Emotional Language Processing in Bilingualism: Subjective Affect and Prosodic Markers in Simultaneous Interpreting

İki Dillilikte Duygusal Dil İşlemleme:
Andaş Çeviride Öznel Duygulanım ve Bürünsel Belirteçler

Alper Kumcu*
Hacettepe University

Abstract
Emotion, as part of the overall sensorimotor, introspective, and affective system, is an essential part of language comprehension within the framework of embodied semantics. As emotional state influences semantic and syntactic processing, emotional language processing has been shown to modulate mood as well. The reciprocal relationship between language and emotion has also been informative in bilingualism. Here we take a relatively underresearched type of bilingual processing, simultaneous interpreting, as a case of extreme bilingualism and investigate the effect of emotional language rendering in the L1 on subjective affect and prosodic markers of L2 output. 18 trainee interpreters were asked to simultaneously interpret three speeches in Turkish that varied in emotionality, valence (negative, neutral, and positive), and difficulty in English. Responses to emotional language processing were analysed based on participants' self-reported positive and negative affect using the Positive and Negative Affect Schedule (PANAS) and three prosodic parameters (intensity, pitch, and fluency). Results showed that interpreting emotionally negative speech increased negative affect, whereas interpreting emotionally positive speech did not modify positive affect. Intensity generally reflected cognitive load. Pitch and fluency, in particular, were more sensitive to changes in the valence of the source speech.
Keywords: emotional language, emotional load, cognitive load, valence, bilingualism, simultaneous interpreting, prosodic parameters

Öz
Duygu; genel duyumotor, içgörü ve duygulanım sisteminin bir parçası olarak, bedenlenmiş anlambilim çerçevesinde dil anlamanın temel bir parçasıdır. Çalışmalar duygudurumun anlamsal ve sözdizimsel işlemlemeyi etkilediği gibi, duygusal dil işlemlemenin de duygudurumu etkilediğini göstermiştir. Dil ve duygu arasındaki karşılıklı ilişki, iki dillilik açısından da açıklayıcıdır. Bu çalışma, görece az çalışılan bir iki dillilik türü olan andaş çeviriyi bir tür "uç iki dillilik" olarak ele almakta ve ana dilde (L1) duygusal dil işlemlemenin öznel duygulanım ve ikinci dilde (L2) bürünsel belirteçler üzerindeki etkisini ortaya koymaktadır. Çalışmada 18 katılımcıdan duygusallık, duygusal değer (olumsuz, nötr ve olumlu) ve zorluk açısından farklılık gösteren üç Türkçe konuşmayı eş zamanlı olarak İngilizceye çevirmeleri istenmiştir. Duygusal dil işlemleme, katılımcıların Olumlu ve Olumsuz Duygulanım Ölçeği (PANAS) ile bildirdikleri olumlu ve olumsuz duygulanım puanları ile üç bürünsel parametre (yoğunluk, tizlik ve akıcılık) temelinde analiz edilmiştir. Sonuçlar, duygusal açıdan olumsuz konuşmanın çevirisinden sonra olumsuz duygulanımın arttığını ancak duygusal açıdan olumlu konuşmanın olumlu duygulanımı etkilemediğini göstermiştir. Seste yoğunluk genel olarak bilişsel yükü yansıtırken tizlik ve özellikle akıcılık kaynak konuşmanın duygusal değerindeki değişikliklere karşı daha hassastır.
Anahtar sözcükler: duygusal dil, duygusal yük, bilişsel yük, duygusal değer, iki dillilik, andaş çeviri, bürünsel parametreler

*Asst. Prof. Dr., Dept of Translation and Interpreting, Hacettepe University
ORCID# 0000-0003-0844-3562; alperkumcu@gmail.com; https://doi.org/10.47777/cankujhss
CUJHSS (ISSN 1309-6761) Dec 2023; 17/2, 194-211. Received Aug 5, 2023; Accepted Oct 17, 2023

Introduction

Previous research has shown that the relationship between language and emotions is much more intertwined than previously thought. That is, the role of language in the experience of emotions goes beyond simply describing feelings with words. Not only does language categorise and construct emotional experiences (Lindquist et al., 2015), but several empirical studies have shown that language can also evoke rather strong emotions (Habermas, 2018). Moreover, emotions evoked by language are simulated as part of comprehension and in turn, can influence judgements and understanding (Bohn-Gettler, 2019; Havas et al., 2007; Johnson & Tversky, 1983). How emotional language, which is language typically composed of emotion-laden words, sentences and text or language that communicates emotions, is processed has also shed much light on bilingual processing. It is generally assumed that emotionality is attenuated in L2, leading to so-called "disembodied" language processing (Pavlenko, 2012). There is indeed much empirical evidence to support this claim, showing reduced physiological responses to emotionally charged words when presented in the non-native second language (L2) compared to the native language (L2) (Harris, 2004; Harris et al., 2003; Heyrani et al., 2022). However, there is also counter-evidence showing similar patterns of emotional activation in both L1 and L2 (e.g., Conrad et al., 2011).

That said, comparatively little attention has been paid to emotional language processing in language-based performative tasks such as simultaneous interpreting (SI), where the language user processes L1 and L2 almost simultaneously with the comprehension and production dimensions. SI is a form of interpreting where the interpreter listens to the speakers' source speech in one language (L1 or L2) and renders it into another language (L1 or L2) in real time. Compared to other language tasks and interpreting modalities, SI is usually described as "extreme bilingualism" (e.g., Nour et al., 2020), because SI has very high parallel processing and executive control demands. Interpreters perform along with the accumulating cognitive load against limited internal resources and under building time pressure (Korpal, 2021; Mellinger & Hanson, 2019). In this regard, SI has been regarded as a window into bilingual processing (see Dong & Li, 2020). Moreover, it is fair to argue that SI is an emotionally taxing activity even without emotional language. Previous studies have identified several external stressors in SI, ranging from speaking rate (Korpal, 2016) to prolonged turns (Klonowicz, 1990; Moser-Mercer et al., 1998), from textual complexity to lexical search under time pressure (Gumul, 2021), from challenging techno-social work environments (Korpal, 2021; Kumcu, 2020; Kurz, 2002) to booth discomfort (Mackintosh, 2003).

Effect of emotional and cognitive load on voice

When an individual is exposed to a stressor, i.e., an upsetting external stimulus or event, the body reacts quickly by gearing up for action and initiating physical responses such as increased muscle tension, increased heart and respiratory rates, or perceived changes in temperature, among others (Selye, 1956). Changes in the vocal output are one of these physiological changes observed during shifting emotional states. Although anger and happiness are relatively more recognizable (Filippa et al., 2022), several studies have repeatedly shown that a wide range of emotions are systematically embodied in the human voice, allowing individuals to communicate and/or recognize emotional experiences.

