a statistical tendency. The trials with stimulus onset delay ≥ 10 ms caused by the hardware delay were discarded from further analysis (1.47% of data). The logatomes were excluded from the analyses. The latency was averaged across all projections of a stimulus type.

RESULTS

The results of the Kolmogorov-Smirnov test showed that projections to the right ear were close to a Gaussian distribution (see Table 1). Therefore, a skewness analysis was mandatory. It showed that in all cases skewness ranged between -2 and 2, allowing for the parametric testing conducted in the next step.

Table 1. Basic descriptive statistics: Correct responses and CAS

	Ear	Emotion	Semantics	*M*	*Me*	*SD*	*Sk*	*Kurt*	*Min*	*Max*	*K-S*	*p*
Correct responses	Left	-	-	26.74	27	2.06	-0.36	-0.98	23	30	0.17	0.009
			+	26.08	27	2.44	-0.65	-0.28	20	30	0.21	<0.001
		+	-	28.03	29	2.01	-1.42	2.62	21	30	0.23	<0.001
			+	25.00	26	2.77	-0.46	-1.18	20	29	0.23	<0.001
	Right	-	-	26.82	27	2.22	-0.84	0.20	21	30	0.17	0.005
			+	26.41	26	1.94	-0.28	-0.57	22	30	0.13	0.116
		+	-	28.13	29	1.59	-0.59	-0.38	24	30	0.22	<0.001
			+	25.36	26	2.49	-0.23	-0.63	20	30	0.16	0.018
CAS		-	-	0.11	0	3.42	0.05	2.79	-10.64	9.80	0.21	<0.001
			+	0.72	0	3.10	0.82	2.21	-5.45	11.11	0.18	0.002
		+	-	0.24	0	3.81	0.90	4.39	-9.43	14.29	0.17	0.009
			+	0.78	0	3.15	-0.10	-0.01	-6.67	7.14	0.16	0.011
Latency	L	-	-	2661.59	2656.95	151.77	0.14	-0.10	2341.88	3027.75	0.09	0.200
			+	2676.25	2688.33	155.46	-0.05	-0.47	2379.52	3011.58	0.11	0.200
		+	-	2634.40	2625.90	162.94	-0.10	-0.52	2268.38	2922.93	0.08	0.200
			+	2619.83	2609.50	147.39	-0.05	-0.34	2270.67	2903.54	0.08	0.200
	P	-	-	2628.52	2670.97	169.41	-0.19	-0.41	2244.17	2978.29	0.11	0.200
			+	2660.73	2667.57	158.58	-0.16	0.05	2280.95	2976.29	0.08	0.200
		+	-	2611.16	2637.62	188.72	0.28	1.00	2210.20	3154.61	0.11	0.200
			+	2601.42	2605.00	181.27	0.39	-0.24	2283.14	3036.83	0.08	0.200

M – mean; Me – median; SD – standard deviation; Sk – skewness; Kurt – kurtosis; Min – minimal value; Max – maximal value; K-S – Kolmogorov-Smirnov test; p – significance

Table 2. Basic descriptive statistics: Reaction times [in ms]

Ear	Emotion	Semantics	*M*	*Me*	*SD*	*Sk*	*Kurt*	*Min*	*Max*	*K-S*	*P*
Left	-	-	2661.59	2656.95	151.77	0.14	-0.10	2341.88	3027.75	0.09	0.200
		+	2676.25	2688.33	155.46	-0.05	-0.47	2379.52	3011.58	0.11	0.200
	+	-	2634.40	2625.90	162.94	-0.10	-0.52	2268.38	2922.93	0.08	0.200
		+	2619.83	2609.50	147.39	-0.05	-0.34	2270.67	2903.54	0.08	0.200
Right	-	-	2628.52	2670.97	169.41	-0.19	-0.41	2244.17	2978.29	0.11	0.200
		+	2660.73	2667.57	158.58	-0.16	0.05	2280.95	2976.29	0.08	0.200
	+	-	2611.16	2937.62	188.72	0.28	1.00	2210.20	3154.61	0.11	0.200
		+	2601.42	2605.00	181.27	0.39	-0.24	2283.14	3036.83	0.08	0.200

