9th Congress of Baltic States SLTs'

Understanding each other

12-13 April, 2024

TARTU ÜLIKOOL
haridusteaduste instituut

EESTI LOGOPEEDIDE ÜHING

LLA

Latvijas Logopēdu asociācija

SLT
baltic countries congress

EESTI
KLIINILISTE
LOGOPEEDIDE
SELTS

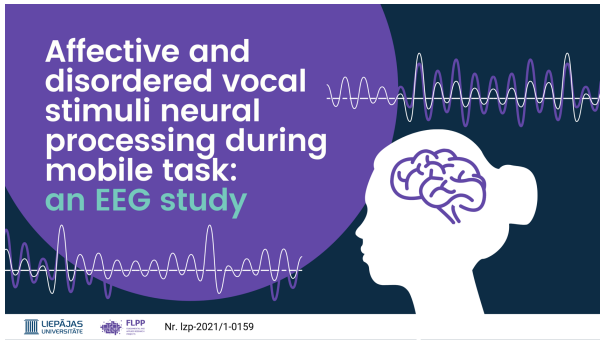
9th Congress of
Baltic States SLTs'
Understanding
each other
12-13 April, 2024

IMPACT OF AGE, GENDER, AND STIMULUS COMPLEXITY ON NEUTRAL VOICE RECOGNITION ACCURACY

Baiba Trinite, Anita Zdanovica, Daiga Kurme, Evija Lavrane, Ilva Magazeina, Anita Jansone

Voice and Speech Research Laboratory

Riga Technical University, Liepaja Academy



The Latvian Science Council funded the study, project No. lzp-2021/1-0159

9th Congress of
Baltic States SLTs'
Understanding each other
12-13 April, 2024

AIM(S) AND METHODS

Aim: The study investigated the impact of listeners' age and gender, as well as the duration (or complexity) of voice samples, on the ability to recognize neutral voice tones when they are played alongside expressive ones.

Methods:

- Thirty-two age-matched listeners listened to 270 semantically neutral voice samples produced in neutral (90), happy (90), and angry (90) intonation by ten professional actors. The participants were required to categorize the auditory stimulus using a customized tablet web interface based on three options.
- The analysis focuses on the recognition accuracy of **neutral voice stimuli** and the factors impacting it. The acoustic analysis determining fundamental frequency (F_0), F_0 standard deviation (SD), sound intensity level (SIL), and SIL SD of neutral voice stimuli was carried out utilizing PRAAT software.

The recording of neutral voice samples. Instructions for actors: Imagine that you are currently not experiencing strong emotions and do not have specific interests or needs. Your task is to repeat the following words/phrases/text in a clear, neutral, and non-emotional voice.

9th Congress of
Baltic States SLTs'
Understanding each other
12-13 April, 2024

INTRODUCTION

- The acoustic properties of a neutral voice include a stable fundamental frequency, reduced frequency range, and less variability in intensity across spoken words (Johnstone & Scherer, 2000).
- These properties contribute to the perception of neutrality, as they lack the inflections and modulations that would otherwise signal particular emotions or communicative intents (Johnstone & Scherer, 2000).
- A neutral voice tone is used in communication for:
  - convey information without bias, maintaining objectivity and credibility;
  - preventing escalating emotions in situations where it is essential to remain calm;
  - reducing the risk of misconstrued messages due to emotional overtones (Scherer, 2003).

INTRODUCTION

- Findings from studies on the effects of age, gender, and stimulus complexity on affective emotion recognition indicate that:
  - The listener and speaker's age may affect the ability to decode vocal affect, with a general decrease in the recognition of vocal emotions as individuals age, often more markedly so for negative emotions (Paulmann, Pell, & Kotz, 2008)
  - vocal affect recognition can be influenced by the gender of both the listener and the speaker (Lausen & Schacht, 2018)
  - Stimulus types and their duration progressively activate emotion-specific knowledge, leading to higher accuracy and confidence ratings. Listening to longer portions of an utterance tends to facilitate the process of explicit recognition and the ability to categorize the meaning of emotional prosody (Pell & Kotz, 2011).