Measures of vocal intensity, fundamental frequency (F0) and pauses are the most commonly used cues to study emotional prosody (Scherer, 1986a). F0 is derived from the rate of vibration of the speaker's vocal cords and is calculated as the average number of vibrations per second in Hertz (Hz.). F0, in turn, is perceived by listeners as pitch. The F0 of an average male voice is 80-150 Hz and that of an average female voice is 150-300 Hz. Intensity, on the other hand, is the rate at which vocal energy is transmitted by the voice wave and is defined in decibels (dB). A normal adult voice intensity is between 60 and 70 dB.

Overall, higher arousal emotions such as happiness or joy are associated with higher mean intensity, higher mean F0 and shorter pauses, whereas lower arousal emotions such as sadness are associated with lower mean intensity, lower mean F0 and longer pauses compared to neutral speech (Juslin & Laukka, 2003; Tisljár-Szabó & Pléh, 2014). However, subtle differences between languages and individual differences due to voice type (i.e., lax/tense voice type) have been documented (Tamuri, 2014). In the context of emotional prosody studies, there is also evidence that vocal cues are reliable indicators of psychiatric disorders. For example, adults with elevated depressive symptoms showed longer total pause duration and shorter total speech duration (Albuquerque et al., 2021), while depressed patients showed greater pitch variability, fewer pauses, and faster speech rate as they responded to treatment (Mundt et al., 2007). Similarly, schizophrenic patients have been shown to speak with less pitch variability (i.e., flat speech) and more pauses compared to healthy controls (Martínez-Sánchez et al., 2015). On the other hand, increased F0 has been reported as an indicator of social anxiety disorder, but only in males (Weeks et al., 2012).

In addition to emotionality and emotional load, difficulty and cognitive load are also reflected in voice (Van Puyvelde et al., 2018). Overall, higher F0 and higher intensity have been evidenced as a result of increased task complexity and thus, difficulty. However, it is important to note that there are large methodological differences between the studies showing the effect of cognitive load on prosody. For example, in Griffin and Williams (1987), participants were asked to perform psychomotor and dichotic listening tasks in an aviation environment. Similarly, Huttunen et al. (2011) conducted a flight simulation study and showed that mean F0 increased by 12 Hz and mean intensity by 1.5 dB during the most cognitively intensive phases of simulator flight. In Rothkrantz et al. (2004), on the other hand, intensity and F0 increased as a function of Stroop test difficulty.

Finally, some studies compare emotional and cognitive load in the same paradigm. When presented with tasks that induce cognitive and emotional stress, males showed higher F0 values under cognitive rather than emotional load (Tolkmitt & Scherer, 1986). However, in another study (Scherer et al., 2002), participants speaking German, English and French showed an increased F0 as a result of emotional rather than cognitive load during several computer-based tasks.

In summary, studies show that the mood of language users leaves its mark on their production, but the direction of the effect is not straightforward due to the different languages and tasks involved.

Emotionality and stress in interpreting

Emotional processing in interpreting has been investigated through a wide array of indicators and measures (see Gieshoff, Lehr, and Hunziker Heeb 2021) from the self-reports (Gumul, 2021) to galvanic skin conductance (Korpal & Jasielska, 2019; Kurz, 2002), from heart rate (Klonowicz, 1994) to blood pressure (Roziner & Shlesinger, 2010) and cortisol levels (Moser-Mercer, 2005).

In one important study combining galvanic skin response and self-reported affect measures (Korpal & Jasielska, 2019), a group of interpreters were asked to simultaneously interpret a speech with negative valence and an emotionally neutral speech from L1 (Polish) to L2 (English) while their galvanic skin responses, as a marker of emotional arousal, were measured. Participants also reported their momentary emotional states through the Positive and Negative Affect Schedule (PANAS) before the first interpretation (as baseline) and immediately after the interpretation of neutral and negative-valence recordings. Results showed that the mean number of galvanic skin responses consistently increased across conditions from baseline to negative-valence speech. Further, participants reported higher negative affect scores immediately after interpreting negative-valence speech compared to neutral speech and the baseline. Results suggested that interpreters indeed empathise with the speaker experiencing negative emotional states by emotionally mimicking her/him which is reflected in both objective and subjective responses. Results also suggest that stress response to the emotional material in SI is activated more or less automatically and that the interpreter has almost no control over it, which makes the negative speech as a stressor particularly challenging compared to other, work-related factors.

It is no surprise that examination of emotional processing and stress response in interpreting based on acoustic markers is rare (see Rojo López et al. 2021) given that phonological studies, in general, are underrepresented in interpreting studies with few exceptions (Ahrens, 2005a, 2005b; Darò, 1990) in contrast to the potential they provide. In one example, Rojo López et al. (2021) showed that phonological variables showing speech rhythm (e.g., the average duration of syllabic intervals) in consecutive interpreting output are correlated with self-reported state anxiety as opposed to trait anxiety but only when interpreting from L1 to L2, suggesting that acoustic measures can reflect emotional states in interpreting.

Indeed, as Scherer (1986) pointed out the vocal-auditory mode of communication is crucial for expressing and conveying emotional and motivational states. Consequently, vocalisations and vocal characteristics have great potential as a means of identifying positive and negative emotional states in interpreting (see also Pisanski & Sorokowski, 2021; Rothkrantz et al., 2004; Sondhi et al., 2015; van den Broek, 2004).

Current Study

The current study aims to investigate the effect of bilingual processing of emotional language on the subjective affect and prosodic markers of emotionality. SI was chosen to capture bilingual processing as an extreme case with very high bilingual processing demands (Hervais-Adelman & Babcock, 2020; Van Der Linden et al., 2018). The study partially follows the procedures in Korpal and Jasielska (2019) and Rojo López et al. (2021) and employs both subjective (i.e., positive and negative affect scores) and objective measures of emotionality response (i.e., prosodic parameters based on pitch, intensity, and fluency), as detailed below. In line with this, we have two main research questions:

(1) Does the simultaneous interpreting of emotional speeches of different valence modulate subjective affect? As discussed above, there is evidence that simultaneous interpretation of negative speeches of low valence leads to an increase in negative affect scores on the PANAS compared to baseline (Korpal & Jasielska, 2019). Thus, we predict that the effect in the current study will be in the same direction, which would replicate the findings. More importantly, however, as the authors themselves point out, it is currently unclear whether the simultaneous interpretation of a positive speech also increases positive affect.

(2) Does the simultaneous interpreting of emotional speeches of different valence modulate prosodic parameters in the auditory output? Emotionality and stress have reliable effects on auditory parameters as discussed above. Overall, the results show that sadness associated with negative valence decreases F0 and intensity but leads to more frequent pauses, and happiness/joy leads to an opposite pattern. With this in mind, we predict that participants will produce higher intensity, higher F0 and lower pause percentage when interpreting a speech with positive valence and lower intensity, lower F0 and higher pause percentage when interpreting a negative-valence speech compared to a neutral speech.

Finally, the present study has an additional research question as a follow-up to Kumcu and Öztürk (2023): (3) Do individual differences in mental imagery, short-term memory (STM) or working memory (WM) have a mediating influence on the predictive relationship between emotionality and subjective affect/prosody parameters? We expect that individuals with better (more vivid) mental imagery will report higher affect scores regardless of the valence since better imagery (especially vividness) functions as an "emotional amplifier" and usually leads to better memory and more emotional involvement (e.g., Holmes & Mathews, 2010).