M – mean; Me – median; SD – standard deviation; Sk – skewness; Kurt – kurtosis; Min – minimal value; Max – maximal value; K-S – Kolmogorov-Smirnov test; p – significance

To test the first hypothesis and verify the influence of the ear (left vs. right) and congruence (positive semantics – positive prosody; positive semantics – negative prosody; negative semantics – negative prosody; negative semantics – positive prosody) on the correctness of responses a two-way analysis of variance was conducted (see Table 3). The main effect of the ear (left vs. right) was observed only on the level of statistical tendency ($F(1, 38) = 3.47$; $p = 0.070$; $\eta^2 = 0.39$), with a higher number of correct responses given after directing stimuli to the right ear. However, the magnitude of the effect was low.

Table 3. Basic statistics: Correct responses

Ear	Prosody	Semantics	M	SE
Left	-	-	26.74	0.33
	-	+	26.08	0.39
	+	-	28.03	0.32
	+	+	25.00	0.44
	all conditions		26.57	0.25
Right	-	-	26.82	0.36
	-	+	26.41	0.31
	+	-	28.13	0.26
	+	+	25.36	0.40
	all conditions		26.68	0.24
Both	-	-	26.78	0.31
	-	+	26.24	0.33
	+	-	28.08	0.24
	+	+	25.18	0.40

The main effect of the semantic-prosodic congruence influencing the correctness of the responses was statistically significant and strong ($F(1, 97) = 24.04$; $p < 0.001$; $\eta^2 = 0.39$). Hence, a post-hoc analysis with the Šidák test was conducted. The highest number of correct responses was observed in the positive prosody – negative semantics condition. This variant differed significantly from all other conditions. The smallest number of correct responses was given in the congruent positive prosody – positive semantics condition. This condition was also significantly different from the rest of alternatives, that is, the participants provided the fewest correct responses in this variant.

The difference between the negative prosody – negative semantics and negative prosody – positive semantics conditions did not reach the threshold of statistical significance. Then, the analysis of the interaction effects was performed. For both ears, the highest number of correct responses was observed after the positive prosody – negative semantics set of stimuli, while the lowest number of correct responses was observed in the positive prosody – positive semantics cluster. Negative congruence as well as positive prosody – negative semantics did not differ even at the level of statistical tendency (see Figure 1).

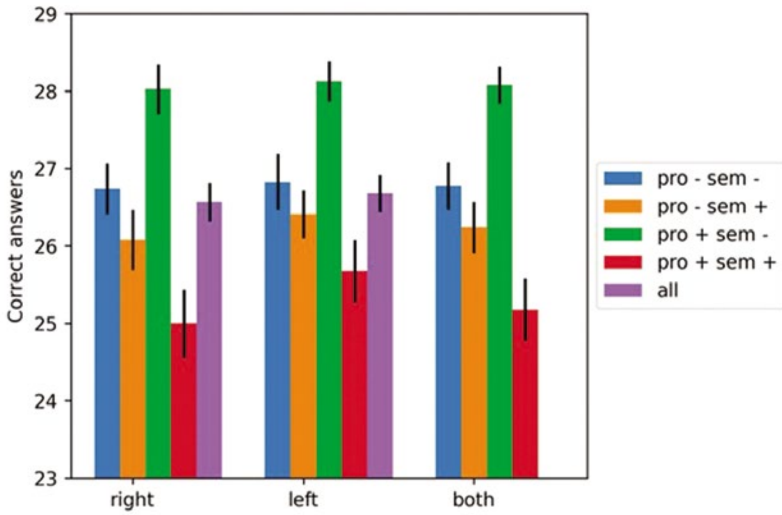

Figure 1. Correct answers in each condition per channel

To test the second hypothesis regarding ear privilege in perceiving four variants of congruence, a one-way ANOVA was conducted. The effect did not reach the threshold of statistical significance ($F(3, 114) = 0.36$; $p = 0.785$, $\eta^2 = 0.01$). Hence, on the basis of the gathered data, it cannot be concluded that the congruence of semantics and prosody affects the ear privilege in a dichotic listening setup.