9th Congress of
Baltic States SLTs'
Understanding each other
12-13 April, 2024

RESULTS

Mean fundamental frequency (F_0), F_0 SD, sound intensity level (SIL), SIL SD in words, phrases, and paragraphs produced in neutral voice tone by professional female actors (n = 5)

M_{age} = 24.6 years,
SD = 1.5 years

Parameter	Words (2 syllables) M (SD)	Phrases (6 syllables) M (SD)	Paragraphs (64 syllables) M (SD)	*p*[1]
F_0 (Hz)	201.6 (16.5)	202.5 (15.5)	202.6 (16.9)	.982
F_0 SD (Hz)	25.1 (25.4)	28.3 (14.7)	38.0 (7.4)	.437
SIL (dB)	50.2 (3.4)	47.4 (4.1)	40.6 (1.2)	< .001
SIL SD (dB)	5.2 (2.6)	8.4 (2.6)	15.5 (1.6)	< .001
Speech rate (syllables/sec)	3.6	4.6	4.0	< .001

[1] One-Way ANOVA

9th Congress of
Baltic States SLTs'
Understanding each other
12-13 April, 2024

RESULTS

Mean fundamental frequency (F_0), F_0 SD, sound intensity level (SIL), SIL SD in words, phrases, and paragraphs produced in neutral voice tone by professional male actors (n = 5)

M_{age} = 26.6 years,
SD = 4.9 years

Parameter	Words (2 syllables) M (SD)	Phrases (6 syllables) M (SD)	Paragraphs (64 syllables) M (SD)	*p*[1]
F0 (Hz)	99.2 (4.6)	103.5 (6.2)	118.4 (1.6)	< .001
F0 SD (Hz)	12.0 (3.7)	16.8 (14.6)	43.5 (13.5)	< .001
SIL (dB)	55.0 (4.3)	51.7 (4.0)	46.8 (3.1)	< .001
SIL SD (dB)	6.1 (2.2)	9.9 (1.5)	15.2 (1.6)	< .001
Speech rate (syllables/sec)	3.7	4.3	3.9	.056

[1] One-Way ANOVA

RESULTS

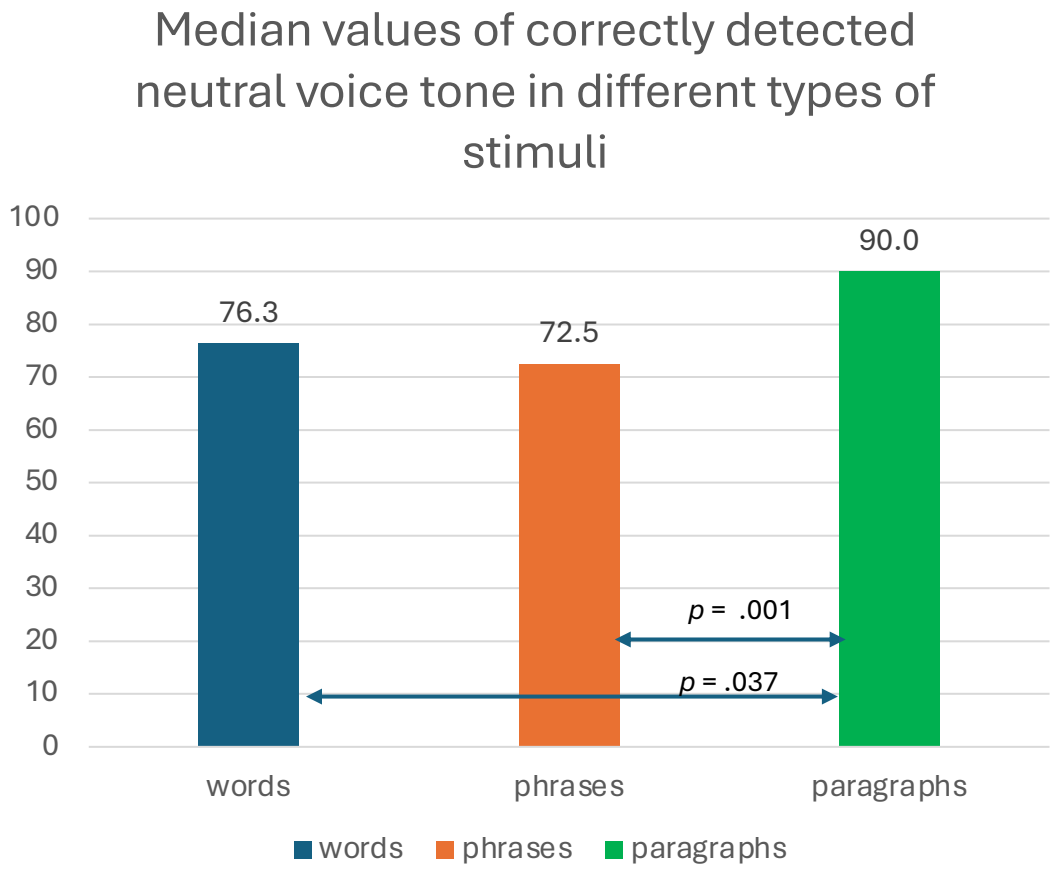
1. Correctly identified were 73.2% of the voice samples expressing a neutral voice tone