Methods

Participants

The study was carried out with 18 4th-year undergraduate translation and interpreting students at Hacettepe University (8 males; M_{age} = 22.61, *SD* = 0.78, range: 21 – 24). All participants were native speakers of Turkish (speaking/learning only Turkish from birth and currently using Turkish as their primary language) with English as the foreign language. In interpreting terms, participants' A language (L1) was Turkish and B language (L2) was English. The pre-test questionnaire documented that all participants had taken undergraduate-level courses in SI where they had practised SI in both directions to the same extent, the main task in the current study. The mean current grade point average (CGPA) of the sample was 3.42 out of 4.00 (*SD* = 0.27, range = 2.85 – 3.76). Participants reported normal or corrected-to-normal vision, no voice, speech or hearing difficulties and no history of any neurological disorder. All participants were fully informed about the details of the procedure and gave online consent. Post-test debriefing revealed that all participants were naïve to the purpose of the study. Larger sample sizes are preferable to increase statistical power and generalizability. However, the present study was conducted with a relatively small number of participants because it focused on a highly specialised population within the general pool of bilinguals, i.e. (student) interpreters with a certain level of experience and expertise in simultaneous interpreting, while coming from similar demographic, linguistic and academic backgrounds, in order to control for extraneous variables. Nevertheless, it is important to emphasise that follow-up studies using the same paradigm but with larger sample sizes are highly recommended for consistent and comparable results.

Stimuli

Three speeches in Turkish (emotionally negative, emotionally neutral, and emotionally positive) were selected to be simultaneously interpreted (see Table 1). The speeches were developed based on the normed speech examples in Korpal and Jasielska (2019). Accordingly, the negative-valence speech was about the experiences of a mother who lost her daughter following an assault and the neutral speech was on telephone etiquette as in Korpal and Jasielska (2019). In addition, a positive-valence speech on memories of a wedding day was included. Speeches were recorded in a sound-attenuated room by a female native speaker of

Turkish. Audio files and full transcripts of the speeches can be accessed at https://osf.io/4ka8p.

Speeches were internally validated on emotionality, valence, imageability, familiarity, difficulty, and speed by all participants of the study on a scale from 1 (not emotional, very negative, not imageable, not familiar, very easy, and very slow) to 5 (very emotional, very positive, highly imageable, very familiar, very difficult, and very fast). Emotionality, valence, imageability, and difficulty were also externally validated by 18 language specialists on the same scale as well. There were reliable differences between the valence and emotionality of positive, negative and neutral speeches in the expected directions based on both internal and external assessments (all *ps* $< .05$) (see Table 1). There was also a significant difference in difficulty as per participant evaluations; $F(2,51) = 3.91$, $p = .02$. The pairwise comparison indicated that neutral speech ($M = 2.83$, $SD = 0.99$) was evaluated as more difficult (to interpret simultaneously) than the negative speech ($M = 2$, $SD = 0.77$); $p = .02$. Further, the Turkish adaptation of Flesch readability index analysis (Ateşman, 1997) based on word and sentence lengths has indicated that neutral speech is in the middle difficulty category (*Readability score* = 60.3), while positive (*Readability score.* = 70.4) and negative (*Readability score* = 79.9) speeches are in the easy category. There was no difference in difficulty as per external specialist evaluations; $F(2,51) = 2.05$, $p = .14$.

Table 1

Descriptive Characteristics (Means and SDs) of Source Speeches as Rated by the Participants (I) (n = 18) and External Assessors (E) (n = 18)

Measure	Positive speech	Neutral speech	Negative speech	*F*	*P*
Overall					
Number of words	406	404	409		
Total duration (s.)	220.16	238.26	236.44		
Participants					
Emotionality (I)	4.33 (0.59)	1.67 (0.97)	**4.67 (0.59)**	88.64	< .0001
Valence (I)	**4.78 (0.43)**	2.72 (0.57)	1.17 (0.38)	268.4	< .0001
Imageability (I)	**4.39 (0.61)**	3.06 (0.94)	3.61 (1.09)	9.92	< .001
Difficulty (I)	2.39 (0.92)	**2.83 (0.99)**	2.00 (0.77)	3.91	.02
Self-performance	2.61 (0.85)	2.44 (0.98)	3.11 (0.96)	2.48	.09
Familiarity	3.56 (1.38)	4.11 (1.02)	3.11 (1.53)	2.56	.09
Speed	2.94 (0.80)	2.94 (0.87)	2.56 (0.92)	1.21	.31
External assessors					
Emotionality (E)	4.22 (0.81)	1.56 (0.78)	**4.78 (0.43)**	110.4	< .0001
Valence (E)	**4.67 (0.49)**	3.11 (0.76)	1.50 (0.99)	76.02	< .0001
Imageability (E)	**4.50 (0.62)**	3.50 (1.04)	3.67 (1.08)	5.86	.005
Difficulty (E)	2.56 (0.70)	2.17 (0.86)	2.67 (0.77)	2.05	.14

Note. The highest values in the measures with significant differences are in bold.

Procedure

The main task of the study was SI from L1 (Turkish) to L2 (English). As spelt out below, the study was composed of five consecutive phases: (1) the baseline, (2) the pre-test questionnaire, (3) simultaneous interpreting (SI), (4) mental imagery and memory tests, and (5) the post-test questionnaire.

(1) The baseline: Two weeks before the interpreting task, participants were asked to describe a black-and-white drawing of a street, which formed the baseline/control language data for comparison with the SI output data. Participants were also asked to complete the PANAS to describe their baseline emotional state using 20 adjectives on a five-point Likert scale (Watson et al., 1988).

(2) Pre-test questionnaire: Participants were asked 12 questions about their demographics (age and gender) and academic background (university, year, CGPA, courses taken and interpreting experience).

(3) Simultaneous interpreting (SI): Participants were asked to simultaneously interpret three speeches varying in emotionality and valence from L1 (Turkish) into L2 (English). Participants received no prior information about the speeches and their terminological background. The speeches were presented in counterbalanced order. 1/3 of the participants started with negative, positive and neutral speeches. The interpretation was recorded. Participants were asked to complete the PANAS immediately after each SI task. The auditory output was analysed as described in 3.1. below.

(4) Mental imagery and memory tests: Participants were also asked to self-report their mental imagery abilities on five scales (see Kumcu & Öztürk, 2023 for further details), and they were tested on forward and backward digit span tests corresponding to STM and WM, respectively, to determine whether there were any imagery or memory differences between participants that would predict their emotional response in the study. The memory tests were developed on PsychoPy (Peirce et al., 2019) and administered on Pavlovia.org. Accuracy and reaction times in the memory tests were used as dependent measures.

(5) Post-test questionnaire: Participants were administered a questionnaire in which they evaluated the perceived emotionality, valence, imageability, familiarity, difficulty, and speed of the source speech and also, their SI performance. They were also asked to qualitatively comment on their output.