To test the third hypothesis regarding the relationship between reaction times and ear (left vs. right) as well as stimulus congruency (a total of four congruency variants), a two-way repeated measures ANOVA was performed. Basic descriptive statistics are shown in Table 4.

Table 4. Basic statistics: Reaction times [in ms]

Ear	Prosody	Semantics	M	SE
Left	-	-	2661.59	24.30
	-	+	2676.25	24.89
	+	-	2634.40	26.09
	+	+	2619.83	23.60
	all conditions		2648.02	23.64
Right	-	-	2628.52	27.13
	-	+	2660.73	25.39
	+	-	2611.16	30.22
	+	+	2625.46	26.62
	all conditions		2625.46	26.62
Both	-	-	2645.06	25.16
	-	+	2668.49	24.41
	+	-	2622.78	27.26
	+	+	2610.63	25.41

The main ear effect (left vs. right) was statistically significant ($F(1, 38) = 9.34$; $p = 0.004$, $\eta^2 = 0.20$). The analysis showed that a longer processing time was required after the projection of stimuli to the left ear. The congruency effect was reported ($F(3, 114) = 14.30$; $p < 0.001$; $\eta2 = 0.27$), therefore a post-hoc Šidák analyses was performed to test the differences across the conditions. The longest latency was measured after the negative prosody – positive semantics variant; whereas the shortest response time was observed in the positive prosody – positive semantics condition. Interestingly, a difference between the other congruent variant, that is, negative prosody – negative semantics, was significantly different, but not when contrasted with the incongruent (positive prosody – negative semantics) condition. Additionally, the analysis of simple effect showed that stimuli in the negative prosody and positive semantics variant caused the longest processing times when projected to the left ear. This finding was significantly different from the positive prosody – negative semantics and positive prosody – positive semantics variants. Furthermore, a difference between the congruent variants directed to the right ear was observed, with shorter latency in the congruent positive condition. For the right ear, only one incongruent cluster was significantly different from the other variants, that is, negative prosody – positive semantics. This condition caused the longest latency (see Figure 2).