Complexity of stimulus	Words (n = 40)	Phrases (n = 40)	Paragraphs (n = 10)
Recognition rate (%)	73.6	71.1	80.3

2. Listener gender differences: (U = 128.5, p = .985; Mann-Whitney U test)
3. Listener age differences: (r_S = .374, p = .035; Spearman Rank correlation)

RESULTS

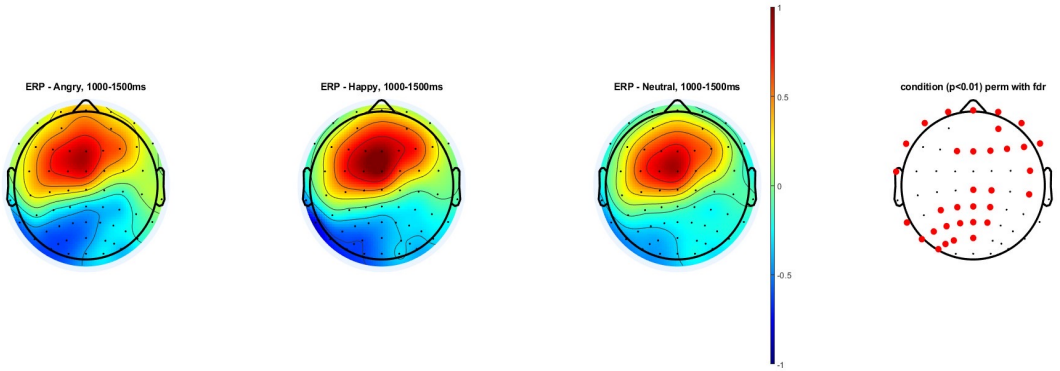
- Recognition accuracy depending on stimulus complexity
  - Words (*Mdn* = 76.3, *IQR* = 31.9)
  - Phrases (*Mdn* = 72.5, *IQR* = 31.9)
  - Paragraphs (*Mdn* = 90.0, *IQR* = 40.0)



9th Congress of
Baltic States SLTs'
Understanding each other
12-13 April, 2024

CONCLUSION

1. Listener age positively correlates with increased accuracy in neutral emotion recognition.
2. Syntactic complexity within spoken utterances influences the discernment of neutral emotions.
3. Longer utterances enhance the recognition of neutral emotional expressions.
4. The upcoming phase of the research will explore the neural underpinnings involved in processing voice stimuli:
   - The brain's response to female and male voices, normophonic and dysphonic voices will be investigated. Electroencephalography (EEG) will measure the neural processing of these varied vocal sounds.



REFERENCES

1. Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed., pp. 220–235). Guilford Press.
2. Lausen, A., & Schacht, A. (2018). Gender differences in the recognition of vocal emotions. *Frontiers in Psychology, 9*(JUN), Article 678.
3. Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language, 104*(3), 262–269.
4. Pell, M. D., & Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE, 6*(11), Article e27256.
5. Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*(1–2), 227–256.