Results

Acoustic Analysis of Auditory Output

Participants' auditory output was cleaned for artefacts (i.e., coughing or throat clearing) and analysed using Praat version 6.3.09 (Boersma & Weenink, 2022) for the acoustic markers of emotional states (Sondhi et al., 2015): pitch as a function of fundamental frequency (F0), intensity, and percentage of pause duration (%) as a function of fluency. F0 range was defined as 75-300 Hz for males and 75-500 Hz for females. The percentage of pause duration was computed as the ratio of the duration of silent pauses to the total duration of interpretations. The *mark pauses* script (Lennes, 2002) was used for the analysis of pause duration. The minimum pause duration was set at 0.5 seconds, the minimum pitch at 100 Hz, and the maximum intensity at 59 dB.

Mixed-Effects Modelling

Data were analysed using linear logit mixed-effects modelling to reveal the predictive factors of SI performance. The contribution of a fixed effect was investigated by comparing a full model containing the effect in question against a reduced model in which only that effect was removed, or a null model without any fixed effects (Winter, 2013). Data were analysed and visualised in the R programming language and environment (R Core Team, 2023). Raw and processed data files and R scripts used to run the analyses can be accessed at https://osf.io/4ka8p. Mixed-effects models were constructed with the *lme4* package (Bates et al., 2015). Significance values of the coefficients in models were computed based on the t-distribution using the Satterthwaite approximation with the *lmerTest* package (Kuznetsova et al., 2017). Multiple comparisons between factors were performed with *multcomp* package (Hothorn et al., 2008) and based on Tukey's Test.

The Effect of Interpreting Emotional Speeches on Subjective Affect

Positive affect

Mixed-effects models were fit with the speech type (baseline, negative speech, neutral speech, and positive speech) as a fixed effect, participants as a random effect and positive affect score as the target variable. The results showed that the speech type significantly improved the null positive affect model; $\chi2(3) = 12.01$, $p = .007$. Participants reported significantly higher positive affect scores at the baseline (M = 28.89, SD = 7.73) as compared to the neutral speech (M = 21.72, SD = 6.29); B = 7.17, t = 3.44, p = .001 and the negative speech (M = 23.61, SD = 5.77); B = 5.28, t = 2.54, p = .01. There was not a significant difference in positive affect between the baseline and the positive speech (M = 26.17, SD = 9.19); p = .2.

When the baseline has been dropped, the speech type did not improve the null positive affect model; $\chi2(2) = 4.66$, $p = .1$, showing that there was not a significant difference between the positive affect score following SI of the positive speech and positive affect score following SI of the neutral speech; p = .35 or negative speech; p = .21 (see Figure 1).

Models with internal and external speech properties, mental imagery scores and memory measures as fixed effects were fit. Results showed that interpreting highly imageable speeches according to both internal; B = 3.33, t = 2.23, p = .03 and external assessments; B = 4, t = 2.13, p = .04 resulted in more positive affect scores. Emotionality, valence or difficulty did not predict positive affect scores. Mental imagery scores or memory performance did not predict positive affect either (all ps > .05).

Negative affect

Similar models were fit with the negative affect score as the target variable. The speech type improved the null negative affect model; $\chi2(3) = 36.43$, $p < .0001$. Participants reported significantly higher negative affect score at the baseline (M = 26.06, SD = 9.97) as compared to the neutral speech (M = 14.28, SD = 7.12); B = 11.78, t = 4.63, p < .0001 and the positive speech (M = 15.67, SD = 6.7); B = 10.39, t = 4.09, p = .0002. There was not a significant difference in negative affect between the baseline and the negative speech (M = 29, SD = 8.53); p = .25.

When the baseline has been dropped, the speech type improved the null negative affect model as well; $\chi2(2) = 34.92$, $p < .0001$. Participants reported significantly higher negative affect scores following SI of the negative speech when compared against both the neutral; B = 14.72, t = 6.84, p < .0001 and the positive speech; B = 13.33, t = 6.20, p < .0001 (see Figure 1).

Interpreting speeches that are evaluated as more emotional; $B = 3.22$, $t = 3.6$, $p = .0007$, more negative; $B = -3.46$, $t = 4.55$, $p < .0001$, less familiar; $B = -14.22$, $t = -5.66$, $p < .0001$, that are perceived more slower; $B = -36.07$, $t = -7.48$, $p < .0001$ and that are associated with better performance; $B = 23.14$, $t = 7.32$, $p < .0001$, that are rated easier by the participants; $B = -17.41$, $t = -5.83$, $p < .0001$ resulted in higher negative affect scores. Mental imagery scores or memory performance did not predict negative affect (all *ps* > .05).