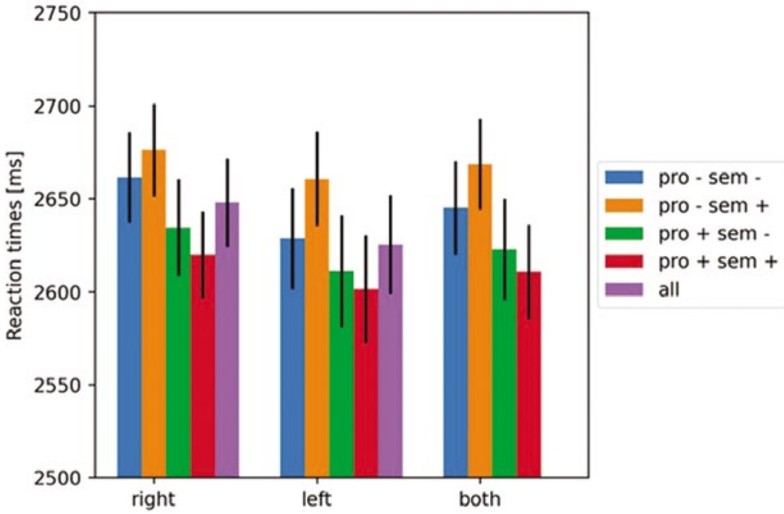

Figure 2. Reaction times in each condition per channel

DISCUSSION

The purpose of this study was to investigate the influence of semantic and prosodic congruence of emotionally affected speech on latency and correctness of responses in a lexical decision task. The non-invasive dichotic listening experiment was conducted to assess the ear advantage for the perception of emotional stimuli and the semantic congruence of lexical items with a speaker's emotional state. The analyses were based on latency and coefficient of asymmetry, computed as the correct-incorrect responses ratio. Similar to previous studies using dichotic listening technique, the right ear advantage for speech processing was observed (Bless et al., 2015; Hugdahl, 2003). The main effect of the ear advantage and correct response rate was reported at the level of statistical tendency. This finding seems to be consistent with a common view on functional lateralization and spoken language processing. Further data collection may increase the effect magnitude and turn the tendency level into significance. The main effect of stimulus congruency on response accuracy was confirmed with the following dependency: the incongruent variant (positive prosody – negative semantics) elicited the largest number of correct responses.

Contrary to initial expectations, the congruent variant (positive prosody – positive semantics) appeared to be the most difficult for the subjects. These results partially contradict previous reports showing that congruence of affective prosody and semantics improves performance and reduces listeners' confusion. Furthermore, these results suggest that the congruency effect does not reduce the speech processing effort assessed on the basis of the ratio of correct to incorrect responses. In the light of the gathered data, the hypothesis regarding the ear advantage and the semantic – prosodic congruence was rejected. The third hypothesis, which related to the influence of stimulus congruency and direction on latency in the lexical decision task, was confirmed. The main effects of congruency and response correctness, as well as the effect of stimulus direction (left ear vs. right ear), were statistically significant and of high magnitude. The analyzed data showed that emotionally affected speech directed to the left ear elicited longer reaction times. In line with initial expectations, the incongruent variant, that is, negative prosody – positive semantics, required the longest processing time and differed significantly from the other conditions.

Another argument in favor of the third hypothesis was provided by the latency data measured in the cluster of congruent samples. The positive prosody – positive semantics set of stimuli required the shortest processing time, but also yielded the fewest correct responses. It can be speculated that the congruent condition may have seemed like an easier task, hence less effort was put into deciding upon the matching samples. This pattern can be attributed to faster reaction times, regardless of the cognitive effort required to make an accurate judgment. The observed outcome may have been influenced by the participants' focus on the spoken stimulus and their simultaneous decoding of the emotional state of the speaker. As a result, encountering the natural-sounding happiness associated with positive meaning may have caused more confusion and therefore resulted in poor performance as measured by the response accuracy.

CONCLUSIONS AND OUTLOOK

The results of this study confirmed the canonical left ear advantage for processing emotional prosody. Furthermore, the premise of increased cognitive effort associated with incongruent stimuli appeared to be accurate. The gathered data showed that latency was increased after exposure to the words with positive semantic connotation pronounced by a speaker conveying a sad emotional state and after the lexemes with negative meaning uttered with happiness. The right ear advantage for language potentially explains the longer reaction times observed after stimuli directed to the left ear. The evidence against the preliminary assumptions was provided by the analyses of the congruent and positive cluster. Exposure to this type of stimulus elicited the shortest latency, but also the smallest number of correct responses. The results of testing the hypothesis regarding the lateralization index and stimulus congruence did not reach the threshold of statistical significance, so no valid conclusions can be drawn based on the available data. The magnitude of the reported effects may increase with further data collection.

The post hoc analyses the reported discrepancies between latency and response accuracymay suggest that these two measures, commonly used in lexical decision tasks, should be treated with special attention. Therefore, these measures should be analyzed separately in lexical decision tasks, especially in experiments with dichotic presentation of the spoken stimuli. The in-depth error analyses may also reveal interesting patterns of dichotic perception of emotionally affected speech. Another methodological criticism may concern the sample of participants tested in the final phase of the experiment. The tested group was not gender balanced, but the available literature does not provide evidence of significant gender disproportions that might affect the dichotic perception of emotionally affected speech to a greater extent than within-group, individual variations. Furthermore, the Weber test is not a true substitute for a thorough hearing assessment, although it is effective in identifying hearing asymmetries, which is the primary concern in this study. Future studies should also supplement such behavioral methods with imaging techniques that would allow tracking functional activity in dichotic perception of emotionally affected speech with greater temporal resolution.