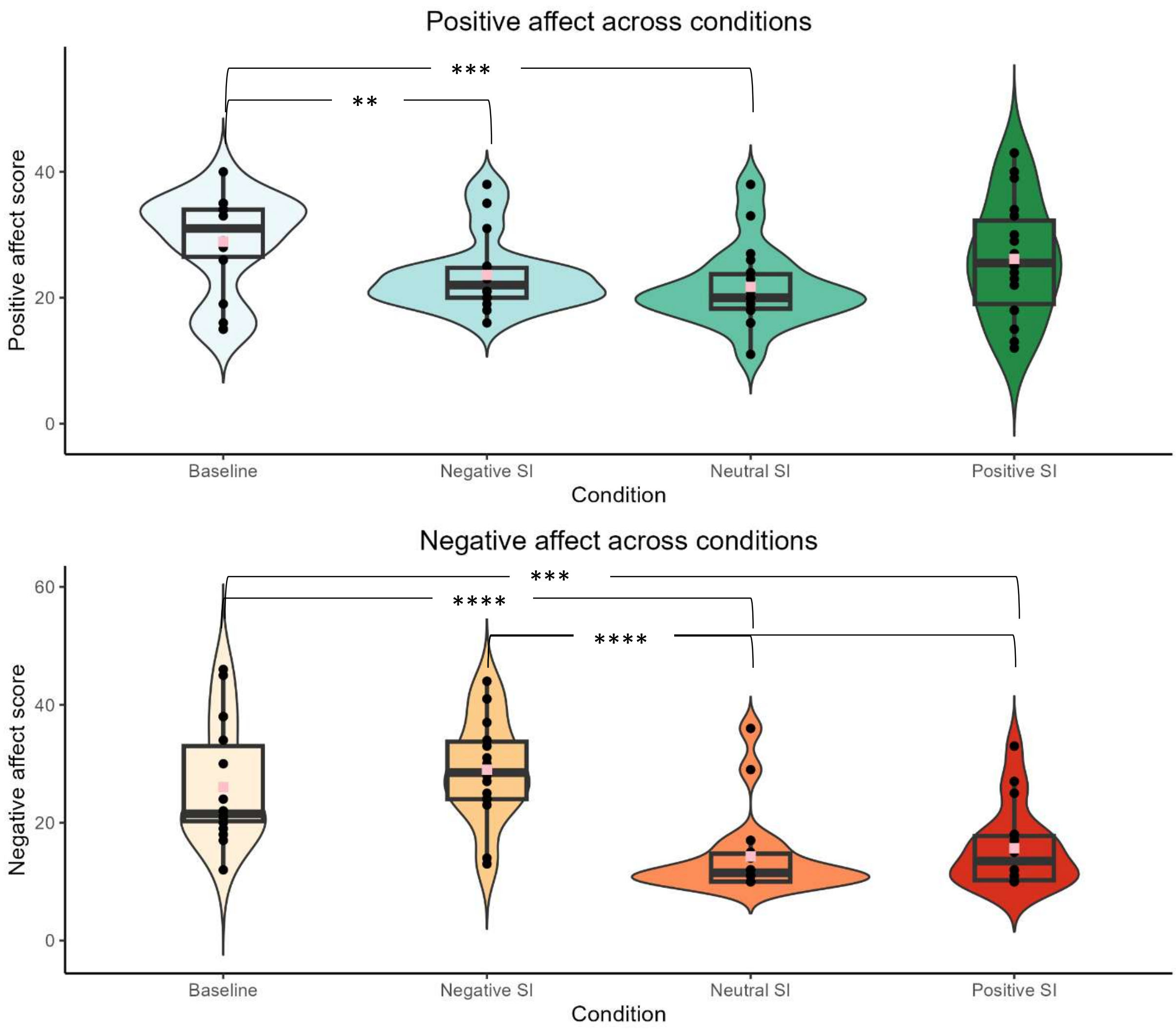

Figure 1. Violin plots show the distribution of positive and negative affect scores at the baseline and immediately after SI tasks. The box plots within the violins show the whole range (vertical line), the mean (pink square), and the median (horizontal line). Black dots represent data points for individual participants.

*$p \leq .05$, **$p \leq .01$, ***$p \leq .001$., ****$p \leq .0001$

The Effect of Interpreting Emotional Speeches on Prosodic Measures

Intensity

Results from the prosodic analysis of the baseline and SI outputs were presented in Table 2. Mixed-effects models were fit with the speech type (baseline, negative speech, neutral speech, and positive speech) as a fixed effect, participants as a random effect and vocal intensity (dB) as the target variable. The results showed that the speech type significantly improved the null intensity model; $\chi2(3) = 23.21$, $p < .0001$. Overall, the vocal intensity was higher in all SI conditions as compared to the baseline (M = 64.52, SD = 5.96), namely, neutral speech (M = 68.45, SD = 3.7); $B = 3.93$, $t = 4.67$, $p < .0001$, positive speech (M = 68.4, SD = 4.2); $B = 3.88$, $t = 4.62$, $p < .0001$, and negative speech (M = 67.43, SD = 3.63); $B = 2.91$, $t = 3.46$, $p = .001$. When the baseline has been dropped, the speech type did not improve the null intensity model; $\chi2(2) = 5.06$, $p = .08$. Post-hoc analyses showed that although the vocal intensity was lower during the interpretation of negative speech as compared to the neutral speech, the difference was not significant; $B = -1.02$, z = -2.06, $p = .1$. Vocal intensity during the interpretation of positive speech was not higher than that of the neutral speech either; $B = 0.04$, $t = 0.09$, $p = .1$ (see Figure 2).

Models with speech properties, mental imagery scores and memory measures showed that higher difficulty (internal); $B = 1.2$, $t = 1.98$, $p = .05$; lower self-performance; $B = -1.62$, $t = -2.27$, $p = .03$, and higher speed; $B = 2.56$, $t = 2.33$, $p = .03$ resulted in a higher voice intensity. Valence did not predict intensity ($p > .05$). Mental imagery scores or memory performance did not predict intensity either (all ps > .05).

Pitch

Mixed-effects models were also fit with the speech type (baseline, negative speech, neutral speech, and positive speech) as a fixed effect, participants as a random effect and F0 as a function of the pitch as the target variable. Following the tradition in prosody research (e.g., Alku & Vilkman, 1996; Tolkmitt & Scherer, 1986; Weeks et al., 2012), the pitch data were analysed by sex for a more precise analysis.

The results showed that the speech type did not improve the general null pitch model either among females; $\chi2(3) = 3.38$, $p = .34$ or males; $\chi2(3) = 5.93$, $p = .12$. When the baseline has been dropped, the speech type improved the null pitch model among males; $\chi2(2) = 6.14$, $p = .05$ but not among females; $\chi2(2) = 0.9$, $p = .63$. Post-hoc analyses in male data showed the F0 was significantly lower during the interpretation of negative speech (M = 119.42, SD = 10.58) as compared to the neutral speech (M = 124.28, SD = 11.27); $B = -4.86$, $z = -2.73$, $p = .02$. F0 during the interpretation of positive speech (M = 121.49, SD = 12.52) was not higher than that of the neutral speech; $B = 2.79$, $z = 1.57$, $p = .26$ (see Figure 2).

Models with speech properties, mental imagery scores and memory measures showed that valence did not predict pitch; $p = .43$. However, higher difficulty (internal) resulted in a higher pitch among males; $B = 8.86$, $t = 2.29$, $p = .04$. Mental imagery scores or memory performance did not predict pitch (all ps > .05).

Fluency

Mixed-effects models were fitted with the speech type in the SI task (negative speech, neutral speech, and positive speech) as a fixed effect, participants as a random effect and percentage of pause duration (ratio of pause duration to total duration) as a function of fluency as the target variable. The results showed that the speech type significantly improved the null fluency model without the baseline condition; $\chi2(2) = 12.68$, $p = .002$. Post-hoc tests showed that the percentage of pause duration during the interpretation of negative speech (M = 35.14, SD = 10.28) was significantly higher than the percentage of pause duration during the interpretation of neutral speech (M = 30.68, SD = 10.16); $B = 4.46$, $t = 2.78$, $p = .009$ and also of positive speech (M = 29.11, SD = 12.1); $B = 6.03$, $t = 3.76$, $p = .0006$. Percentage of pause duration in positive speech was not significantly lower than neutral speech; $p = .6$. In line with this, lower valence; $B = -1.62$, $t = 3.58$, $p = .001$ but also lower difficulty (internal); $B = -5.17$, $t = 2.42$, $p = .02$ predicted higher pause percentage. Mental imagery scores or memory performance did not predict pause percentage (all ps > .05).