ACKNOWLEDGMENTS

Special thanks to Jordan Bićanić for his help in the recording studio.

DATA AVAILABILITY

The experimental data are publicly available in the following Open Science Framework repository (https://osf.io/u276n/).

REFERENCES

Ahlsén, E. (2006). *Introduction to Neurolinguistics.* Amsterdam – Philadelphia: John Benjamins Publishing.

Alba-Ferrara, L., Hausmann, M., Mitchell, R. L., & Weis, S. (2011). The neural correlates of emotional prosody comprehension: disentangling simple from complex emotion, *PloS one*, 6(12), e28701.332.

Barić, E., Lončarić, M., Malić, D., Pavečić, S., Peti, M., Zečević, V., & Znika, M. (1997). *Hrvatska gramatika.* Zagreb: Školska knjiga.

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & van Lieshout, P. H. (2016). Prosody and semantics are separate but not separable channels in the perception of emotional speech: Test for rating of emotions in speech. *Journal of Speech, Language, and Hearing Research*, 59(1), 72–89.

Berl, M. M., Duke, E. S., Mayo, J., Rosenberger, L. R., Moore, E. N., VanMeter, J., Ratner, N. B., Vaidya, C. J., & Gaillard, W. D. (2010). Functional anatomy of listening and reading comprehension during development, *Brain and language*, 114(2), 115–125.

Bless, J. J., Westerhausen, R., Torkildsen, J. v. K., Gudmundsen, M., Kompus, K., & Hugdahl, K. (2015). Laterality across languages: Results from a global dichotic listening study using a smartphone application. *Brain and Cognition*, 20(4), 434–452.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature neuroscience*, 8(3), 389–395.

Bookheimer, S., Zeffiro, T., Blaxton, T., Malow, B., Gaillard, W., Sato, S., Kufta, C., Fedio, P., & Theodore, W. (1997). A direct comparison of pet activation and electrocortical stimulation mapping for language localization. *Neurology*, 48(4), 1056–1065.

Borod, J. C., Bloom, R. L., Brickman, A. M., Nakhutina, L., & Curko, E. A. (2002). Emotional processing deficits in individuals with unilateral brain damage. *Applied neuropsychology*, 9(1), 2–36.

Broadbent, D. E. (1956). Successive responses to simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 8(4), 145–152.

Bryan, K. L. (1989). Language prosody and the right hemisphere. *Aphasiology*, 3(4), 285–299.

Bryden, M., & Sprott, D. (1981). Statistical determination of degree of laterality. *Neropsychologia*, 19(4), 571–581.

Cancelliere, A. E., & Kertesz, A. (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and Cognition*, 13(2), 133–147.

Dobrić, A. (2012). Lateralizacija jezičnih funkcija u odraslih višejezičnih govornika mjerena dihotičkim slušanjem. *Govor*, 29(2), 121–148.

Erhan, H., Borod, J. C., Tenke, C. E., & Bruder, G. E. (1998). Identification of emotion in a dichotic listening task: event-related brain potential and behavioral findings. *Brain and Cognition*, 37(2), 286–307.

Frühholz, S., & Grandjean, D. (2013). Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neuroscience & Biobehavioral Reviews*, 37(10), 2847–2855.

Geschwind, N., & Levitsky, W. (1968). Human brain: left-right asymmetries in temporal speech region. *Science*, 161(3837), 186–187.

Grandjean, D. (2021). Brain networks of emotional prosody processing. *Emotion Review*, 13(1), 34–43.

Grandjean, D., Bänziger, T., & Scherer, K. R. (2006). Intonation as an interface between language and affect. *Progress in brain research*, 156, 235–247.

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.

Hervé, P.-Y., Razafimandimby, A., Vigneau, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2012). Disentangling the brain networks supporting affective speech comprehension. *NeuroImage*, 61(4), 1255–1267.

Hirnstein, M., Hugdahl, K., & Hausmann, M. (2014). How brain asymmetry relates to performance–a large-scale dichotic listening study. *Frontiers in psychology*, 4, 997.