Table 2

Self-Reported Affect Scores and Acoustic Measures at the Baseline and across Task Conditions (Means and SDs)

Measure	Baseline	SI Positive speech	SI Negative speech	SI Neutral speech
Positive affect score	28.89 (7.73)	26.17 (9.19)	23.61 (5.77)	21.72 (6.29)
Negative affect score	26.06 (9.97)	15.67 (6.70)	29.00 (8.53)	14.28 (7.12)
Total duration (s.)	154.74 (46.59)	227.12 (6.81)	232.59 (34.49)	245.78 (9.70)
Mean F0 (Hz.)	165.37 (47.32)	169.63 (48.18)	168.44 (49.23)	171.82 (49.06)
Mean intensity (dB)	64.52 (5.96)	68.40 (4.20)	67.43 (3.63)	68.45 (3.70)
Pause duration (s.)	65.00 (40.43)	65.90 (26.85)	81.62 (27.09)	75.28 (24.55)
Pause percentage (%)	40.16 (19.47)	29.11 (12.10)	35.14 (10.28)	30.68 (10.16)

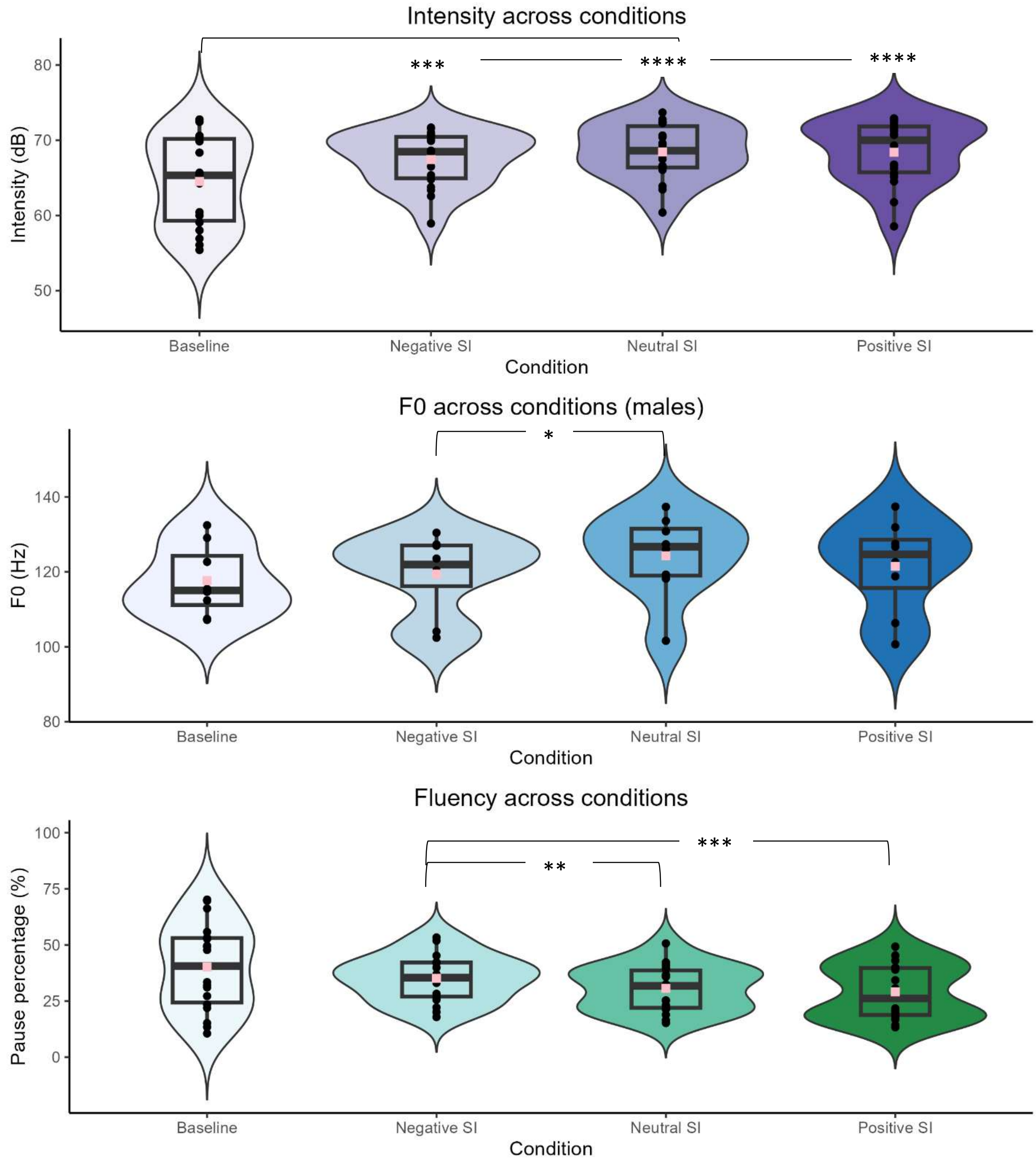

Figure 2. Violin plots show the distribution of intensity, F0 and pause percentage at the baseline and immediately after SI tasks. The box plots within the violins show the whole range (vertical line), the mean (pink square), and the median (horizontal line). Black dots represent data points for individual participants.

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$., **** $p \leq .0001$

Discussion

The current study investigated the predictive effect of simultaneously interpreting speeches varying in emotional valence on self-reported positive and negative affect and three prosodic measures (i.e., intensity, F0, and pause duration) to shed light on the emotional language processing in extreme bilingualism. Overall, our results demonstrated that processing emotional language in a performative language task with very high language control demands (Hervais-Adelman & Babcock, 2020; Van Der Linden et al., 2018) has consequences on both subjective affect and acoustic measures of emotional state. Below we summarise the results as per the research questions and discuss them in light of previous research.

(1) Does the simultaneous interpreting of emotional speeches of different valence modulate subjective affect? Subjective affect scores collected via the PANAS showed that interpreting an emotional speech with a negative valence, which can be considered as "sad", amplified the negative psychological state of the participants in the current study. Such an effect was previously reported by Korpal and Jasielska (2019). In the current study, the findings were replicated, suggesting the reliability of the “negative speech effect”. More importantly, our results have provided an answer to an open question: Does interpreting a so-called "happy" speech also increase positive affect? Our results answer this question in the negative. Positive affect scores did not increase after interpreting a positive speech. In conclusion, the present study shows that SI modulates self-reported negative emotional state, but not positive emotional state. Interestingly, participants' positive and negative mood scores were higher at baseline than when interpreting a neutral speech, suggesting that SI without emotional content somehow dampened the emotional response in the present study.

To explain our findings in subjective affect, we can turn to emotion research, where there is ample evidence for a “negativity bias” in affect, proposing that negative emotions are more salient than positive emotions and that negative experiences are remembered better than positive experiences, even when they are equally intense (Kensinger, 2009; Vaish et al., 2008). More importantly, there is also evidence that the negativity bias is more prevalent in younger adults and that cognitive load enhances the effect (Carstensen & DeLiema, 2018), which applies to the task conditions and the sample in the current study. One finding that requires further investigation is the reason why the imageability of the source speech predicted positive affect but not negative affect. Our findings as to the first research question have important implications for interpreting as well considering that emotional stability largely modulates performance (Albl-Mikasa, 2014; Bontempo & Napier, 2011) and output quality deteriorates as a result of stress in SI (Mackintosh, 2003). Further, community interpreters typically working in highly emotional settings are vulnerable to developing vicarious trauma due to prolonged emotional fatigue (Splevins et al., 2010).

(2) Does the simultaneous interpreting of emotional speeches of different valence modulate prosodic parameters in the auditory output? First, the results showed that participants’ vocal output exhibited higher intensity in all SI conditions compared to baseline. However, there was no difference in intensity according to the emotional valence of the source speech. Although mean intensity was lower when interpreting negative speech than when interpreting neutral speech, the difference was not significant. Furthermore, while higher difficulty and lower self-performance predicted higher intensity, valence did not. Taken together, these results suggest that intensity in SI appears to be modulated by cognitive rather than emotional load. On the other hand, male participants produced a lower F0 when interpreting the negative speech compared to the neutral speech. That said, higher difficulty

rather than higher valence predicted a higher pitch in males, as reported in intensity. A higher F0 was not documented when interpreting positive speech compared to neutral speech. Finally, the percentage of pauses during negative speech was higher for both neutral and positive speech, but the same difference was not found between positive and neutral speech. In contrast to intensity and pitch, lower valence, together with lower difficulty, predicted a higher percentage of pauses.

Overall, the prosody results suggest that the percentage of pause duration, in particular, and pitch as a function of F0 are modulated by the emotionality of source speech to a larger extent. The direction of the effect is in line with the literature discussed above. It is important to note that although males produced lower F0 during the negative speech as to the neutral speech, speech difficulty still predicted F0. This is an important area of research and future studies can be developed on a 2 by 2 factorial design to test the effect of both emotional load and cognitive load in SI. The fact that the effect of pitch was reported only in males complies with the previous results showing sex-related differences in findings based on pitch (e.g., Weeks et al., 2012). A more detailed analysis showed that as highly expected, F0 among females (M = 208.55, SD = 25.13) was significantly higher than F0 among males (M = 121.73, SD = 11.16) across all SI conditions; B = 86.81, t = 9.44, $p < .0001$ due to anatomical/physiological and even sociocultural reasons (Van Bezooijen, 1995). However, the difference between males and females in mean F0 was larger in negative speech (M_{diff} = 88.22) as compared to neutral speech (M_{diff} = 85.57) suggesting that it was the even lower pitch during the interpretation of negative speech among males which led to a significant difference. In any case, the results of the current study should be viewed and interpreted with caution due to the limited sample size, which is a common problem in experiments testing either trainee or professional interpreters (Liu, 2011).

Finally, our results showed that individual differences in mental imagery, STM or WM did not play a mediating role in the relationship between emotional language processing and subjective affect/prosody parameters. Our previous results (Kumcu & Öztürk, 2023) showed that mental imagery control is a crucial skill for consecutive interpreting accuracy. In this regard, future studies could investigate the relationship between mental imagery and emotional processing on performance metrics.

In conclusion, the current study extends previous findings on emotional language processing in bilingualism to an underrepresented language pair (Turkish-English) and to an underrepresented task (SI). We show that competent bilinguals transfer mostly negative emotions processed in the L1 to the L2. In other words, interpreters not only imitate the speaker's discourse intentions (Amos & Pickering, 2020) but also her/his emotions to a certain extent. The study also shows that certain vocal cues can be considered as potential indicators of emotional (and also non-emotional) processing in interpreting, which calls for further research.

References

Albuquerque, L., Valente, A. R. S., Teixeira, A., Figueiredo, D., Sa-Couto, P., & Oliveira, C. (2021). Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan. *PLOS ONE*, *16*(4), e0248842. https://doi.org/10.1371/journal.pone.0248842

Alku, P., & Vilkman, E. (1996). A comparison of glottal voice source quantification parameters in breathy, normal and pressed phonation of female and male speakers. *Folia Phoniatrica et Logopaedica*, *48*(5), 240-254. https://doi.org/10.1159/000266415

Ateşman, E. (1997). Türkçe'de okunabilirliğin ölçülmesi. *Dil Dergisi*, *58*, 71-74.

Bates, D. M., Mächler, M., Bolker, B., Walker, S., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.3.03) [Computer software]. http://www.praat.org/

Bohn-Gettler, C. M. (2019). Getting a grip: The PET framework for studying how reader emotions influence comprehension. *Discourse Processes*, *56*(5-6), 386-401. https://doi.org/10.1080/0163853X.2019.1611174

Carstensen, L. L., & DeLiema, M. (2018). The positivity effect: A negativity bias in youth fades with age. *Current Opinion in Behavioral Sciences*, *19*, 7-12. https://doi.org/10.1016/j.cobeha.2017.07.009

Conrad, M., Recio, G., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: A cross cultural ERP study with German-Spanish bilinguals. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00351

Dong, Y., & Li, P. (2020). Interpreting: A window into bilingual processing. *Bilingualism: Language and Cognition*, *23*(4), 703-705. doi.org/10.1017/S1366728920000346

Filippa, M., Lima, D., Grandjean, A., Labbé, C., Coll, S. Y., Gentaz, E., & Grandjean, D. M. (2022). Emotional prosody recognition enhances and progressively complexifies from childhood to adolescence. *Scientific Reports*, *12*(1), 17144. https://doi.org/10.1038/s41598-022-21554-0

Griffin, G. R., & Williams, C. E. (1987). The effects of different levels of task complexity on three vocal measures. *Aviation, Space, and Environmental Medicine*, *58*(12), 1165-1170.

Gumul, E. (2021). Reporting stress in simultaneous interpreting. The analysis of trainee interpreters' retrospective reports and outputs. *Onomázein Revista de Lingüística Filología y Traducción*, *8*. https://doi.org/10.7764/onomazein.ne8.04

Habermas, T. (2018). *Emotion and narrative: Perspectives in autobiographical storytelling* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781139424615

Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development*, *25*(2-3), 223-247. https://doi.org/10.1080/01434630408666530

Harris, C. L., Ayçiçegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*(04), 561-579. https://doi.org/10.1017/S0142716403000286

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review*, *14*(3), 436-441. https://doi.org/10.3758/BF03194085

Hervais-Adelman, A., & Babcock, L. (2020). The neurobiology of simultaneous interpreting: Where extreme language control and cognitive control intersect. *Bilingualism: Language and Cognition*, *23*(4), 740-751. https://doi.org/10.1017/S1366728919000324

Heyrani, R., Nejati, V., Abbasi, S., & Hartwigsen, G. (2022). Laterality in emotional language processing in first and second language. *Frontiers in Psychology*, *12*, 736359. https://doi.org/10.3389/fpsyg.2021.736359

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349-362. https://doi.org/10.1016/j.cpr.2010.01.001

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346-363. https://doi.org/10.1002/bimj.200810425

Huttunen, K., Keränen, H., Väyrynen, E., Pääkkönen, R., & Leino, T. (2011). Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. *Applied Ergonomics*, *42*(2), 348-357. https://doi.org/10.1016/j.apergo.2010.08.005

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*(1), 20-31. https://doi.org/10.1037/0022-3514.45.1.20

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770-814. https://doi.org/10.1037/0033-2909.129.5.770

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1*(2), 99-113. https://doi.org/10.1177/1754073908100432

Klonowicz, T. (1990). A psychophysiological assessment of simultaneous interpreting: The interaction of individual differences and mental workload. *Polish Psychological Bulletin*, *21*, 37-48.