Hugdahl, K. (2003). Dichotic Listening in the study of auditory laterality. In K. Hugdahl & R. J. Davidson (Eds.), *The asymmetrical brain* (pp. 441–475). Cambridge: MIT Press.

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

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 15(3), 166.

Kotz, S. A., Meyer, M., & Paulmann, S. (2006). Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design. *Progress in brain research*, 156, 285–294.

Kreitewolf, J., Friederici, A. D., & Kriegstein, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage*, 102, 332–344.

Mayer, J., Wildgruber, D., Riecker, A., Dogil, G., Ackermann, H., & Grodd, W. (2002). Prosody production and perception: converging evidence from fmri studies. *Speech Prosody*, *2002*.

Mildner, V. (1995). Reprezentacija jezičnih i govornih procesa u mozgu, *Govor*, 12(2), 85–109

Mildner, V. (2000). Is the hand to speech what speech is to the hand? *Brain and Cognition*, 43, 345–349.

Mildner, V. (2004). Hemispheric asymmetry for linguistic prosody: A study of stress perception in Croatian. *Brain and Cognition*, 55(2), 358–361.

Mildner, V. (2008). *The cognitive neuroscience of human communication.* New York: Taylor & Francis.

Mildner, V. (2013). Experimental and quasi-experimental research in clinical linguistics and phonetics. In N. Müller & M. J. Ball (Eds.), *Research Methods in Clinical Linguistics and Phonetics: A practical guide.* (pp. 28–47). Wiley-Blackwell.

Mildner, V., & Golubić, A. (2003). Funkcionalna mozgovna asimetrija pri obradi jezičnih podataka na materinskom i stranom jeziku. *Govor*, 20 (1–2), 277–288.

Mitchell, R. L., & Ross, E. D. (2008). fMRI evidence for the effect of verbal complexity on lateralisation of the neural response associated with decoding prosodic emotion. *Neuropsychologia*, 46(12), 2880–2887.

Papousek, I., Murhammer, D., & Schulter, G. (2011). Intra-and interindividual differences in lateralized cognitive performance and asymmetrical eeg activity in the frontal cortex. *Brain and Cognition*, 75(3), 225–231.

Papousek, I., Reiser, E. M., Weber, B., Freudenthaler, H. H., & Schulter, G. (2012). Frontal brain asymmetry and affective flexibility in an emotional contagion paradigm. *Psychophysiology*, 49(4), 489–498.

Pisoni, D. B., & Remez, R. E. (2008). *The handbook of speech perception*. John Wiley & Sons.

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.

Raithel, V., & Hielscher-Fastabend, M. (2004). Emotional and linguistic perception of prosody. *Folia Phoniatrica et Logopaedica*, 56(1), 7–13.

Rouse, M. H. (2019). *Neuroanatomy for speech-language pathology and audiology*. Burlington: Jones & Bartlett Learning.

Scott, S. K., & Wise, R. J. (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition*, 92(1-2), 13–45.

Stemmer, B., & Whitaker, H. A. (2008). *Handbook of the Neuroscience of Language*. London: Academic Press.

Szaflarski, J. P., Binder, J. R., Possing, E. T., McKiernan, K. A., Ward, B. D., & Hammeke, T. A. (2002). Language lateralization in left-handed and ambidextrous people: fmri dana. *Neurology*, 59(2), 238–244.

Van der Haegen, L., & Cai, Q. (2019). Lateralization of language. In G. I. De Zubicaray & N. O. Schiller (Eds.), *The Oxford Handbook of Neurolinguistics*. New York: Oxford University Press.

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody – Influence of acoustic parameters and arousal. *Neuroimage*, 39(2), 885–893.

Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., & Ackermann, H. (2004). Distinct Frontal Regions Subserve Evaluation of Linguistic and Emotional Aspects of Speech Intonation. *Cerebral Cortex*, 14(12), 1384–1389.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fmri. *Neuroimage*, 24(4), 1233–1241.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral cortex*, 11(10), 946–953.

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.