Korpal, P. (2016). Interpreting as a stressful activity: Physiological measures of stress in simultaneous interpreting. *Poznan Studies in Contemporary Linguistics*, *52*(2). https://doi.org/10.1515/psicl-2016-0011

Korpal, P. (2021). Stress and emotion in conference interpreting. In M. Albl-Mikasa, & E. Tiselius (Eds.), *The Routledge handbook of conference interpreting* (pp. 401-413). Routledge.

Korpal, P., & Jasielska, A. (2019). Investigating interpreters' empathy: Are emotions in simultaneous interpreting contagious? *Target*, *31*(1), 2-24. https://doi.org/10.1075/target.17123.kor

Kumcu, A. (2020). Konferans salonunda dağıtık biliş: Andaş çeviri sürecine yeni bir bakış. *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, *37*(1), 170-185. https://doi.org/10.32600/huefd.621553

Kumcu, A., & Öztürk, A. (2023). Visual mental imagery and verbal working memory: Evidence from consecutive interpreting. *Journal of Cognitive Psychology*, 1-16. https://doi.org/10.1080/20445911.2023.2216917

Kurz, I. (2002). Physiological stress responses during media and conference interpreting. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century: Challenges and opportunities* (pp. 195–202). John Benjamins Publishing Company.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Lennes, M. (2002). *Mark_pauses* [Computer software]. http://phonetics.linguistics.ucla.edu/facilities/acoustic/mark_pauses.txt

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00444

Liu, M. (2011). Methodology in interpreting studies: A methodological review of evidence-based research. In B. Nicodemus & L. Swabey (Eds.), *Advances in interpreting research: Inquiry in action* (pp. 85–120). John Benjamins Publishing Company. https://doi.org/10.1075/btl.99.08liu

Mackintosh, J. (2003). The AIIC workload study. *FORUM. Revue Internationale d'interprétation et de Traduction / International Journal of Interpretation and Translation*, *1*(2), 189-214. https://doi.org/10.1075/forum.1.2.09mac

Martínez-Sánchez, F., Muela-Martínez, J. A., Cortés-Soto, P., García Meilán, J. J., Vera Ferrándiz, J. A., Egea Caparrós, A., & Pujante Valverde, I. M. (2015). Can the acoustic analysis of expressive prosody discriminate schizophrenia? *The Spanish Journal of Psychology*, *18*, E86. https://doi.org/10.1017/sjp.2015.85

Mellinger, C. D., & Hanson, T. A. (2019). Meta-analyses of simultaneous interpreting and working memory. *Interpreting. International Journal of Research and Practice in Interpreting*, *21*(2), 165-195. https://doi.org/10.1075/intp.00026.mel

Moser-Mercer, B., Künzli, A., & Korac, M. (1998). Prolonged turns in interpreting: Effects on quality, physiological and psychological stress (Pilot study). *Interpreting. International Journal of Research and Practice in Interpreting*, *3*(1), 47-64. https://doi.org/10.1075/intp.3.1.03mos

Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralts, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics*, *20*(1), 50-64. https://doi.org/10.1016/j.jneuroling.2006.04.001

Nour, S., Struys, E., & Stengers, H. (2020). Adaptive control in interpreters: Assessing the impact of training and experience on working memory. *Bilingualism: Language and Cognition*, *23*(4), 772-779. https://doi.org/10.1017/S1366728920000127

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology*, *47*(6), 405-428. https://doi.org/10.1080/00207594.2012.743665

Pisanski, K., & Sorokowski, P. (2021). Human stress detection: Cortisol levels in stressed speakers predict voice-based judgments of stress. *Perception*, *50*(1), 80-87. https://doi.org/10.1177/0301006620978378

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/.

Rojo López, A. M., Foulquié-Rubio, A.-I., Espín López, L., & Martínez Sánchez, F. (2021). Analysis of speech rhythm and heart rate as indicators of stress on student interpreters. *Perspectives*, *29*(4), 591–607. https://doi.org/10.1080/0907676X.2021.1900305

Rothkrantz, L. J. M., Wiggers, P., Wees, J.-W. A., & Vark, R. J. (2004). Voice stress analysis. In P. Sojka, I. Kopeček, & K. Pala (Eds.), *Text, speech and dialogue* (pp. 449-456). https://doi.org/10.1007/978-3-540-30120-2_57

Scherer, K. R. (1986a). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*(2), 143-165. https://doi.org/10.1037/0033-2909.99.2.143

Scherer, K. R. (1986b). Voice, stress, and emotion. In M. H. Appley & R. Trumbull (Eds.), *Dynamics of stress* (pp. 157-179). Springer US. https://doi.org/10.1007/978-1-4684-5122-1_9

Sondhi, S., Khan, M., Vijay, R., Salhan, A. K., & Chouhan, S. (2015). Acoustic analysis of speech under stress. *International Journal of Bioinformatics Research and Applications*, *11*(5), 417. https://doi.org/10.1504/IJBRA.2015.071942

Splevins, K. A., Cohen, K., Joseph, S., Murray, C., & Bowley, J. (2010). Vicarious posttraumatic growth among interpreters. *Qualitative Health Research*, *20*(12), 1705-1716. https://doi.org/10.1177/1049732310377457

Tamuri, K. (2014). Fundamental frequency in Estonian emotional read-out speech. *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, *6*(1), 9-21. https://doi.org/10.12697/jeful.2015.6.1.01

Tisljár-Szabó, E., & Pléh, C. (2014). Ascribing emotions depending on pause length in native and foreign language speech. *Speech Communication*, *56*, 35-48. https://doi.org/10.1016/j.specom.2013.07.009

Tolkmitt, F. J., & Scherer, K. R. (1986). Effect of experimentally induced stress on vocal parameters. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 302-313. https://doi.org/10.1037/0096-1523.12.3.302

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383-403. https://doi.org/10.1037/0033-2909.134.3.383

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, *38*(3), 253-265. https://doi.org/10.1177/002383099503800303

van den Broek, E. L. (2004). Emotional Prosody Measurement (EPM): A voice-based evaluation method for psychological therapy effectiveness. *Studies in Health Technology and Informatics*, *103*, 118-125.

Van Der Linden, L., Van De Putte, E., Woumans, E., Duyck, W., & Szmalec, A. (2018). Does extreme language control training improve cognitive control? A comparison of professional interpreters, L2 teachers and monolinguals. *Frontiers in Psychology*, *9*, 1998. https://doi.org/10.3389/fpsyg.2018.01998

Van Puyvelde, M., Neyt, X., McGlone, F., & Pattyn, N. (2018). Voice stress analysis: A new framework for voice and effort in human performance. *Frontiers in Psychology*, *9*, 1994. https://doi.org/10.3389/fpsyg.2018.01994

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063-1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weeks, J. W., Lee, C.-Y., Reilly, A. R., Howell, A. N., France, C., Kowalsky, J. M., & Bush, A. (2012). "The Sound of Fear": Assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. *Journal of Anxiety Disorders*, *26*(8), 811-822. https://doi.org/10.1016/j.janxdis.2012.07.005

Winter, B. (2013). *Linear models and linear mixed effects models in R with linguistic applications*. http://arxiv.org/abs/1308.5499